In: Proceedings of the 64th Annual Mtg. of the Association for Computational Linguistics. Berlin, Germany, 2016.

Convergence of Syntactic Complexity in Conversation

Yang Xu and **David Reitter**
College of Information Sciences and Technology
The Pennsylvania State University
University Park, PA 16802, USA
yang.xu@psu.edu, reitter@psu.edu

Abstract

Using corpus data of spoken dialogue, we examine the convergence of syntactic complexity levels between interlocutors in natural conversations, as it occurs within spans of topic episodes. The findings of general convergence in the Switchboard and BNC corpora are compatible with an information-theoretic model of dialogue and with Interactive Alignment Theory.

1 Introduction

According to *Interactive Alignment* theory (Pickering and Garrod, 2004), mutual understanding in dialogue is helped by a variety of interconnected adaptation processes. Over the course of a conversation, interlocutors' linguistic productions assimilate at multiple levels, such as phonemes (Pardo, 2006), lexical choice (Garrod and Anderson, 1987), syntactic structures (Pickering and Branigan, 1998; Branigan et al., 2000; Reitter et al., 2006) and so on. The alignment at these levels contributes to the establishment of aligned *situation models* between speakers, which is the ultimate goal of a successful conversation (Pickering and Garrod, 2004; Reitter and Moore, 2007, 2014).

Alignment does not only refer to the mimicking and repetition of particular linguistic structures; it also includes the convergence at the statistical and ensemble level, which is known as *distributional matching* (Abney et al., 2014). Speech rates (Webb, 1969), probability distributions over syntactic forms (Jaeger and Snider, 2008), power law distributions of acoustic onset events (Abney et al., 2014), and social intent of the speech act (Wang et al., 2015) were all found to match between interlocutors.

An aspect of accommodation that presumably very much helps dialogue partners understand each other's language is *syntactic complexity*. Despite rich investigation of alignment in conversation, this property has been largely overlooked in the analysis of dialogue.

The general concept of syntactic complexity has, of course, been addressed in various ways. In educational psychology and applied linguistics, it is often defined as the degree of sophistication of language forms. It has broad application in the assessment of second language acquisition (Ortega, 2003; Lu, 2010, 2011), the readability test (MacGinitie and Tretiak, 1971), and elementary education (Abedi and Lord, 2001). In computational linguistics, previous studies have shown that the syntactic complexity of a sentence is closely related to the amount of information being transmitted (Genzel and Charniak, 2002, 2003; Jaeger and Levy, 2006; Jaeger, 2010). However, as far as we know, syntactic complexity as a high level feature of language production has not been investigated under the theoretical lens of the Interactive Alignment Model (Pickering and Garrod, 2004).

Therefore, the focus of this study is to track the syntactic complexity of different interlocutors as the conversation develops. A convergence of sentence complexity between interlocutors would be compatible with two pertinent theories. The first is the Interactive Alignment Model. The second is the Uniform Information Density hypothesis (Jaeger and Levy, 2006; Jaeger, 2010), as it applies to syntactic structure. It postulates that speakers will strive to keep information density approximately constant. In other words, if one interlocutor decreased their rate of information transmission, the other one would increase it in response. As far as syntactic complexity is proportional to the amount of information, this would imply that if one interlocutor changes their syntactic complex-

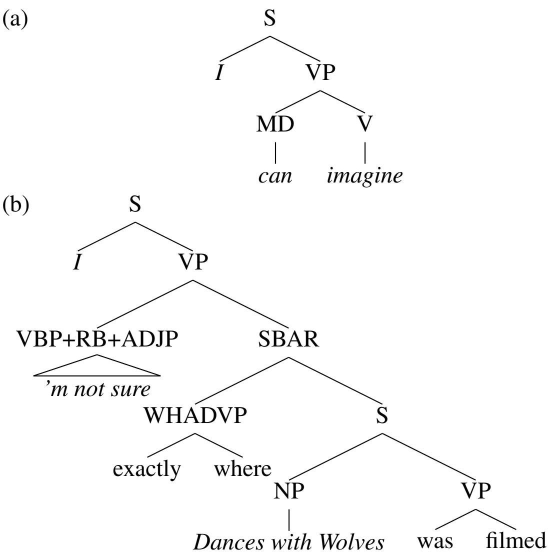


Figure 1: Contrast the syntactic complexity of a simple sentence (a) vs. a complex sentence (b). The tree depth of (a) is 4, while the value of (b) is 7. The branching factor of (a) is 1.38, while the value of (b) is 1.48

ity, their dialogue partner is likely to make the opposite change.

2 Methods

2.1 Corpus data

We use the Switchboard corpus (Godfrey et al., 1992) and the British National Corpus (BNC) (BNC, 2007) in this study. Switchboard contains 1126 conversations over the telephone, where each conversation features exactly two native American English speakers. From the BNC, we use only a subset of the data that contains spoken conversations with exactly two participants so that the dialogue structures are consistent with Switchboard.

2.2 Metrics of syntactic complexity

We consider three candidate statistics to measure the syntactic complexity of a sentence: *sentence length* (number of words in a sentence), *tree depth*, and *branching factor*. The first two are straightforward: syntactically complex sentences are typically used to express complex meanings, and thus are more likely to contain more words than simple ones. More complex syntactic structures, such as relative clauses and noun clauses, also have deeper parse trees (see Figure 1).

The third statistic, branching factor, is defined as the average number of children of all non-leaf nodes in the parse tree of a sentence. In contrast to tree depth, it measures the *width* of a tree structure, thus a sentence with a larger branching factor looks flatter.

These three statistics are inter-correlated. For instance, tree depth has an almost linear correlation with sentence length. To come up with a measure that solely characterizes the complexity of a sentence in terms of its tree structure, we normalize tree depth and branching factor by excluding the effect of sentence length. We adopt the method proposed by Genzel and Charniak (2002). Let f be a complexity measure of a sentence (tree depth or branching factor). We compute the average measure $\bar{f}(n)$ for sentences of the same length n $(n = 1, 2, \dots)$:

$$\bar{f}(n) = {}^{1}\!/_{|S(n)|} \sum\nolimits_{s \in S(n)} f(s) \qquad (1)$$

where s denotes a sentence, and $S(n) = \{s | l(s) = n\}$ is the set of sentences of length n. The normalized complexity measure is:

$$f'(s) = \frac{f(s)}{\bar{f}(n)} \qquad (2)$$

This normalized measure f' is not sensitive to sentence length. This gives us five metrics of complexity: *sentence length* (SL), *tree depth* (TD), *branching factor* (BF), *normalized tree depth* (NTD), and *normalized branching factor* (NBF).

2.3 Topic segmentation and speaker role assignment

To verify the hypothesized convergence of a certain statistic between two speakers in dialogue, one possible method is to measure whether the difference in that statistic becomes smaller as the conversation progresses. However, this design is overly simplistic in this case for several reasons. For instance, previous studies have found that sentence complexity in written text increases with its position (Genzel and Charniak, 2003); thus even if we observed that the difference of complexity becomes smaller, a ceiling effect could be a simpler explanation.

Additionally, the syntactic complexity of a sentence largely depends on the amount of meaning that is conveyed. Intuitively, when a speaker has a large amount of information to express, she tends to use more sophisticated syntactic constructions Linking this consideration to another very common scenario in dialogue: one interlocutor *leads* the conversation by steering the on-going topics,

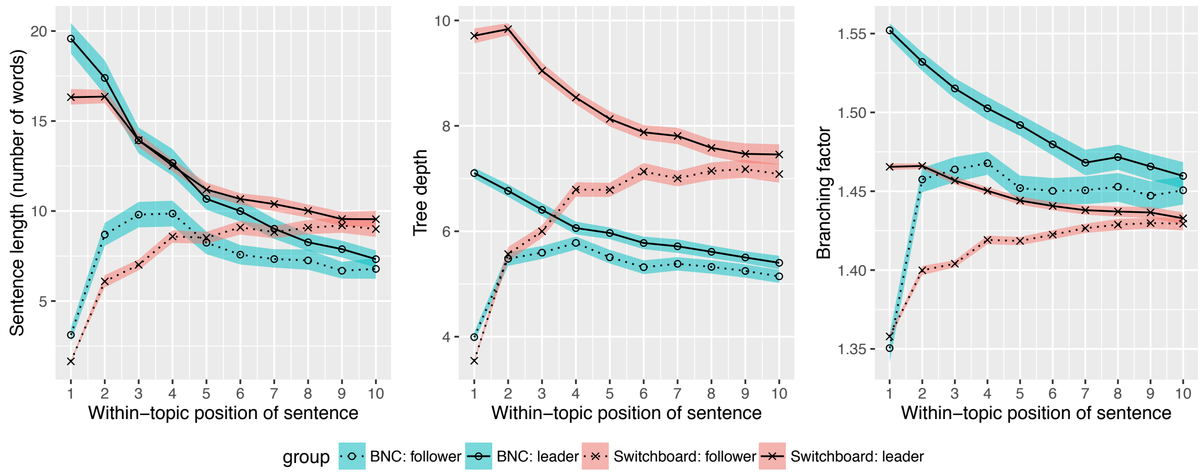


Figure 2: *Sentence length* (SL), *tree depth* (TD) and *branching factor* (BF) against within-topic sentence position (the relative position of a sentence from the beginning of the topic episode), grouped by speaker role, *leader* vs. *follower*. Shaded areas: bootstrapped 95% confidence intervals.

while the other participant *follows* along. Here, we are not talking about the *turn-taking* mechanism in dialogue, which describes the shift at the utterance level. Rather, we are describing the shift at a higher level in conversation, the *topic* level, which is formally referred to as *topic shift* in Conversation Analysis (Ng and Bradac, 1993; Linell, 1998). According to these related theories, a complete conversation consists of several *topic episodes*. Some speakers play a more active role in *leading* the unfolding of new topic episodes, while others play a more passive role by *following* the topic shift. Beginning a new topic means bringing in new information, thus it is reasonable to infer that the interlocutor's syntactic complexity would partially depend on whether he is playing the *leader* or the *follower*. Considering the fact that the leader vs. follower roles are not fixed among interlocutors (a previous leader could be a follower later and vise versa), we should not examine the convergence of syntactic complexity within the whole conversation. Rather, we want to zoom in to the finer scale of topic episodes, in which the interlocutors' roles are relatively stable.

Based on these considerations, we use the TextTiling algorithm (Hearst, 1997) to segment the conversation into several topic episodes. This is a sufficient topic segmentation method for our research questions, though it is less sophisticated compared to Bayesian models (Eisenstein and Barzilay, 2008) or Hidden Markov Models (Blei and Moreno, 2001).

Within each topic episode that resulted from the segmentation operation, we assign roles to the two speakers. This is based on which of the interlocutors is leading this topic episode, as previously explained. We use two rules to determine this *leader* and *follower* differentiation:

Rule I: If the topic episode starts in the middle of the speaking turn of speaker A, then let A be the *leader* of this topic.

Rule II: If the topic episode starts with a complete speaking turn, then let the first speaker who contributes a sentence greater than N words in length in this episode be the *leader*.

Note that the purpose of Rule II is to select the most *probable* topic leader, based on the intuition that longer sentences are more *likely* to initiate a new topic. Thus the determination of the N words threshold here is totally empirical. We use $N = 5$ as the final threshold, because for $N \geq 5$ our experiments draw similar results.

3 Results

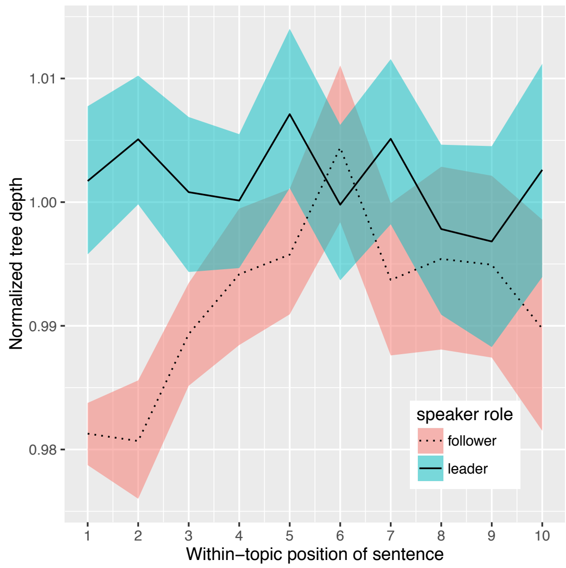
For each sentence in conversation, we compute the five earlier-discussed metrics of syntactic complexity: SL, TD, BF, NTD, and NBF.

For the first three metrics, SL, TD and BF, we observe convergence between topic leaders and followers, for both corpora (Fig. 2). Basically, topic leaders have higher syntactic complexity measures at the early stage of a topic episode, which drops gradually as the topic develops. The converse holds for topic followers. We fit 12 linear

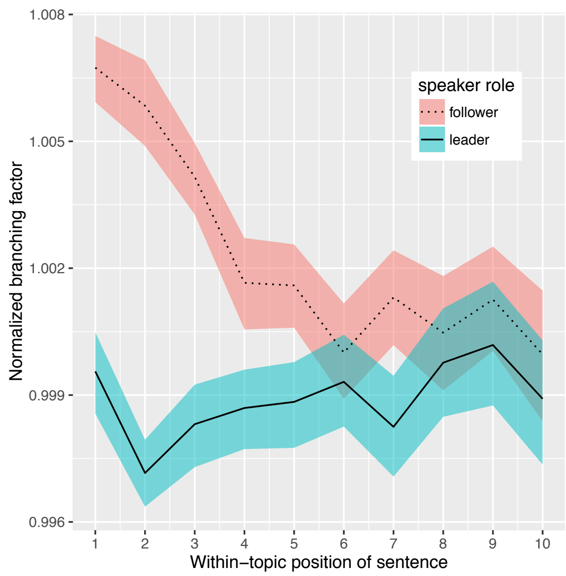
Table 1: β coefficients of the fixed effect (within-topic position) of the linear mixed models.

group	SL	TD	BF
Switchboard leader	0.363***	-0.129***	-1.82×10^{-3}***
Switchboard follower	0.188***	0.104***	2.141×10^{-3}***
BNC leader	-0.166***	-0.030***	-1.88×10^{-3}***
BNC follower	0.012	9.45×10^{-3}***	5.51×10^{-4}***

***$p < 0.001$



(a) NTD



(b) NBF

Figure 3: Two normalized metrics of syntactic complexity, tree depth (NTD) (a) and branching factor (NBF) (b), vs. within-topic position of sentences in Switchboard. Shaded areas: bootstrapped 95% confidence intervals.

mixed models (3 metrics $\times$ 2 roles $\times$ 2 corpora) using metrics as the respective response variables, the within-topic position as a fixed effect, and a random intercept grouped by individual speakers. We find a positive effect of within-topic position for leaders, and a reliably negative effect for followers (except SL of BNC follower), which confirms the observation of convergence trend (See Table 1).

For NTD and NBF, we observe convergence patterns in Switchboard, but not reliably in BNC (Figure 3). Linear mixed models are fit in similar ways, and the β coefficients are: for NTD, $\beta_{\text{leader}} = -2.2 \times 10^{-5}$, $\beta_{\text{follower}} = 9.7 \times 10^{-4***}$; for NBF, $\beta_{\text{leader}} = 6.8 \times 10^{-5*}$, $\beta_{\text{follower}} = -2.9 \times 10^{-4***}$ (*** indicates $p < 0.001$, and * indicates $p < 0.05$). Thus, a general trend seems supported. As NBF is the only metric that is lower in leaders and higher in followers, it could actually be an index for syntactic *simplicity*.

4 Discussion and Conclusion

By segmenting a conversation into several topic episodes, and then differentiating the interlocutors in terms of their roles in initiating the topic, leader or follower, we show that the syntactic complexity of the two interlocutors converges within topic episodes. The syntactic complexity of the topic leader decreases, while the complexity of the topic follower increases.

From an information-theoretical point of view, the syntactic complexity of a sentence is closely related to its amount of lexical information or negative entropy (Genzel and Charniak, 2002, 2003). By starting a new topic in conversation, the leading speaker brings novelty to the existing context, which often involves relatively long and complex utterances. On the other hand, the following speaker has to accommodate this change of context, by first producing short acknowledging phrases at the early stage, and gradually increase his contribution as the topic develops. Therefore, the convergence of syntactic complexity within a topic episode is a reflection of the process in which two interlocutors contribute jointly to build up common ground (Clark and Brennan, 1991) with respect to a certain topic.

We find our results explained the theoretical frameworks of common ground (Clark, 1996) and the Interactive Alignment Model (IAM, Pickering and Garrod, 2004), models which are sometimes characterized as opposing accounts of coordination in dialogue. From the common-ground perspective of language-as-activity, interlocutors play different roles in dialogue, and the coordination between these roles facilitates the successful unfolding of dialogue. Our account identifies two such macro-level roles: topic *leader* vs. *follower*. From the perspective of Interactive Alignment, interactions between interlocutors in a dialogue are accompanied by the alignment of linguistic elements at multiple levels, including syntactic rules. Thus, the micro-level convergence of syntactic complexity is predicted by the IAM. Therefore, our findings point to the possibility of a unified perspective that combines the two theories.

It is worth pointing out that we present some novel ideas about the scope of convergence. Existing studies focus on the alignment effect that is observable throughout the whole conversation. In our case, the convergence of syntactic complexity occurs within smaller scope: the topic episodes. Note that the direction of convergence is dynamic: a speaker of higher complexity in one episode might be of lower complexity in the next episode, depending on her role. The next questions arising from these patterns mirror those asked of other types of alignment: is complexity alignment purposeful, is it controlled by individual differences or situational goals, and can it predict task success? We leave these questions for future work.

Acknowledgments

This work has been funded by the National Science Foundation under CRII IIS grant 1459300.

References

Jamal Abedi and Carol Lord. 2001. The language factor in mathematics tests. *Applied Measurement in Education* 14(3):219–234.

Drew H Abney, Alexandra Paxton, Rick Dale, and Christopher T Kello. 2014. Complexity matching in dyadic conversation. *Journal of Experimental Psychology: General* 143(6):2304–2315.

David M Blei and Pedro J Moreno. 2001. Topic segmentation with an aspect hidden Markov model. In *Proceedings of the 24th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM, pages 343–348.

BNC. 2007. The British National Corpus, version 3 (BNC XML Edition).

Holly P Branigan, Martin J Pickering, and Alexandra A Cleland. 2000. Syntactic co-ordination in dialogue. *Cognition* 75(2):B13–B25.

Herbert H Clark. 1996. *Using language*. Cambridge University Press.

Herbert H Clark and Susan E Brennan. 1991. Grounding in communication. *Perspectives on socially shared cognition* 13(1991):127–149.

Jacob Eisenstein and Regina Barzilay. 2008. Bayesian unsupervised topic segmentation. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, pages 334–343.

Simon Garrod and Anthony Anderson. 1987. Saying what you mean in dialogue: A study in conceptual and semantic co-ordination. *Cognition* 27(2):181–218.

Dmitriy Genzel and Eugene Charniak. 2002. Entropy rate constancy in text. In *Proc. 40th Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics, pages 199–206.

Dmitriy Genzel and Eugene Charniak. 2003. Variation of entropy and parse trees of sentences as a function of the sentence number. In *Proc. 2003 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, pages 65–72.

John J Godfrey, Edward C Holliman, and Jane McDaniel. 1992. Switchboard: Telephone speech corpus for research and development. In *Acoustics, Speech, and Signal Processing, 1992. ICASSP-92., 1992 IEEE International Conference on*. IEEE, volume 1, pages 517–520.

Marti A Hearst. 1997. Texttiling: Segmenting text into multi-paragraph subtopic passages. *Computational Linguistics* 23(1):33–64.

T Florian Jaeger. 2010. Redundancy and reduction: Speakers manage syntactic information density. *Cognitive Psychology* 61(1):23–62.

T Florian Jaeger and Roger P Levy. 2006. Speakers optimize information density through syntactic reduction. In *Advances in Neural Information Processing Systems*. pages 849–856.

T Florian Jaeger and Neal Snider. 2008. Implicit learning and syntactic persistence: Surprisal and cumulativity. In *Proc. 30th Annual Meeting of the Cognitive Science Society*. pages 1061–1066.

Per Linell. 1998. *Approaching dialogue: Talk, interaction and contexts in dialogical perspectives*, volume 3. John Benjamins Publishing.

Xiaofei Lu. 2010. Automatic analysis of syntactic complexity in second language writing. *International Journal of Corpus Linguistics* 15(4):474–496.

Xiaofei Lu. 2011. A corpus-based evaluation of syntactic complexity measures as indices of college-level ESL writers' language development. *Tesol Quarterly* 45(1):36–62.

Walter H MacGinitie and Richard Tretiak. 1971. Sentence depth measures as predictors of reading difficulty. *Reading Research Quarterly* pages 364–377.

Sik Hung Ng and James J Bradac. 1993. *Power in language: Verbal communication and social influence*. Sage Publications, Inc.

Lourdes Ortega. 2003. Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing. *Applied Linguistics* 24(4):492–518.

Jennifer S Pardo. 2006. On phonetic convergence during conversational interaction. *The Journal of the Acoustical Society of America* 119(4):2382–2393.

Martin J Pickering and Holly P Branigan. 1998. The representation of verbs: Evidence from syntactic priming in language production. *Journal of Memory and Language* 39(4):633–651.

Martin J Pickering and Simon Garrod. 2004. Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences* 27(02):169–190.

David Reitter and Johanna D. Moore. 2007. Predicting success in dialogue. In *Proc. 45th Annual Meeting of the Association of Computational Linguistics*. Prague, Czech Republic, pages 808–815.

David Reitter and Johanna D. Moore. 2014. Alignment and task success in spoken dialogue. *Journal of Memory and Language* 76:29–46.

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 (CogSci)*. Cognitive Science Society, Vancouver, Canada, pages 685–690.

Yafei Wang, John Yen, and David Reitter. 2015. Pragmatic alignment on social support type in health forum conversations. In *Proc. Cognitive Modeling and Computational Linguistics (CMCL)*. Association for Computational Linguistics, Denver, CO, pages 9–18.

James T Webb. 1969. Subject speech rates as a function of interviewer behaviour. *Language and Speech* 12(1):54–67.