Priming of Syntactic Rules in Task-Oriented Dialogue and Spontaneous Conversation

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

Abstract

Previous work provided corpus evidence for structural priming for specific syntactic constructions. The present paper extends these results by investigating priming effects involving arbitrary syntactic rules in spoken dialogue corpora. We demonstrate the existence of within- and between-speaker priming in both spontaneous conversation (the Switchboard corpus) and task-oriented dialogue (the Map Task corpus). We also find that between-speaker priming is stronger in the Map Task corpus. This supports the hypothesis that in task-oriented dialog, low-level priming is linked to higher-level alignment of situation models.

Keywords: Structural priming; dialogue; task-orientation; language production; language comprehension; speech

Introduction

Priming is a wide-spread phenomenon in both language comprehension and language production. A classical priming phenomenon is that a word (the target) is recognized more quickly and more accurately if it is semantically similar to a preceding word (the prime). Similar priming effects have also been demonstrated for syntactic constructions (Bock, 1986; Branigan et al., 2000). Here, the key finding is that speakers tend to repeat a given syntactic choice (e.g., active vs. passive) in the target, if the same choice was made in the prime. However, such structural priming effects have mostly been demonstrated in carefully controlled psycholinguistic experiments, thus raising the question of whether priming can also occur in natural, fully spontaneous conversation. Recent work addressed this question, providing evidence for priming effects in corpus data (Gries, 2005; Szmrecsanyi, 2005; Dubey et al., 2005).

In spite of corroborating experimental and corpus evidence, all current studies on structural priming share a serious limitation. They only deal with a small set of syntactic rules or constructions such as active vs. passive voice or direct object vs. prepositional object (e.g., as in *give your friend the book* vs. *give the book to your friend*). This raises the question of whether syntactic priming only affects these specific constructions, rather than being a more wide-spread phenomenon, perhaps applying to arbitrary syntactic rules. The present paper addresses this question.

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

One reason for the repeated use of linguistic structure is cognitive economy, which may well translate to computational economy. Repetition at lexical and syntactic levels leads to global alignment at semantic and situation model levels. Interaction partners establish a common sub-language over the course of a conversation: they agree on terms to use for objects, and methods to interpret the real world, creating common ground via the basic and automatic mechanism of priming, all without the expensive requirement to explicitly model their respective interlocutor's perspective (Pickering and Garrod, 2004).

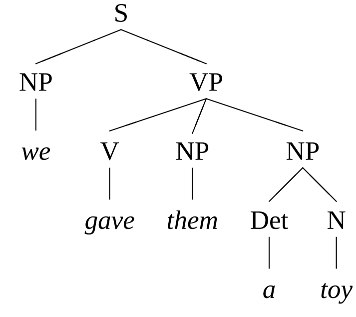
Pickering & Garrod's (2004) Interactive Alignment Model assumes that higher-level (semantic, situation-level) alignment is due to lower-level alignment (including syntactic priming). This predicts more priming in task-oriented dialogue, because situation-level alignment is typically required to perform a given task. In this paper, we test this prediction by comparing priming in spontaneous and task-oriented dialog. We distinguish *comprehension-production* (CP) priming, where the speaker first comprehends the prime (uttered by his/her interlocutor) and then produces the target, and *production-production* (PP) priming, where both the prime and the target are produced by the same speaker.

Methodology

In this study, we examined two spoken-language corpora with respect to structural repetition. The Switchboard (Marcus et al., 1994) and HCRC Map Task (Anderson et al., 1991) corpora both contain transcriptions of spoken dialogue and phrase-structure-based syntactic tree annotation.

Corpus processing

The trees were converted into phrase structure rules in order to list the rules that *license* the trees. For example, the (hypothetical) tree



would have been converted to three phrase structure rules:

(R1) S $\rightarrow$ NP VP,

(R2) VP $\rightarrow$ V NP NP and
(R3) NP $\rightarrow$ Det N.
This conversion is unique.[1]

Given the phrase structure rules for each utterance, we can now identify the repeated use of rules. A certain amount of repetition will obviously be coincidental. But structural priming would predict that a rule (*target*) occurs more often closely after a potential *prime* of the same rule (stimulus) than further away. Therefore, we can correlate the probability of repetition with the distance between prime and target.

As syntactic structure, we count each syntactic rule which licenses part of the syntactic analysis for a tree. For example, if a sentence-level conjunction leads to the rule *S* $\rightarrow$ *S conj S*, and such a conjunction occurs in utterances 3 and 11, we would observe a repetition at distance 8. This way, every syntactic rule is counted as a potential prime and (almost always) as a target for priming. Because interlocutors tend to stick to a topic during a conversation for some time, we exclude cases of syntactic repetition that are solely due to repetition of an entire phrase.

Generalized Linear Mixed Effects Regression

There are several ways to identify an effect of distance on repetition probability. One can normalize the number of observed repetitions by the number of expected repetitions for each syntactic rule by taking its prior probability of occurrence into account. The disadvantage of this is that for rare rules, we will see a grossly higher error than for rules with higher frequency. Such a data set would be difficult to model. Alternatively, one can examine the distribution of repetition counts over prime-target-distances and use a sampling technique to balance the number of trials across distances. Thirdly, we can contrast cases of structural repetition and cases where no repetition occurs between two speech units that occurred a chosen distance apart. We adopt the latter technique.

In this study, we use 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 exactly d units apart. GLMMs with a binary response variable can be considered a form of *logistic regression*.[2]

Regression allows us to fit a *model* to our data. A *model* is simply a choice of coefficients β_i, one for each explanatory variable i (and one for each of their interactions). β_i expresses the contribution of i to the probability of the outcome event, that is, in our case, successful priming. Our data is represented by extracted features – in our context, we will call them factors (discrete) and predictors (continuous explanatory variables).

For example, the β_i estimates allow us to predict the decline of repetition probability with increasing distance between prime and target, or other variables such as corpus choice. If we see priming as a form of pre-activation of syntactic nodes, it indicates the decay rate of pre-activation. The scale for this coefficient is the logarithmic distance in number of utterances.[3]

The fitting algorithms for GLMMs allow non-normally distributed response variables, as in our case with the binary variable indicating priming / non-priming. We trained our models using Penalized Quasi-Likelihood (Venables and Ripley, 2002). The reported experiments were conducted on random samples of the corpora.

Table 1 summarizes a GLMM along with further figures that allow us to estimate whether the coefficients obtained are reliable (statistically significant).

Syntactic repetitions

Every pair of two equal syntactic rules up to a maximal distance is a potential case of priming-enhanced production. Consider the example shown in Figure 1, where a small subset of the rules that license constituents are marked. Two syntactic repetitions shown here are data points for our analysis. Repetitions a and b are both at distance 2, because the occurrences (prime and target) are two utterances apart. Repetition c would be included at distance 1, if the lexical content of prime and target differed. In c, however, we see a syntactic repetition that is due to lexical repetition. Repetitions of unary rules such as the one marked as d are not included. The third sentence lends the opportunity to include another repetition (of the prepositional phrase rule PP $\rightarrow$ IN NP), but unlike Dubey et al. (2005), this study is not concerned with within-utterance repetitions.

The following analysis shows the distribution of repetition probability over distance from the repetition (target) to the prime. In our data, each repetition occurrence of a syntactic rule R at distance d counts as *priming*. Each case where R occurs, but isn't primed d units beforehand in the dialog, is counted as *non-priming*.

Our goal is to model $\hat{p}(prime|target, n)$, that is, the sampling probability that a *prime* is present in the n-th utterance before *target* occurs. Without syntactic priming in the general case, we would assume that

$$\hat{p}(prime|target, n) = \hat{p}(prime|target).$$

In order to eliminate cases of lexical repetition of a phrase, e.g., names or lexicalized noun phrases, which we consider topic-dependent or cases of lexical priming, we only collect syntactic repetitions with at least one differing word.

For instance (Figure 1), we would have two cases of priming for the rule PP $\rightarrow$ IN-NP, namely at distance 2 (a,b), and two of non-priming at distance 1 (two occurrences of that rule and their non-occurrence in the previous utterance).

The distance between stimulus and target (DIST) is initially counted in utterances (Experiments 1 – 3), but later in seconds (Experiments 4 & 5), which also includes within-utterance priming. Additive priming by a stimulus that is repeated several times is not captured by the model. We looked for repetitions within windows of 25 utterances or 15 seconds. So,

[1] Obviously, when dealing with speech, we encounter constructions that cannot be analyzed with a traditional phrase-structure rules. The annotation of both corpora commonly assigns ad-hoc rules with flat derivations in such cases. This leads to a large set of extracted rules. Such rules are unlikely to be repeated. For the analysis of repetition, they represent no theoretical obstacle.

[2] The data are assumed to be binomially distributed. We will not generally give classical R^2 figures, as this metric is not appropriate to such GLMMs.

[3] In our analysis, we will focus on the coefficients rather than on the intercept β_0 because long-term adaptation effects and the granularity of syntactic annotations will show up in β_0. Both lie out of the scope of this study.

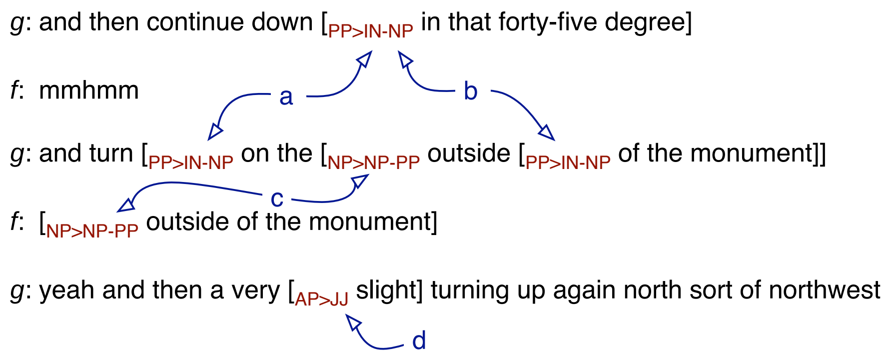


Figure 1: Two instances of syntactic repetitions (a,b), a lexical-syntactic one (c) and a preterminal rule (d) from Map Task.

each rule occurrence in the dialog can lead to up to 25 (or 15) data points for the various distances.

From our analysis, we drop all hapax rules (frequency $f = 1$) as well as outliers, that is 15 highly frequent rules ($f > 2,000$, out of 759) in the case of Map Task, and accordingly 9 ($f > 12,000$, out of 4695) in the larger Switchboard corpus.

We include the target utterance as a random factor in our model, grouping the several measurements (up to 25 for utterances or 15 for time) as *repeated measurements*, since they depend on the same target rule occurrence and are partially inter-dependent.

Again: without priming, one would expect that there are equally many cases of syntactic repetition, no matter the distance between first (*prime*) and second (*target*) occurrence. The analysis attempts 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 translate to more cases for shorter repetition distances, since priming effects usually decay rapidly (Branigan et al., 1999). (cf. Figure 3, which illustrates the decay.)

Additionally, we distinguish cases of *self-priming* (PP) and *priming between speakers* (CP) using the factor ROLE.

A predictor FREQ is included to express the logarithm of the normalized frequency of the repeated syntactic rule in the corpus (Experiments 3 – 5).

Exp. 1: Repetition in spontaneous conversation

Switchboard is a corpus of spontaneous spoken telephone dialogue among randomly paired, North American speakers who were given a general topic, but otherwise remained unrestricted. 80,000 utterances of the corpus have been annotated with syntactic structure. We use time-aligned (per word) data from the Paraphrase project (Carletta et al., 2004). $1,293,000$ repetitions could be found in $472,000$ extracted phrase structure rules, $4,700$ of which distinct.

Results

Syntactic rules (targets) are used more frequently when they occur shortly after the same rule (prime). The closer prime and target occur to one another, the stronger the preference is to repeat. Priming is present within a speaker (PP) and it decays rapidly, but there is a negative effect for priming between speakers (CP).

The model shows a reliable effect of $ln(\text{DIST})$: there are more repetition pairs with short distances than long ones ($t = -7.2, p < 0.0001$).

ROLE interacts with the decay coefficient for $ln(\text{DIST})$ ($t = 8.8, p < 0.0001$). The concrete result of that interaction is that the parameter for $ln(\text{DIST})$ in our model is -0.14 in PP priming, but 0.02 in CP priming. In Switchboard, we find evidence for PP priming, but for CP priming, the resulting decay coefficient (0.19) is actually positive, suggesting that speakers try to avoid repeating their interlocutor's sentence structure. Recall that high-frequency outliers had been dropped from the analysis. If we include them, we see that the difference between CP and PP is even stronger. Thus, in Experiments 3 and 5, we include the rule frequency as a predictor to evaluate the effect of frequency on priming strength.

Exp. 2: Repetition in task-oriented dialogue

To determine whether the type of dialogue affects syntactic repetition effects, we also analyzed the HCRC Map Task corpus. Map Task comprises more than 110 dialogs with a total of $20,400$ utterances, using 759 different phrase structure rules. Using exactly the same methodology as for Switchboard, we find $402,000$ syntactic repetitions in Map Task between the $157,000$ rules extracted from its syntactic analyses.

Like Switchboard, HCRC Map Task is a corpus of spoken, two-person dialogue in English. Unlike Switchboard, Map Task contains *task-oriented dialogue*: interlocutors work together to achieve a task as quickly and efficiently as possible. Subjects were asked to work together to find a route on a map. The interlocutors are in the same room, but have separate maps and are unable to see each other's maps. One of them, the Instruction Giver, is to describe a route, while the other one, the Instruction Follower, is to follow it on her own map. Their maps differ with respect to names of some locations, certain features (potential waypoints), and missing or displaced labels. Interlocutors were in the same room, while (in Switchboard) they used a telephone connection.

Syntactic priming as an instance of general priming or pre-activation is an almost universal effect. We know, however, that some control is exerted by the conditions of the dialogue and possibly by speakers tailoring their utterances to match the needs of their audience. Still, we would expect to find syntactic priming in the task-oriented dialogue of Map Task.

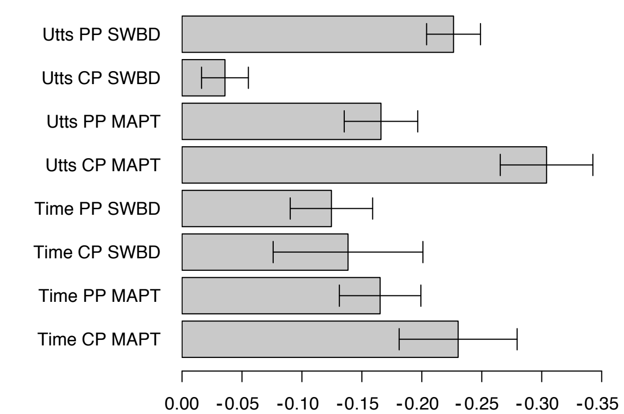


Figure 2: Priming effect sizes ($\ln(\text{DIST})$) under different ROLE and SOURCE situations. Prime-target distance by number of utterances (Exp. 3) and seconds (Exp. 5). 95% CI. Effects estimated from separately fitted nested regression models on separately sampled datasets.

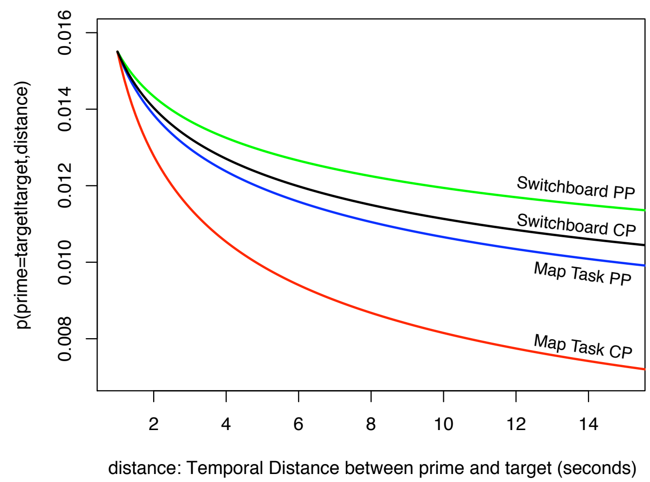


Figure 3: Decaying repetition probability estimates depending on the increasing distance between prime and target, contrasting different ROLE and SOURCE situations. (Exp. 5)

Again, a GLMM was built to correlate priming condition with the set of factors and predictors.

Results

Once again we find that repetition is more likely the shorter the distance between prime and target utterances is. Unlike in Switchboard, interlocutors repeat each other's syntactic structures more readily and more similarly to the way they repeat their own structures.

The model showed a reliable effect of $\ln(\text{DIST})$ ($t = -71.2, p < 0.005$).

ROLE had a reliable constant effect on repetition rates ($t = -11.0$, $p < 0.0001$), but there was no interaction between ROLE and DIST ($p = 0.92$).

This finding confirms experimental results by Bock and Griffin (2000) and Branigan et al. (1999), who find syntactic priming over longer distances, even though the effect decays. (The effect of ROLE on bias may be related to speaker idiosyncracies, i.e. more chance repetition within speakers.)

To determine whether there is a significant influence of dialogue type on priming, comparing the effects we have seen in experiments 1 and 2, we built a further model, described in the next section.

Exp. 3: Comparing corpora

With their Interactive Alignment Model, (Pickering and Garrod, 2004) argue that the situation-model alignment of speakers is due to lower-level priming effects. In task-oriented dialogue, and in the task carried out by participants in Map Task, speakers need to align in order to successfully complete their tasks. Thus, the theory would predict that syntactic priming between speakers (CP) is greater in task-oriented dialogue.

We test this hypothesis by fitting a model of the joint data set with SOURCE as a binary factor, indicating whether a repetition stems from Map Task (task-oriented) or Switchboard (not task-oriented). From Map Task, only dialogues in which interlocutors could not see one another where included.

Results

As seen in the previous experiments, it can make a difference whether a speaker primes themself or is primed by their interlocutor. Interestingly, the gap between CP and PP priming is substantially affected by the choice of corpus (last two interactions in Table 1). In both corpora, we find a positive PP priming effect. However, in Map Task, CP and PP priming cannot be distinguished (cf. Experiment 2), while in Switchboard, there is little CP priming (cf. Experiment 1). Figure 2 (first four bars) provides the resulting priming strength estimates for the four factorial combinations of ROLE and SOURCE at increasing distance. Also, priming is stronger for less frequent rules.

For Switchboard, the model estimates a higher coefficient for $\ln(\text{DIST})$, suggesting that there was faster decay in Map Task (Baseline effect of LN(DIST): $\beta_{lnDist} = -0.092, p < 0.0001$; $\beta_{lnDist:CP} = 0.083, p < 0.0001$; $\beta_{lnDist:MapTask} = -0.044, p = 0.05$; $\beta_{lnDist:CP:MapTask} = -0.140, p < 0.0001$). Frequency is negatively correlated with decay ($\beta_{lnDist:lnFreq} = 0.049, p < 0.0001$).

Finding the marked difference between CP and PP priming, and also a clear PP priming effect in spontaneous conversation, extends Dubey et al. (2005), who do not find reliable evidence of adaptation within speakers in Switchboard for selected syntactic rules in coordinate structures.

Thus, the data is consistent with the hypothesis that semantic alignment in dialogue is based on lower-level (syntactic) priming. However, when comparing data across corpora, we need to be careful to ensure that differences in genre and annotation are not the primary cause of the effect at hand. The coefficient for pre-activation decay is sensitive to utterance length, which becomes an issue for instance when utterances are not consistently marked or if decay occurs over time and not with utterances. Indeed, most utterances in Switchboard are actually *dialogue turns*, and given the genre, they are usually longer than those in Map Task. Therefore, it makes sense

	coefficients (β_i)	Std. Error
Intercept	-3.778	0.025 ***
$\ln(\textsc{Dist}_{Time})$	-0.057	0.015 **
$\ln(\textsc{Freq})$	0.538	0.190 ***
$\ln(\textsc{Dist}) : \ln(\textsc{Freq}))$	0.083	0.010 ***
$\ln(\textsc{Dist}) : (\textsc{Role} = CP)$	-0.031	0.012 *
$\ln(\textsc{Dist}) : (\textsc{Role} = PP) : (\textsc{Source} = MapTask)$	-0.050	0.014 **
$\ln(\textsc{Dist}) : (\textsc{Role} = CP) : (\textsc{Source} = MapTask)$	-0.137	0.018 ***

Table 1: The regression model for the joint data set of Switchboard and Map Task (Exp. 5). This is the minimal model without insignificant covariates. * $p < 0.01$, ** $p < 0.005$, *** $p < 0.0001$.

to verify the hypothesis using *time* as the relevant decay correlate. We do so in Experiments 4 and 5.

Exp. 4: Pre-activation decay: over time, or with each utterance?

While the previous experiments have shown that repetition probability decays soon after any stimulus, it is unclear whether the pre-activation diminishes with time, or with actual linguistic activity. To some extent, corpora can help make that distinction.

The differences between conversational and task-oriented dialogue that we pointed out (Experiment 3) are founded on the correlation of distance between prime and target and repetition likelihood. This correlation is likely to be sensitive to the *scale* of DISTANCE. As an alternative, we can use the delay between the left boundaries of the priming and target phrases as the relevant predictor.

The models discussed measure the distance between prime and target in *utterances*. In this experiment, we fitted a second regression model, estimating decay over *time*.

To compare the two (obviously interrelated) predictors $\textsc{Dist}_{Time}$ and $\textsc{Dist}_{Utts}$, we estimated two simple linear regression models, one for time, the other one for number of utterances as predictor. Such regression models can, as opposed to GLMMs, produce a meaningful R^2 measure. In these models, we include the maximum-likelihood estimate of the number of chance repetitions, which is calculated from the overall frequency of each syntactic rule (this is in addition to the covariates discussed before). The response variable here is not binary, as in the other experiments, but a count of actual rule repetitions. The complete interaction term is $rep \sim ln(\textsc{Dist}_{Utts}) * \textsc{Role} * \textsc{Source} + \textsc{Expected}$.[4]

The goodness-of-fit measure R^2 helps us determine how much of the variance in our data is explained by the model.

Results

For distance over utterances, R^2 is 0.91, for time (in 1-second buckets) it is 0.89, a similar size.

Thus, there is no compelling empirical evidence to assume $\textsc{Dist}_{Time}$ as a predictor over the work-load-based one (using utterance distance) chosen before. Because we cannot reasonably opt for one of the alternatives, we will reevaluate the effect of corpus choice seen in Experiment 3, this time using $\textsc{Dist}_{Time}$.

[4]These models assume a normal distribution as opposed to the appropriate Poisson one.

Exp. 5: Priming over time

While time- and utterance-based models fit their respective data similarly well, *time* is a theoretically attractive measure of distance, in particular because the *utterance* is difficult to delineate in the context of speech.

The methodology of this experiment is as in Experiment 3, except that $\textsc{Dist}_{Time}$ is the distance predictor, instead of the $\textsc{Dist}_{Utts}$ used previously.

Results

The interaction of corpus type and priming decay found in Experiment 3 holds. CP priming is stronger in task-oriented dialogue. Table 1 contains the estimated model.

The model based on temporal distance makes essentially comparable predictions. The SOURCE has an interaction effect on the priming decay $ln(\textsc{Dist})$, both for CP priming ($\beta_{lnDist:CP:MapTask} = -0.137, t = -7.6, p < 0.0001$) and for PP priming ($\beta_{lnDist:PP:MapTask} = -0.050$, $t = -3.7, p < 0.0005$). Figures 2, 3 provide the predictions for the four combinations of ROLE and SOURCE.

Discussion

Both corpora of spoken dialogue we investigated showed an effect of distance between prime and target in syntactic repetition, thus providing evidence for a structural priming effect for arbitrary syntactic rules. In both corpora, we also found reliable effects of both production-production (PP) priming (self-priming) and comprehension-production-priming. But only in the Map Task, a corpus of task-oriented dialogue did we find evidence for stronger CP priming than PP priming.

A possible explanation for these results is the reduced cognitive load that we can reasonably assume for spontaneous, everyday conversation (as in the Switchboard corpus). Pickering and Garrod (2004) suggest that interlocutors reduce their workload by aligning their linguistic and semantic representations, as re-using structure is easier than creating it. As cognitive load in non-task oriented, spontaneous conversion is low, speakers reduce the amount of priming that is required in dialogue that related to a difficult difficult task. The fact that we consistently see stronger priming for less frequent syntactic rules supports the cognitive-load explanation: frequently used rules are more accessible, hence their representations need less pre-activation.

Another reason may simply be that interlocutors in Switchboard (as in all spontaneous dialogue) switch topics frequently, engaging in longer turns in between. Such a sequence of monologues may, in general, be less affected by priming. The hypothesis that topic switches reduce priming may be tested in a future study.

On the other hand, one could expect that the impoverished single channel (phone line in Switchboard) leads speakers to make an effort to at least accept more self-priming (PP), designing their message so that they could be easily understood. Such *audience design* would be in line with work by Pearson et al. (2004), who found that speakers use less alignment (or priming) when talking to an (artificial) interlocutor that was perceived to have better linguistic capabilities. However, we see little actual evidence of speakers having difficulty understanding each other over the phone line, and they only show self-priming effects in the time-distance based models.

The Interactive Alignment Model (Pickering and Garrod, 2004) provides a viable explanation for the different effects that the two corpora expose. What we observe is the reciprocal boosting of syntactic priming and the alignment of the situation models present in task-oriented dialogue. The interaction partners synchronize their situation models in the task-oriented setting, which co-occurs with cross-speaker priming (CP) on other communicative levels. While self-priming may have to do with reduced cognitive load in production, the CP priming may be enhanced by sharing a situation model.

Conclusions

Reliable syntactic priming effects can detected in dialogue even when the full range of syntactic rules is taken into account instead of selected constructions with known strong priming effects. We have modelled syntactic priming as the decay of repetition probability of syntactic rules, either in the course of linguistic activity, or over time.

The parameters of priming vary with the setting of the conversation. In particular, we believe that the task-orientedness of the dialogue and increased cognitive load may boost alignment between speakers.

Since dialogue systems are often task-oriented, they may leverage the effect to resolve ambiguities or to produce better aligned output. Priming phenomena could also be exploited to aid automated processing, for instance in Automatic Speech Recognition using Cache Models (Kuhn and de Mori, 1990) and also in parsing (Charniak and Johnson, 2005; Dubey et al., 2006).

Acknowledgments

The authors would like to thank Roger Levy and Martin Pickering for their helpful comments. A preliminary version of Experiment 5 on separately sampled data has been described in Reitter et al. (2006). The first author was supported by 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.

J. Kathryn Bock and Zenzi Griffin. 2000. The persistence of structural priming: transient activation or implicit learning? *Journal of Experimental Psychology: General*, 129:177–192.

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.

Jean Carletta, S. Dingare, Malvina Nissim, and T. Nikitina. 2004. Using the NITE XML toolkit on the Switchboard corpus to study syntactic choice: a case study. In *Proc. 4th Language Resources and Evaluation Conference*. Lisbon.

Eugene Charniak and Mark Johnson. 2005. Coarse-to-fine n-best parsing and MaxEnt discriminative reranking. In *Proc. 43th Annual Meeting of the Association for Computational Linguistics*. Ann Arbor, Michigan.

Amit Dubey, Frank Keller, and Patrick Sturt. 2005. Parallelism in coordination as an instance of syntactic priming: Evidence from corpus-based modeling. In *Proc. HLT/EMNLP-2005*, pages 827–834. Vancouver.

Amit Dubey, Frank Keller, and Patrick Sturt. 2006. Integrating syntactic priming into an incremental probabilistic parser, with an application to psycholinguistic modeling. In *Proceedings of the 21st International Conference on Computational Linguistics and 44th Annual Meeting of the Association for Computational Linguistics*. Sydney.

Stefan Th. Gries. 2005. Syntactic priming: A corpus-based approach. *Journal of Psycholinguistic Research*, 34(4):365–399.

Roland Kuhn and Renato de Mori. 1990. A cache-based natural language model for speech recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(6):570–583.

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*. Plainsboro, New Jersey.

J. Pearson, M. Pickering, H. Branigan, J. McLean, C. Nass, and J. Hu. 2004. The influence of beliefs about an interlocutor on lexical and syntactic alignment: Evidence from human-computer dialogues. In *10th Annual Conference Architectures and Mechanisms of Language Processing*. Aix en Provence, France.

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

David Reitter, Frank Keller, and Johanna D. Moore. 2006. Computational modelling of structural priming in dialogue. In *Proc. Human Language Technology conference - North American chapter of the Association for Computational Linguistics annual meeting '06*. New York City.

Benedikt Szmrecsanyi. 2005. Creatures of habit: A corpus-linguistic analysis of persistence in spoken english. *Corpus Linguistics and Linguistic Theory*, 1(1):113–149.

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