Structural Divergence in Dialogue

1 Introduction

It is often observed that people tend to match each others body movements, speech style, and patterns of language use during conversation (Giles Coupland and Coupland, 1991). Recently, it has been proposed that 'structural priming' should be added to this list (Pickering and Ferreira, 2008). Structural priming occurs when processing of a linguistic structure is facilitated by prior exposure to the same structure. The main evidence for these effects comes from experimental studies of individuals processing sequences of sentences. However, the Interactive Alignment Model has proposed that *cross*-person structural priming is a part of an automatic, resource-free priming mechanism that helps to underpin all successful human interaction (Pickering and Garrod, 2004, 2006).

We present evidence from a corpus analysis that, in fact, people tend to diverge in their use of syntactic structures in ordinary conversation. There are three problems with the corpus evidence normally provided for cross-speaker structural priming in conversation:

1. Automatic priming predicts an increase in matching of all structures but this claim has not been directly tested for ordinary dialogue. Prior work has mixed genres (e.g. Gries (2005)) and used subsets of syntactic constructions (e.g. Gries (2005); Reitter, Moore and Kelleher (2006)).

2. Prior studies have not used a control condition so the chance level of structural matching is not directly estimated and effects such as conversational genre or register cannot be discounted (cf. Tannen (2007)).

3. Lexical and Syntactic repetition are correlated (Gries (2005); Brannigan, Pickering and Cleland (2000)) but previous analyses have not adjusted for this.

2 Methods

We analyse the levels of cross-turn, cross-participant similarity and compare this with chance. Two corpora are used: the Diachronic Corpus of Present-Day Spoken English (DCPSE) which has a full set of hand-checked parse trees and the British National Corpus (BNC) combined with a set of machine-coded parse trees we generated using a CCG parser.

Spontaneous dyadic conversations are predicted to show the strongest cross-person priming (Pickering and Garrod 2004, p.187) so we restrict our sub-sample to two-person informal conversations as classified by the corpus authors. This results in a sample of 6,616 turns by 92 people in 47 conversations in the DCPSE and 95,169 turns by 389 people in 2884 conversations in the BNC.

Two conversation types are used; 'Real' and 'Control'. The Real conversations consist of the original interactions. The Control conversations are created from these by separating each person's turns in each conversation and interleaving them with turns randomly sampled from the rest of the

corpus. This ensures that the sample of words, syntactic constructions and conversation lengths in the two samples are identical.

For the calculation of lexical similarity (S_{lex}), we look at pairs of turns A and B and calculate the number of word pairs N_{AB} that match. This is then normalised for the total number of words in the two turns combined: $S_{lex} = N_{AB}/\sqrt{N_{AA} \times N_{BB}}$. This gives a score between 0 and 1, where 1 corresponds to verbatim lexical matching across turns and 0 corresponds to no matching.

For the calculation of syntactic similarity (S_{syn}) we look at the number of non-terminal syntactic rules shared by the analyses of pairs of successive turns (A and B) again normalised to account for the total number of rules in both turns: $S_{syn} = N_{AB}/\sqrt{N_{AA} \times N_{BB}}$.

3 Results

Two Linear Mixed Models analyses are used to test for unadjusted differences in average turn-by-turn syntactic and lexical similarity for each person's conversations. Subjects and Conversation are entered as random factors and Conversation Type (Real vs. Control) as a within-subjects fixed factor.

The hand-coded corpus (DCPSE) shows no effect of Conversation Type on syntactic similarity ($F_{(1,140)}$ = 1.76, p = 0.19), but a reliable effect on lexical similarity ($F_{(1,140)}$ = 98.3, p = 0.00). The machine-coded corpus (BNC) also shows no effect of Conversation Type on syntactic similarity ($F_{(1,3777)}$ = 2.11, p = 0.15), but a reliable effect on lexical similarity ($F_{(1,3777)}$ = 531, p = 0.00). People are no more likely to match syntactic structures in adjacent conversational turns than would be expected by chance but they are more likely to use the same words. On average this difference accounts for just over 3% of the words used in an utterance.

The preceding comparisons treat lexical and syntactic similarity as independent however these measures are correlated. To adjust for this two Linear Mixed Models analyses of syntactic similarity are carried out with Subjects and Conversation as random factors, Conversation Type as a fixed factor and lexical similarity as a covariate. The DCPSE shows a main effect of Conversation Type on syntactic similarity $F_{(1,168)}$ = 6.27, p = 0.01) and a significant Conversation Type $\times$ Lexical Similarity interaction $F_{(2,182)}$ = 49.5, p = 0.00). The BNC follows the same pattern with a main effect of Conversation Type on syntactic similarity $F_{(1,4319)}$ = 67.7, p = 0.00) and a significant Conversation Type $\times$ Lexical Similarity interaction $F_{(2,5056)}$ = 179, p = 0.00).

Parameter estimates show that people's level of syntactic matching in adjacent turns is reliably *lower* than chance when lexical similarity is taken into account (DCPSE $= -0.037, t = -2.50, p = 0.01$, BNC $= -0.024, t = -8.23, p = 0.00$). The marginal means in Table 1 indicate around 1.6% less syntactic matching across adjacent turns.

Table 1: Marginal Means for Syntactic Similarities Corrected for Lexical Similarity

Corpus	Real Mean	S.E.	Control Mean	S.E.
Hand-Coded (DCPSE)	0.184	0.004	0.211	0.005
Machine-Coded (BNC)	0.118	0.001	0.123	0.001

The interaction between Conversation Type and Lexical Similarity shows that the effect of lexical similarity on syntactic similarity is larger in the Real conversations because of the higher level of syntactic matching (see above).

4 Conclusion

These results follow the opposite pattern to that predicted by the priming mechanism proposed in the Interactive Alignment model. Although people are more likely to use the same words in adjacent turns they have a tendency to avoid using the same syntactic structures. The pattern for cross-person, cross-turn co-ordination in dialogue is thus different from the pattern established for within-person sentence processing.

Our results are consistent with a view of conversation as an activity in which people collaborate to move each topic forward by repeating some of the same words but placing them in different contexts. For example, by asking and answering questions or by making, appraising and elaborating on proposals. We find this more intuitive than a priming model in which the primary focus is repetition and which consequently has difficulties accounting for the forward momentum of interaction.

References

Brannigan, H., Pickering, M., & Cleland, A. (2000). Syntactic co-ordination in dialogue. *Cognition*, *75*(B), 13-25.

Giles, H., Coupland, N., & Coupland, J. (1991). Accommodation theory: Communication, context, and consequence. In H. Giles, J. Coupland, & N. Coupland (Eds.), *Contexts of accommodation* (pp. 1–68). Cambridge: Cambridge University Press.

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

Pickering, M., & Ferreira, V. (2008). Structural priming: A critical review. *Psychological Bulletin*, *134*(3), 427–459.

Pickering, M., & Garrod, S. (2004). The interactive alignment model. *Behavioral and Brain Sciences*, *27*(2), 169–189.

Pickering, M., & Garrod, S. (2006). Alignment as the basis for successful communication. *Research on Language and Computation*, *4*, 203–228.

Reitter, D., Moore, J., & Keller, F. (2006). Priming of syntactic rules in task-oriented dialogue and spontaneous conversation. In R. Sun (Ed.), *Proceedings of the 28th conference of the cognitive science society* (pp. 685–690).

Tannen, D. (2007). *Talking voices: Repetition, dialogue and imagery in conversational discourse*. Cambridge: Cambridge University Press. (Second Edition)