

Syntactic Alignment is an Index of Affective Alignment: An Information-Theoretical Study of Natural Dialogue

Fermín Moscoso del Prado Martín (fmoscoso@linguistics.ucsb.edu)

Department of Linguistics, University of California – Santa Barbara
Santa Barbara, CA 93106 USA

John W. Du Bois (jdubois@linguistics.ucsb.edu)

Department of Linguistics, University of California – Santa Barbara
Santa Barbara, CA 93106 USA

Abstract

We present an analysis of a treebank of spontaneous English dyadic conversations, investigating whether the degree of syntactic priming found across speakers is a function of the degrees of affective alignment and overall positivity of the speakers. We use information theory to measure the proportion of overlap between the syntactic structures of the speakers. The affective state of the speakers is indexed by aggregated measures of the affective valences of the words they use. We find that there is a positive relation between syntactic priming and affective alignment, over and above any lexical repetition effects. This constitutes evidence for the percolation of inter-speaker alignment across multiple levels of representation. This also illustrates the indexical value of syntactic alignment, as has been proposed in modern functional theories of grammar such as Dialogic Syntax.

Keywords: Affective Alignment; Corpus Study; Information Theory; Natural Dialogue; Spoken Language; Syntactic Priming; Treebank; Resonance; Valence

Introduction

Spoken dialogue is the natural home of human language. It is the only naturally evolved form of language, and the only one that is acquired from infancy across all human cultures. It follows that the structure of the human mind, and that of languages themselves, should be expected to be particularly fine-tuned to spoken interactions involving more than one speaker. This privileged status of dialogue has attracted increasing attention in recent years. Modern theories on the nature of human dialogue stress the importance of multi-level alignment between speakers. From the point of view of psychology, Pickering and Garrod (2004) propose the Interactive Alignment Theory. This theory states that the linguistic representations of speakers are automatically aligned across many levels. Reitter and Moore (2014) showed that increased syntactic alignment leads to increased performance in collaborative map tasks (possible through more successful communication). More recently, Abney, Paxton, Kello, and Dale (in press) found that speakers in dyadic conversations also mirror the complexity of each other's language. This alignment extends beyond language: Shockley, Santana, and Fowler (2003) report increased postural alignment between dialogue participants, and Chartrand and Bargh (1999) report that mirroring of gestures and postures across di-

alogue participants facilitates interactions and increases the affective bonding between participants.

From the point of view of linguistics, Du Bois (2014) introduces the framework of Dialogic Syntax, directly addressing Pickering and Garrod (2004)'s desideratum of "a grammatical framework that is designed to deal with language in dialogue rather than monologue". Du Bois' theory proposes a "resonance principle" by which speakers would strive, insofar as possible, to maximize resonance in their choice of syntactic structures. In Dialogic Syntax, even if there may be a large automatic component in the lexico-syntactic alignment between speakers, speakers do retain some degree of control over the alignment, and they might *choose* to misalign themselves lexically or syntactically to signal things such as "stance differentials". In other words, speakers do have some notion of the indexical value of the alignment itself. This is in line with the broad family of functional (e.g., Chafe, 1970) and cognitive (e.g., Goldberg, 2006) theories of grammar positing that grammatical structures are "fully invested with meaning at all levels" (Du Bois, 2014).

The view that alignment or resonance extends from linguistic levels of representation into other aspects of the speakers' minds, and that alignment at one level reinforces alignment at other levels (Pickering & Garrod, 2004) suggests that this could also extend to the affective states of dialogue participants. There should be a relationship between the similarity of the speakers' affective states and the alignment between the linguistic structures they use. Moreover, from the results of Chartrand and Bargh (1999) one could infer that the more positive affective status would lead to higher degrees of linguistic alignment. Here, we present a corpus study (i.e., a 'natural experiment'; e.g., Gries, Hampe, & Schönefeld, 2005) investigating whether one can detect such relations between linguistic and affective resonance. In particular, we test whether the degree of overlap between the syntactic structures used by two speakers is correlated with the degree to which they are in more or less positive mental states, and with the extent to which their mental states differ. If this were the case, it would provide evidence for an indexical function of syntactic structure, supporting the notion of a meaning- and affect-invested

dialogic syntax.

In what follows we present an analysis of a treebank of spontaneous natural conversations. Moscoso del Prado Martín (2015) already proved that there is a significant amount of priming both within and across speakers in this same corpus, contrary to what is claimed by some authors (Healey, Purver, & Howes, 2014). Therefore we proceed directly to investigate the relation between syntactic priming and affective expression. In order to measure the degrees of affective alignment and overall positivity of the participants in dialogues, we use –as a proxy– the affective valences of the words that they each use. This is in line with substantial evidence (cf., Pearl & Enverga, 2014) indicating that the affective states of speakers are reflected in the affective characteristics of the words they use, leaving a sort of ‘mindprint’ on language.

Measures of Lexical and Syntactic Alignment

We follow the information-theoretical approach developed by Moscoso del Prado Martín (2015) for estimating the degree of lexical and syntactic overlap between the speakers in each dialogue. The frequency with which each production rule r is used across a whole corpus is denoted by $f(r)$. The frequencies of occurrence of each non-lexical (i.e., non-terminal or pre-terminal) symbol n used in the phrase-structure trees are denoted by $f(n)$. Using these frequencies, we obtain a maximum likelihood estimate of the probability that symbol n will be expanded by a rule r having n in its left-hand side:

$$\hat{p}(r) = \frac{f(r)}{f(n)}, \text{ where } n = \text{lhs}(r), \quad (1)$$

with the functor $\text{lhs}(x)$ denoting the left-hand side of rule x . This corresponds to the probability of rule r in a probabilistic context-free grammar (PCFG; Booth & Thompson, 1973) induced from the treebank by maximum likelihood. We can use this probability to estimate the point-wise information or surprisal (Shannon, 1948) produced by encountering rule r :

$$\hat{I}(r) = -\log_2 \hat{p}(r). \quad (2)$$

This measure can easily be applied to a list L of productions rules, as –for instance– the productions rules of the tree in Figure 1, which are listed in Table 1.¹ If $f(r; L)$ denotes the number of times that a rule r occurs in the list L , then, the *informational content* of L is estimated by:

$$\hat{I}(L) = \sum_{r \in L} f(r; L) \hat{I}(r). \quad (3)$$

¹Notice that, if the list corresponds to the productions of a phrase-structure tree, then this is equivalent to computing the surprisal of the tree given the induced PCFG.

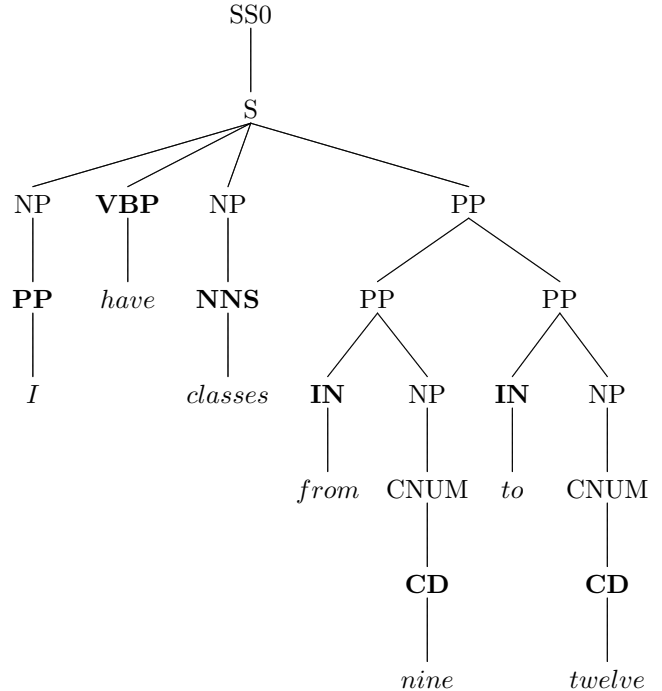


Figure 1: Example of a phrase-structure tree extracted from the TüBa-E/S corpus. Normal font denotes non-terminal nodes, bold font denotes pre-terminal nodes, and italic font denotes terminal nodes.

$\hat{I}(L)$ measures the *total information content* of the productions in list L . Notice that there are three types of nodes in the tree in Figure 1:²

- Non-terminal nodes (denoted in normal font), which correspond to grammar-internal symbols.
- Pre-terminal nodes (denoted in bold font), which correspond to part-of-speech annotations.
- Terminal nodes (denoted in italic font), which correspond to English words.

This distinction allows us to decompose the information content of Eq. 3 into two components, a *lexical information content* ($\hat{I}_{\text{lex}}(L)$) measuring the information due to the words used, and a *syntactic information content* ($\hat{I}_{\text{syn}}(L)$) measuring the information due to the use of particular syntactic constructions. The syntactic component is computed by applying Eq. 3 only to the subset of the *non-lexical productions* (rules that do not include terminals on their right-hand side) in L , whereas the lexical component is computed by considering only the *lexical productions* in L (those that have a non-terminal symbol in their right-hand side). Table 1 show how lexical and non-lexical rules are separated for the example

²Hence, in Figure 1, the non-terminal PP (a prepositional phrase) and the pre-terminal **PP** (a pronoun) are different symbols.

in Figure 1. If computed this way, it holds that the total information content is such the sum of the lexical and syntactic information contents,

$$\hat{I}(L) = \hat{I}_{\text{lex}}(L) + \hat{I}_{\text{syn}}(L). \quad (4)$$

Table 1: Phrase-structure production rules from the tree in Figure 1 split into lexical and non-lexical. Normal font denotes non-terminal symbols, bold font denotes pre-terminal symbols, and italic font denotes terminal symbols.

Lexical Productions	Non-Lexical Productions
PP → <i>I</i>	SS0 → S
VBP → <i>have</i>	S → NP VBP NP PP
NNS → <i>classes</i>	NP → PP
IN → <i>from</i>	NP → NNS
CD → <i>nine</i>	PP → PP PP
IN → <i>to</i>	PP → IN NP
CD → <i>twelve</i>	NP → CNUM
	CNUM → CD
	PP → IN NP
	NP → CNUM
	CNUM → CD

Consider now having two lists of productions L_1 and L_2 , corresponding for instance to the productions contained in two utterances or sets thereof by different speakers. One can now measure the amount of information that is shared between the two lists, the *shared information*,

$$\hat{I}(L_1; L_2) = \sum_{r \in L_1 \cap L_2} [f(r; L_1) + f(r; L_2)] \hat{I}(r). \quad (5)$$

This measure is bounded

$$0 \leq \hat{I}(L_1; L_2) \leq \hat{I}(L_1) + \hat{I}(L_2),$$

taking a value of zero iff the lists are fully disjoint, having no common productions, and being equal to the sum of their total informations iff every production that occurs in one list occurs also in the other. As was the case for the total information, the shared information can also be decomposed into its lexical and syntactic components, by counting only the lexical or non-lexical rules respectively.

The lexical shared information is a measure of the amount of lexical overlap between two sets of productions (i.e., utterances, dialogue turns, ...), and the syntactic shared structure measures the degree to which the two sets use the same syntactic constructions. They can therefore be used as measures of lexical and syntactic priming within or across speakers. As discussed in Moscoso del Prado Martín (2015), these are more adequate measures of priming than other measures that

have been used in corpus analyses (e.g., Healey et al., 2014; Pietsch, Buch, Kopp, & de Ruiter, 2012; Reitter, Moore, & Keller, 2006; Reitter & Moore, 2014). In addition, in order to safely compare utterances with very different lengths, it is useful take advantage of the bounds for the shared information, and define the *shared information ratio* (SIR) between two lists of productions as the percentage

$$\text{SIR}(L_1; L_2) = 100 \frac{\hat{I}(L_1; L_2)}{\hat{I}(L_1) + \hat{I}(L_2)}. \quad (6)$$

As above, the SIR can also be computed separately for the lexical and syntactic components (SIR_{lex} and SIR_{syn}).

Corpus Analysis

Materials and Methods

Corpus We obtained the Tübingen Treebank of Spoken English (TüBa-E/S; Hinrichs, Bartels, Kawata, Kordoni, & Telljohann, 2000). This corpus contains 649 transcribed English two-participant dialogues. The dialogues are segmented into conversational *turns* (uninterrupted contributions by one dialog participant), which are further broken down into approximately 30,000 *syntactic units* (corresponding to either full sentences or phrases).³ An HPSG-style phrase-structure tree is provided for each syntactic unit in the corpus.

The dialogues in the corpus are short (an average of 41.23 syntactic units per dialogue, ranging between four and 293 units) spontaneous interactions between two native speakers, concerning travel arrangements, appointment negotiations, and computer maintenance. The phrase-structure parses were constructed using a probabilistic parser, the outputs of which were manually corrected by human annotators

Pre-processing For each dialogue in the corpus, we grouped all phrase-structure trees by the participant who produced them. We extracted from the phrase-structure trees the phrase structure production rules that were used in the tree. Figure 1 provides an example of the phrase-structure trees contained in the corpus. The corresponding (lexical and non-lexical) production rules are listed in Table 1. In this way, we constructed two lists of phrase-structure productions for each dialogue, each corresponding to the productions used by one of the two participants. Each of these lists was further divided into two lists: one of lexical, and another of non-lexical productions.

Estimation of Relevant Measures

Syntactic and lexical overlaps For each of the 649 dialogues in the treebank, we computed the lexical and syntactic SIRs (i.e., SIR_{lex} and SIR_{syn}) using Eqs. 1–6.

³See Hinrichs et al. (2000) for more details.

Affective valence values For each word in a participant’s list of lexical productions, we obtained the affective valence rating from the database described by Warrimer, Kuperman, and Brysbaert (2013). This database contains average human affective ratings for about 14,000 English words. The ratings are on a 1–9 scale, where 1 denotes maximum positive valence and 9 the maximum negative valence. This process involved looking up whether the word –as contained in the corpus– was contained in the database, and otherwise looking up the word after lemmatization (using the WordNet lemmatizer; Miller, Beckwith, Fellbaum, Gross, & Miller, 1990). Words that were not present in the database even after lemmatization were ignored. This produced a list of affective valence values associated to each participant in each dialogue.

For each dialogue, as a measure of affective alignment we computed the absolute value of the difference in the median affective values of the two participants. As a measure of overall valence of the dialogues, we also computed the median valence of the concatenation of the valence list of both participants in each dialogue. To diminish the sensitivity of the measure to outliers (i.e., words with exceedingly positive or negative valences) we used the median values instead of the mean. The mean is most often used to compute the valences of texts as a function of the words they contain (cf., Heise, 1965; Leveau, Jhean-Larose, Denhière, & Nguyen, 2012). It should be noted that our results do not hinge on this choice.

Results and Discussion

The calculations described above produced a table with 649 entries, one for each dialogue in the corpus. For each entry, the table includes its total length in number of syntactic units (mean: 41.23 units/dialogue; range: [4 – 296]), the lexical (mean: 41.92%, range: [3.96 – 68.17]) and syntactic (mean: 40.78%, range: [2.01 – 67.46]) SIR values, the median affective valence across the two speakers (mean: 6.08, range: [5.51 – 6.33]), and the difference in median affective valences between the two speakers (mean: .14, range: [0.00 – 1.27]). We fit a linear regression model with the syntactic SIR as the dependent variable, and the lexical SIR, (log) length in sentences, median valence, and difference in median valences, as independent variables, as well as all possible interactions between them. A fast backwards elimination of factors using the AIC retained all main effects, which were also significant according to a sequential ANOVA (log length: $F[1, 643] = 1582.36, p < .0001$; SIR_{lex}: $F[1, 643] = 321.73, p < .0001$; difference in median valences: $F[1, 643] = 5.96, p = .0149$; median valence: $F[1, 643] = 12.74, p = .0004$). The backwards elimination of factors also suggested keeping the interaction between log length and the difference in median valences.

However, this interaction did not reach significance according to the ANOVA ($F[1, 643] = 2.50, p = .1141$).⁴ The corrected R^2 value for this regression was 75%. The residuals and predicted values of the regression did not exhibit major non-normalities or correlations.

Figure 2 summarizes the significant main effects found in the regression analysis. Panel (a) shows a positive correlation between the length of the dialogue and the syntactic SIR, indicating that longer dialogs exhibit more syntactic priming across the speakers. Panel (b) shows a very strong main effect of the lexical SIR on the syntactic SIR. This effect lends itself to two possible explanations: On the one hand, this strong correlation between the lexical and syntactic SIRs could be the effect of the “percolation” of alignment across levels (Pickering & Branigan, 1998; Pickering & Garrod, 2004). On the other hand, the correlation between the two forms of inter-speaker alignment could be a by-product of the reuse of certain multi-word expressions. Therefore, in this respect, one cannot argue with much certainty that these results provide evidence for percolation of inter-speaker alignment between the syntactic and lexical levels.

Panels (c) and (d) in Figure 2 illustrate the crucial results of this study. Panel (c) shows how dialogues whose speakers are more aligned in terms of the affective valence of the words they use (i.e., lower difference in median valence ratings), are also more aligned in terms of syntactic priming. This effect remains over and above the effect of dialogue length and that of the degree of lexical priming. One could perhaps argue that this is not so surprising because the reuse of lexical items from one speaker to another would naturally decrease the difference in median valence scores. However, such an interpretation would also predict either

- that the effect should disappear when entering lexical alignment directly into the regression, or at least
- that there is a significant two-way interaction between the effect of lexical priming and that of the difference in valences.

Since neither of the above is true, we interpret this effect as evidence for the percolation of alignment between the affective and syntactic levels: Speakers who are more aligned in affective terms (as evidenced by their use of more positive words), also tend to be more aligned in terms of the syntax they use. This interpretation is further supported by the effect of the affective valence itself, plotted in panel (d). As we predicted, this effect indicates that the more positive terms are used overall in a conversation, the more the speakers tend to align with

⁴Besides not reaching significance, the two-way interaction had not been predicted a priori. Further considering that it was the result of testing 24 possible (unpredicted) interactions between the main effects, motivated us to discard further consideration of this interaction.

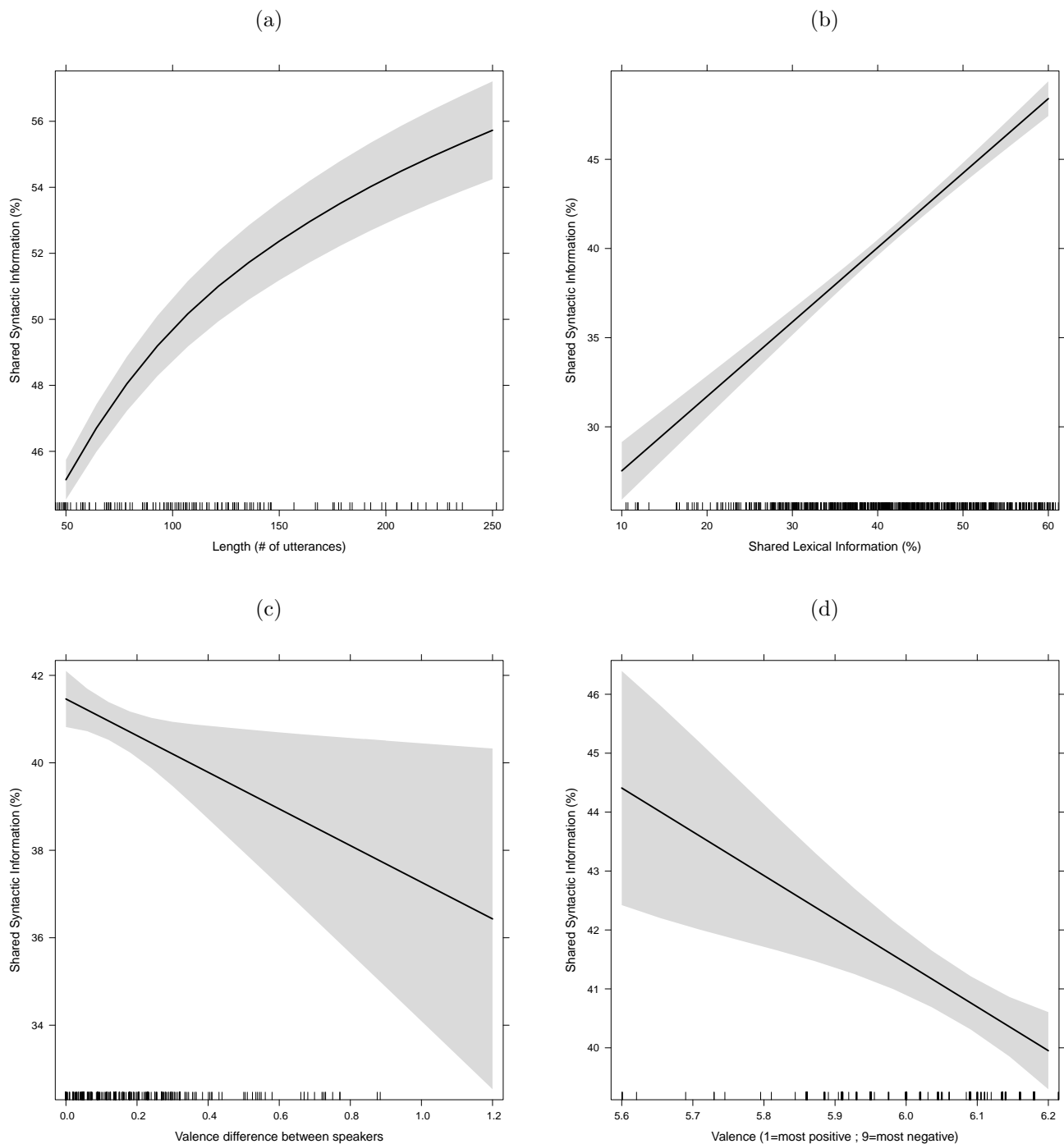


Figure 2: (a) Effect of dialogue length on SIR_{syn} . (b) Effect of SIR_{lex} on SIR_{syn} . (c) Effect of the difference in valence across speakers on SIR_{syn} . (d) Effect of the overall affective valence of the dialogue on SIR_{syn} .

each other syntactically. Notice that this effect cannot be discarded as a side-effect of lexical repetition, as this would require saying that positive words are more likely to be repeated (which would itself in any case support the percolation of alignment interpretation).

Our results provide evidence of a direct link between

syntactic priming and the degree to which the speakers are aligned in their use of positive or negative words, as well as with the overall level of positivity of the words they use. In turn, recent research shows that the affective aspects of these words are themselves indicators of the affective states of the speakers (Pearl &

Enverga, 2014). These results are important in two respects. First, they provide support for Pickering and Garrod (2004)'s percolation of inter-speaker alignment across levels, showing that the relation between alignments extends beyond purely linguistic aspects. Second, and equally important, our results indicate –in line with the predictions of functional/cognitive theories of grammar (e.g., Chafe, 1970; Du Bois, 2014; Goldberg, 2006)– that apparently abstract syntactic structure is nevertheless invested with indexical function, enabling speakers to signal and reinforce their degree of affective alignment. In other words, the choice of syntactic structure used by a speaker in dialogue is in itself meaningful in how it relates to the syntactic choices made by its interlocutor. This also contradicts early accounts (e.g., Bock, 1986) characterizing syntactic priming as plain subconscious priming on the choice between two structures, but devoid in itself of any communicative function.

Acknowledgements

We are indebted to Kevin Schäfer for helpful suggestions on this paper.

References

- Abney, D., Paxton, A., Kello, C., & Dale, R. (in press). Complexity matching in dyadic conversation. *Journal of Experimental Psychology: General*.
- Bock, J. K. (1986). Syntactic persistence in language production. *Cognitive Psychology*, 18, 355–387.
- Booth, T. L., & Thompson, R. A. (1973). Applying probability measures to abstract languages. *IEEE Transactions on Computers*, C-22, 442–450.
- Chafe, W. L. (1970). *Meaning and the structure of language*. Chicago, IL: The University of Chicago Press.
- Chartrand, T. L., & Bargh, J. A. (1999). The Chameleon Effect: The perception-behavior link in social interaction. *Journal of Personality and Social Psychology*, 76, 893–910.
- Du Bois, J. W. (2014). Towards a dialogic syntax. *Cognitive Linguistics*, 25, 359–410.
- Goldberg, A. (2006). *Constructions at work: The nature of generalization in language*. Oxford, UK: Oxford University Press.
- Gries, S. T., Hampe, B., & Schönefeld, D. (2005). Converging evidence: Bringing together experimental and corpus data on the association of verbs and constructions. *Cognitive Linguistics*, 16, 635–676.
- Healey, P. G. T., Purver, M., & Howes, C. (2014). Divergence in dialogue. *PLoS ONE*, 9, e98598.
- Heise, D. (1965). Semantic differential profiles for 1,000 most frequent English words. *Psychological Monographs*, 79, 1–31.
- Hinrichs, E. W., Bartels, J., Kawata, Y., Kordoni, V., & Telljohann, H. (2000). The Tübingen Treebanks for spoken German, English, and Japanese. In W. Wahlster (Ed.), *Verbmobil: Foundations of speech-to-speech translation* (pp. 550–574). Berlin – Heidelberg, Germany: Springer.
- Leveau, N., Jhean-Larose, S., Denhière, G., & Nguyen, B. (2012). Validating an interlingual metanorm for emotional analysis of texts. *Behavior Research Methods*, 44, 1007–1014.
- Miller, G. A., Beckwith, R., Fellbaum, C. D., Gross, D., & Miller, K. (1990). WordNet: An online lexical database. *International Journal of Lexicography*, 3, 235–244.
- Moscoso del Prado Martín, F. (2015). *Syntactic priming across speakers in natural dialogue is a lagged process: Evidence from English, German, and Japanese*. Manuscript submitted for publication.
- Pearl, L., & Enverga, I. (2014). Can you read my mind-print? automatically identifying mental states from language text using deeper linguistic features. *Interaction Studies*, 15, 359–387.
- Pickering, M. J., & Branigan, H. P. (1998). The representation of verbs: Evidence from syntactic priming in language production. *Journal of Memory and Language*, 39, 633–651.
- Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, 27, 169–226.
- Pietsch, C., Buch, A., Kopp, S., & de Ruiter, J. (2012). Measuring syntactic priming in dialogue corpora. In B. Stolterfoht & S. Featherston (Eds.), *Empirical approaches to linguistic theory: Studies in meaning and structure* (pp. 29–42). Mouton de Gruyter.
- Reitter, D., & Moore, J. (2014). Alignment and task success in spoken dialogue. *Journal of Memory and Language*, 76, 29–49.
- Reitter, D., Moore, J., & Keller, F. (2006). Priming of syntactic rules in task-oriented dialogue and spontaneous conversation. In *Proceedings of the 28th annual conference of the cognitive science society* (pp. 685–690). Austin, TX: Cognitive Science Society.
- Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, 27, 379–423, 623–656.
- Shockley, K., Santana, M., & Fowler, C. (2003). Mutual interpersonal postural constraints are involved in cooperative conversation. *Journal of Experimental Psychology: Human Perception and Performance*, 29, 326–332.
- Warrimer, A. B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods*, 45, 1191–1207.