

Structural Repetition in Restricted and Unrestricted Dialogue

Patrick G. T. Healey, Matthew Purver
Cognitive Science Research Group
Queen Mary University of London
{p.healey,m.purver}@qmul.ac.uk

Christine Howes
Centre for Language Technology
University of Gothenburg
christine.howes@gu.se

Language Processing in Dialogue

Can structural priming help to explain language processing in dialogue? Dialogue is the environment in which language is first encountered, acquired and used and arguably represents the most important challenge for models of language processing.

The literature on structural priming provides a large body of evidence that exposure to a particular syntactic structure, in production or comprehension, makes individuals more likely to repeat that structure. However, the strongest evidence for these effects comes from experimental studies of *individual* language processing, i.e. by individual participants who are not actively engaged with a conversational partner. In a typical experimental paradigm, participants are exposed to sequences of specific types of sentences. For example, participants are presented with sentences to read in one of two forms, such as a prepositional object ("A rock star sold some cocaine to an undercover agent.") or double object construction ("A rock star sold an undercover agent some cocaine."). The form of the prime sentence influences the form of an unrelated sentence produced on a subsequent picture description task (Bock, 1986).

It has become common to claim that the structural priming effects found in these studies extend to dialogue, and even provide the basis for successful communication (Pickering and Garrod, 2004; Pickering and Ferreira, 2008; Jaeger and Snider, 2013, and the call for papers for this special issue a.o.). However, there are problems with this generalisation. Common experimental paradigms such as picture description tasks significantly restrict the naturalness of the interaction by limiting people's responses (Branigan et al., 2000, 2006) and face the difficulties associated with using confederates in dialogue experiments (Kuhlen and Brennan, 2013). Corpus studies which are cited to support for the idea that syntactic priming is prevalent in naturally occurring dialogue either also focus on a few specific structures (Gries, 2005; Szmrecsanyi, 2005), or if they have broader coverage in terms of syntactic structures often focus on task-based dialogues (Reitter et al., 2006).

Healey et al. (2014) examined patterns of repetition in unrestricted natural dialogues and found that when the effects of lexical repetition are taken into account people systematically avoid using the same constructions in the next turn i.e. a pattern of structural divergence. They concluded that in natural dialogue structural priming effects are overwhelmed by the demands of constructive engagement with a conversational partner. Repetition alone would quickly bring ordinary conversation to a halt, preventing conversations from moving forward. Productive conversation involves people providing answers to questions, responding to proposals, and, of course, by changing topics and adding new material. According to Healey et al. this leads to the reuse of lexical items in contrasting syntactic contexts.

These findings are controversial and there are several possible alternative explanations. Possibly the most obvious contrast with prior work is the focus on unrestricted natural dialogue. Patterns of repetition in restricted or task-oriented dialogues are likely to be different. In many of the standard dialogue tasks described in the literature, participants encounter recurring situations or materials which should be expected to foster repetition. In these situations structural priming might provide a useful, low-cost way of promoting co-ordination. A second possible explanation, drawn from the literature, is that repetition effects are specific to certain kinds of syntactic structures; in particular it has been suggested that low frequency constructions are more likely to be repeated (Jaeger and Snider, 2013).

In this paper we test these two explanations by comparing patterns of lexical and syntactic repetition in samples of context restricted and unrestricted dialogues, and by comparing general syntactic repetition to that for only low-frequency syntactic constructions. We calculate the actual levels of observed repetition and the chance levels of repetition for each dialogue genre and we analyse for independent effects of syntactic repetition following the method described in Healey et al. (2014).

Methods

Three corpora are used; The British National Corpus (BNC), the Diachronic Corpus of Present-Day Spoken English (DCPSE) and the Switchboard Corpus (SWBD). The syntactic parses for the BNC were obtained automatically using a CCG parser (Clark and Curran, 2007). Switchboard and the DCPSE have full hand-coded syntactic parses. Switchboard has only dyadic dialogues and, following Healey et al. (2014) we use only the dyadic dialogues from the BNC and DCPSE as these are where priming effects are predicted to be strongest (Pickering and Garrod, 2004).

One sample of restricted contexts is provided by the BNC 'Educational' dialogues which consist principally of classroom interactions and home tutorials. A second is the 'Broadcast Interviews' section of the DCPSE. We also consider the Switchboard dialogues which are partially restricted in that they involve strangers talking on the telephone about pre-defined topics but who are otherwise relatively unconstrained. We compare these with the unrestricted dyadic dialogues from the 'Demographic' portion of the BNC and the 'Informal Conversations' portion of the DCPSE.

For each sample we compare patterns of repetition across turns in 'Real' dialogues and in 'Chance' dialogues. The Chance dialogues are created by randomly re-shuffling turns from the real dialogues as transcribed in the corpora (see Healey et al., 2014). This keeps the language sample constant while ensuring that there can be no effects of interaction or turn sequence. This enables us to separate out the amount of lexical and syntactic repetition specifically attributable to language processing in dialogue. We compare each person's turn with the last turn produced by themselves or by their conversational partner (or for the 'Chance' case, with a randomly chosen turn produced by themselves or another), for each dialogue genre.

Results

The raw means for next turn lexical and syntactic 'other' repetition in the different corpora are shown in Table 1. Note that the greater variety of parse trees for the machine-coded BNC produces lower overall levels of syntactic (but not lexical) match in all genres than the hand-coded dialogues of the DCPSE and Switchboard.

Table 1 shows firstly that levels of cross-speaker repetition tend to be higher in the restricted dialogues, especially in the Broadcast Interviews. In both the BNC and DCPSE the context restricted dialogues show over 40 % more syntactic repetition and over 30% more lexical repetition than in the unrestricted dialogues. This suggests that genre does have a substantial impact on the likelihood of

Table 1: Average Syntactic and Lexical Repetition of Others

Corpus	Real		Chance	
	Syntactic	Lexical	Syntactic	Lexical
Switchboard	0.21	0.07	0.28	0.07
Broadcast Interview (DCPSE)	0.37	0.12	0.36	0.08
Unrestricted Dialogue (DCPSE)	0.21	0.08	0.19	0.05
Educational Dialogue (BNC):	0.23	0.15	0.22	0.12
Unrestricted Dialogue (BNC):	0.12	0.07	0.12	0.04

repetition even for the relatively weak forms of restriction represented by Broadcast Interviews and Educational dialogues.

Secondly, Table 1 shows that chance levels of repetition are consistently high. Averaging across the BNC and DCPSE, the real dialogues show only 1% more syntactic repetition and only 3% more lexical repetition than would be expected by chance. Switchboard is the exception with chance levels of syntactic repetition actually higher than real levels. We return to this point below.

Thirdly, the summary data show that the chance levels of repetition track the real levels of repetition quite closely indicating that more restricted genres involve more limited linguistic repertoires independently of any additional effects that could be attributed to language processing in dialogue.

Table 2 displays the parallel data for self-repetition. This follows broadly the same pattern as other-repetition with syntactic and lexical similarity much higher in the restricted dialogue genres. Chance levels of repetition are again generally high and track real levels quite closely for each genre: syntactic repetition is only 1% above chance and lexical repetition 3% above chance.

Table 2: Average Syntactic and Lexical Self Repetition

Corpus	Real		Chance	
	Syntactic	Lexical	Syntactic	Lexical
Switchboard	0.34	0.13	0.31	0.10
Broadcast Interview (DCPSE)	0.35	0.13	0.35	0.12
Unrestricted Dialogue (DCSPE)	0.23	0.10	0.21	0.07
Educational Dialogue (BNC):	0.26	0.18	0.25	0.15
Unrestricted Dialogue (BNC):	0.16	0.12	0.15	0.08

Comparing Tables 1 and 2 it seems that, with the exception of broadcast interviews, levels of self-repetition are consistently higher –in both real and chance dialogues– than levels of other-repetition.

The raw means in Tables 1 and 2 do not take account of the intercorrelation between lexical repetition and syntactic repetition. To test whether there is a pattern of syntactic repetition independent of lexical repetition in the different genres, analyses are carried out with lexical repetition as a covariate. This enables us to contrast patterns of syntactic in the Real with the Chance dialogues and, in addition, contrasting the genre samples in the BNC and DCPSE (Restricted vs. Unrestricted).

Other Repetition

BNC and DCPSE: GLMM analyses with Conversation (Real vs. Chance), Lexical Repetition and Genre (Restricted vs. Unrestricted) and Genre \times Conversation as fixed factors and Participants as a random intercept show reliable main effects of Genre (DCPSE: $F_{(1,251)} = 152, p < 0.001$ BNC: $F_{(1,780)} = 65.0, p < 0.001$) Lexical Repetition (DCPSE: $F_{(1,251)} = 102, p < 0.001$; BNC: $F_{(1,780)} = 173, p < 0.001$) and Conversation (DCPSE: $F_{(1,251)} = 20.7, p < 0.001$ BNC: $F_{(1,780)} = 5.04, p = 0.02$) but no Conversation \times Genre Interaction (DCPSE: $F_{(1,251)} = 0.23, p = 0.63$, BNC: $F_{(1,780)} = 0.17, p = 0.68$).

In both the BNC and DCPSE the restricted dialogues show substantially higher levels of syntactic repetition than the unrestricted dialogues. This is true for both the actual dialogues and the artificially constructed Chance dialogues. As the estimated marginal means in Table 3 shows, once we take the effect of lexical repetition into account the real dialogues show lower levels of syntactic repetition than the chance dialogues.

For Switchboard there is no genre contrast so the analysis is simplified. GLMM analysis of syntactic repetition of the other participant with Conversation (Real vs. Chance) as a fixed factor, lexical repetition as a covariate and participants as a random intercept shows main effects of conversation ($F_{(1,2597)} = 1,749, p < 0.001$) and lexical repetition ($F_{(1,2597)} = 998, p < 0.001$). As for the other two corpora Syntactic repetition is higher in the Chance dialogues than the Real dialogues (see Table 3)

Table 3: Estimated Marginal Means and Standard Errors for Syntactic Other-Repetition

	Real	Chance
DCSPE	0.26 (0.006)	0.28 (0.005)
BNC	0.15 (0.005)	0.16 (0.004)
SWBD	0.21 (0.001)	0.28 (0.001)

Self Repetition:

For self repetition both the BNC and DCPSE show reliable main effects of Genre (DCPSE: $F_{(1,251)} = 61.3, p < 0.001$, BNC: $F_{(1,767)} = 30.0, p < 0.001$) Lexical Repetition (DCPSE: $F_{(1,251)} = 65.2, p < 0.001$, BNC: $F_{(1,767)} = 163, p < 0.001$) but no effect of Conversation (DCPSE: $F_{(1,251)} = 0.03, p = 0.85$, BNC: $F_{(1,767)} = 1.13, p = 0.29$) and no Conversation \times Genre Interaction (DCPSE: $F_{(1,251)} = 0.7, p < 0.001$, BNC: $F_{(1,767)} = 0.17, p = 0.67$).

Self-repetition of syntactic constructions in Switchboard follows the same pattern. There is no main effect of Conversation ($F_{(1,2597)} = 2.34, p = 0.17$, Estimated Means: Chance = 0.33, Real = 0.33) however there is a lexical repetition effect ($F_{(1,2597)} = 1,540, p < 0.001$, Fixed Effect Coefficient: 1.04, $t = 39.3, p < 0.001$).

Overall, syntactic self-repetition is higher in the context restricted dialogues but does not rise above chance.

Structure Type

To assess whether these patterns of syntactic repetition vary depending on the frequency of the particular syntactic constructions involved we separate the compare the 10 most frequent syn rules (which

make up about 60% of observations) with the remainder.

Syntactic other repetition in Switchboard is compared in a GLMM analysis with Lexical Repetition as a covariate, Structure Type (Common vs Rare) and Conversation (Real vs Chance) and Conversation \times Structure Type as fixed factors and participant as a random intercept. This shows, as above, main effects of Lexical Repetition ($F_{(1,5190)} = 841, p < 0.001$) and Conversation ($F_{(1,5190)} = 129, p < 0.001$). It also shows a main effect of Structure Type ($F_{(1,5190)} = 58,895, p < 0.001$) and a Conversation \times Structure Type interaction ($F_{(1,5190)} = 154, p < 0.001$)

Table 4: Estimated Marginal Means for Conversation \times Structure Type interaction (SWBD)

	Real	Chance
Frequent	0.28	0.36
Rare	0.02	0.02

As Table 4 shows, there is a substantial difference in the levels of syntactic repetition for rare and frequent constructions. The interaction arises from the fact that divergence in switchboard is due to the most commonly used constructions.

Discussion

The basic pattern of results across corpora is consistent. Genre makes a substantial difference to the likelihood of repetition. Even the relatively weakly restricted dialogues sampled here show much more self- and other-repetition than the unrestricted samples. Second there seems to be a general tendency for more self-repetition than other-repetition. Third, in all cases there is a strong and consistent interrelationship between lexical and syntactic repetition (e.g. fixed coefficients for the BNC for other-repetition: = 0.66, for self-repetition = 0.53). This underlines the need to take account of this correlation when estimating levels of structural repetition.

Levels of lexical repetition are consistently higher than chance. This makes sense since people must repeat at least some words in dialogue to maintain topical coherence. However, levels of syntactic repetition appear to be much closer to chance. Moreover, when levels of lexical repetition are taken into account we find that regardless of the dialogue context, people do not repeat their own syntactic structures more often than would be expected by chance and they systematically *diverge* from their conversational partners choice of syntactic construction i.e. in all contexts people repeat each other’s syntactic constructions less than would be expected by chance. This replicates and extends the structural divergence effects reported in Healey et al. (2014).

This brings us back to the question of whether structural priming in general can be used to help explain the patterns of language in natural dialogue. The results presented here create difficulties for the claim that structural priming mechanisms can provide a general account of language processing in dialogue. If the preceding analysis is correct, people engaged in dialogue typically avoid structural repetition. However, the results for the analysis by structure type indicate that this depends on the specific constructions being considered. In the full paper we will explore whether the contrast between repetition effects for high and low-frequency constructions can reconcile the divergence findings with the repetition of specific constructions found reported in the literature. Some additional explanation is needed to account for why less frequent constructions should be more likely to repeat when they do occur.

It seems that the prospects for structural priming as a general mechanism for explaining successful dialogue are limited. The present results highlight a general problem for priming-based approaches to understanding dialogue. Outside of experimental contexts, language processing in natural dialogue is characterised not by repetition but by contrast. As noted, the difference between the observed levels of repetition and those that could be expected by chance is small: around 1% for syntax and 3% for lexicon. Ultimately, we need a model of language processing that can address the 97% of natural dialogue which does not repeat.

References

- Bock, J. K. (1986). Syntactic persistence in language production. *Cognitive psychology*, 18(3):355–387.
- Branigan, H., Pickering, M., and Cleland, A. (2000). Syntactic co-ordination in dialogue. *Cognition*, 75:13–25.
- Branigan, H., Pickering, M., McLean, J., and Cleland, A. (2006). Syntactic alignment and participant role in dialogue. *Cognition*, 104(2):163–197.
- Clark, S. and Curran, J. (2007). Wide-coverage efficient statistical parsing with CCG and log-linear models. *Computational Linguistics*, 33(4):493–552.
- Gries, S. (2005). Syntactic Priming: A Corpus-based Approach. *Psycholinguistic Research*, 34(4):365–399.
- Healey, P. G. T., Purver, M., and Howes, C. (2014). Divergence in dialogue. *PLoS ONE*, 9(6):e98598.
- Jaeger, T. F. and Snider, N. E. (2013). Alignment as a consequence of expectation adaptation: Syntactic priming is affected by the primes prediction error given both prior and recent experience. *Cognition*, 127(1):57–83.
- Kuhlen, A. K. and Brennan, S. E. (2013). Language in dialogue: when confederates might be hazardous to your data. *Psychonomic bulletin & review*, 20(1):54–72.
- Pickering, M. and Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, 27:169–226.
- Pickering, M. J. and Ferreira, V. S. (2008). Structural priming: A critical review. *Psychological Bulletin*, 134(3):427–459.
- Reitter, D., Moore, J., and Keller, F. (2006). Priming of syntactic rules in task-oriented dialogue and spontaneous conversation. In Sun, R., editor, *Proceedings of the 28th Conference of the Cognitive Science Society*, pages 685–690.
- Szmrecsanyi, B. (2005). Language users as creatures of habit: A corpus-based analysis of persistence in spoken English. *Corpus Linguistics and Linguistic Theory*, 1(1):113–150.