

# Investigating Linguistic Alignment in Collaborative Dialogue: A Study of Syntactic and Lexical Patterns in Middle School Students

Language and Speech

1–24

© The Author(s) 2024

Article reuse guidelines:

[sagepub.com/journals-permissions](https://sagepub.com/journals-permissions)

DOI: 10.1177/00238309241234565

[journals.sagepub.com/home/las](https://journals.sagepub.com/home/las)**Xiaoyi Tian** , **Amanda E. Griffith**

Department of Computer &amp; Information Science &amp; Engineering, University of Florida, USA

**Zane Price**

Tansa Systems; Department of Spanish and Portuguese Studies, University of Florida, USA

**Kristy Elizabeth Boyer**

Department of Computer &amp; Information Science &amp; Engineering, University of Florida, USA

**Kevin Tang** 

Department of English Language and Linguistics, Institute of English and American Studies, Heinrich Heine University Düsseldorf, Germany; Department of Linguistics, University of Florida, USA

## Abstract

Linguistic alignment, the tendency of speakers to share common linguistic features during conversations, has emerged as a key area of research in computer-supported collaborative learning. While previous studies have shown that linguistic alignment can have a significant impact on collaborative outcomes, there is limited research exploring its role in K–12 learning contexts. This study investigates syntactic and lexical linguistic alignments in a collaborative computer science–learning corpus from 24 pairs (48 individuals) of middle school students (aged 11–13). The results show stronger effects of self-alignment than partner alignment on both syntactic and lexical levels, with students often diverging from their partners on task-relevant words. Furthermore, student self-alignment on the syntactic level is negatively correlated with partner satisfaction ratings, while self-alignment on lexical level is positively correlated with their partner’s satisfaction.

## Keywords

Linguistic alignment, partner satisfaction, collaborative learning, middle school, pair programming, computer science education

---

## Corresponding author:

Kevin Tang, Department of English Language and Linguistics, Institute of English and American Studies, Faculty of Arts and Humanities, Heinrich Heine University Düsseldorf, Universitätsstr. 1, Düsseldorf 40225, North Rhine-Westphalia, Germany.

Email: [kevin.tang@hhu.de](mailto:kevin.tang@hhu.de)

## Introduction

Conversation plays a crucial role in human interaction, allowing us to communicate, collaborate, and build relationships. During conversation, linguistic *alignment*, also known as *entrainment*, *convergence*, or *adaptation* (Friedberg et al., 2012; Gallois et al., 2005; Lubold et al., 2021), arises when speakers (*interlocutors*) share common linguistic features, such as word choice or syntactic structure. Linguistic alignment is especially crucial in collaborative learning settings, where learners interact together and share their knowledge and skills to reach a specific learning goal (So & Brush, 2008). In this process, interlocutors exchange ideas constantly and construct shared mental models and mutual understanding of the task (Ludvigsen et al., 2018). This alignment process creates a “conceptual pact” where interlocutors establish a common way of referring to objects, which they utilize throughout their discourse (Brennan & Clark, 1996; Garrod & Pickering, 2004).

Linguistic alignment has been linked to various social and learning outcomes in collaborative contexts, such as task success (Friedberg et al., 2012; Hirschberg et al., 2008), interpersonal rapport (Duran et al., 2019; Madaio et al., 2018; Sinha & Cassell, 2015), and learning (Mitchell et al., 2012; Sinha & Cassell, 2015). One type of collaborative outcome is *peer satisfaction* between collaborative partners. In this study, *peer satisfaction*, or partner satisfaction, refers to the speaker’s satisfaction with the dynamics, cooperation, or outcome of a collaborative effort. This outcome plays an important role in collaborative tasks. Student groups with higher satisfaction can have better team dynamics (Ku et al., 2013) and task performance (Jung et al., 2012). One indicator of peer satisfaction is through alignment: when students speak more similarly to their partners, they achieve better outcome for the joint task (Giles et al., 1991; Nguyen et al., 2016). In another study, Sagi and Diermeier (2017) found that in a negotiation task, dialogue parties demonstrating higher linguistic alignment (although through a text-based communication channel) are more likely to reach an agreement. Angus et al. (2012) observed that in medical conversations, a greater alignment (accommodative communication) from the doctor is linked to increased patient satisfaction and better treatment outcome. However, other studies found no significant relationship between alignment and outcomes: a study by Schoot et al. (2016) showed that syntactic priming does not influence the perception by the conversation partner (although this is more related to a personal, affective outcome rather than a task-specific one); in another collaborative problem-solving context, lexical and syntactic alignment also did not lead to improved team performance but less semantic alignment predicts better performance (Duran et al., n.d.). These studies show that linguistic alignment can have different effects depending on the task and the communicative context. Our research aims to investigate the role of linguistic alignment and its relationship with partner satisfaction in K–12 collaborative computer science learning.

Linguistic alignment can occur at multiple levels of communication (Duran et al., 2019). Interlocutors have been found to adapt their prosodic cues such as speaking rate, pitch, and vocal intensity (Abel & Babel, 2017; Babel, 2012; Lubold et al., 2021), as well as the content of their speech such as sentence structure and vocabulary (Branigan et al., 2000; Brennan & Clark, 1996; Friedberg et al., 2012; Reitter & Moore, 2014). In this work, we are interested in linguistic alignment on the syntactic and lexical levels.

Syntactic alignment is the tendency to repeat a syntactic structure similar to a previously experienced sentence (J. Bock, 1986; Garrod & Pickering, 2004). For example, “We are good friends.” and “They are lovely pets.” are syntactically aligned, as both sentences share the same grammatical structure. Lexical alignment refers to the repetition of words between interlocutors. For example, Speaker A says, “How do I *delete* this?” and Speaker B responds, “Wait, *delete* what?” Speaker A then clarifies, “I want to *delete* this costume,” where the word *delete* is repeated by both speakers. Lexical alignment can be calculated in two ways: (1) *indiscriminate* (i.e., based on all words)

(Reitter & Moore, 2014), or (2) based on a selective set of words (Friedberg et al., 2012; Fusaroli et al., 2012). Studies have shown that task-relevant words may be strongly correlated with a positive collective benefit outcome, while general lexical alignment is negatively correlated with this result (Fusaroli et al., 2012). Similarly, Friedberg et al. (2012) found that lexical entrainment measured from project-related words reveals a significant difference between the high- and low-performance groups of undergraduate engineering students. In this study, we consider both types of lexical alignment, where we select a set of task-specific words related to the students' collaborative problem-solving tasks.

There are two main theoretical frameworks to explain the mechanisms behind linguistic alignment: **grounding** and **priming** (Rasenberg et al., 2020). These two mechanisms offer different perspectives on the role of conscious effort in the alignment process. Grounding refers to the deliberate and coordinated process in which interlocutors aim to establish mutual understanding and reach a shared belief about the meaning of what has been said (Brennan & Clark, 1996; Holler & Wilkin, 2011). Grounding in conversation is essential for effective communication, as it helps to avoid misunderstandings and promote mutual understanding. Ostrand and Chodroff (2021) found speakers aligned on some linguistic levels but not others (e.g., syntactic vs. phonetic) or aligned on partial linguistic features but not all features (e.g., temporal vs. global) within the same level. This flexible relationship between alignments at different linguistic levels can be guided by communicative and social factors. According to communication accommodation theory (CAT) (Giles, 2016), individuals align with their partner to achieve social approval. As conversation unfolds, each speaker's linguistic features may become more similar (convergent) or less similar (divergent). Interlocutors converge their linguistic features to reduce the social distance between themselves, making themselves appear more favorable and cooperative (Nguyen et al., 2016).

Priming, on the contrary, suggests that linguistic alignment occurs unconsciously and automatically (Pickering & Ferreira, 2008; Reitter et al., 2006). A notable theory in the context of priming and linguistic alignment is the interactive alignment model (IAM) by Garrod and Pickering (2004). The model posits an automatic alignment mechanism that enables speakers to reuse and expand on prior language patterns, simplifying mutual understanding.

Scholars argue different cognitive origins of priming, and research points to different origins for short- versus long-term priming. Chang et al. (2000) introduced the theory of implicit learning to explain long-term priming. Here, the idea is that after frequently encountering specific sentence patterns, individuals subconsciously favor those patterns even without recent exposure. Hartsuiker et al. (2008) and Reitter et al. (2011) outlined two mechanisms driving the two types of syntactic priming: spreading activation and base-level learning. The short-term version relies on recent linguistic experiences and the immediate accessibility of sentence structures, closely linked to the words recently used. This mechanism is called spreading activation. Alternatively, long-term priming is not anchored to specific lexical items but is rooted in implicit learning, as represented by Chang et al. (2000) and K. Bock et al. (2007). These two types of priming—short-term and long-term—also correlate with different outcomes. Reitter and Moore (2007) found long-term priming correlates with task success in task-oriented dialogue, whereas short-term priming does not.

In summary, both grounding and priming mechanisms are crucial in promoting effective communication and mutual understanding between interlocutors. However, different theories suggest different explanations for the importance of linguistic alignment. Grounding theories suggest a flexible relationship between behavioral alignment (in different modalities or linguistic levels) and alignment of conceptual representations (Giles et al., 1991; Giles & Ogay, 2007; Holler & Wilkin, 2011). On the contrary, priming theories posit a causal link between priming and alignment across multiple linguistic levels (Garrod & Pickering, 2004; Rasenberg et al., 2020; Reitter et al., 2006). Thus, if the processing benefits of alignment vary for lexical and syntactic structures, this would

be better explained by grounding theory, which posits that interlocutors intentionally converge to achieve effective communication and social outcomes. In contrast, if there is a consistent relationship between different levels of alignment and collaboration outcomes, the priming mechanism would provide greater explanatory power, as it suggests a universal relationship between perception and exhibited behavior.

While much research has focused on how the alignment of linguistic patterns relates to the quality of collaboration, it is important to note that speakers can also diverge from each other's language use, emphasizing differences in their communicative behaviors (Gallois et al., 2005). Divergence can occur in complex tasks to contrast views and update information, allowing interlocutors to make complementary and different contributions (Fusaroli et al., 2014; Mills, 2014). In everyday conversations, individuals deliberately diverge their choice of syntactic structures, offering reformulations of their partner's utterances and new ideas to further the discussion (Healey et al., 2014). Divergence of ideas has been recognized as a crucial factor in shaping collaborative interactions (Hoadley & Enyedy, 1999; Puntambekar, 2006). However, there has been limited research on linguistic divergence compared with convergence (Schegloff, 2007). The phenomenon of linguistic divergence is not universally evident and might depend on conversation types and task complexity. Further exploration is needed to understand this phenomenon.

### 1.1 Linguistic alignment and partner satisfaction outcome

In collaborative learning, linguistic alignment has been linked to collaboration outcomes, highlighting the importance of understanding the mechanisms underlying this phenomenon. For example, Gweon et al. (2013) found that "other-oriented" contributions (contributions based "on a prior contribution of a conversational partner") are important for knowledge integration during collaboration. Dillenbourg et al. (2016) found that partners completing tasks together modeled one another and that this modeling was predictive of outcomes such as task performance and learning gains. In peer-tutoring context, a tutor's lexical convergence metrics were a significant predictor of learning, mental effort, and emotional state (Mitchell et al., 2012). However, while most research on linguistic alignment focuses on adults (Friedberg et al., 2012; Rahimi et al., 2017; Reitter & Moore, 2014), its precise dynamics and relationship with collaborative outcomes are still understudied in children, who are known to have different sociolinguistic conventions than adults (Cook-Gumperz & Corsaro, 1977). Furthermore, social settings can affect how children express uncertainty, and age plays a role in the clarity of such expressions (Visser et al., 2014). Understanding how young learners converge or diverge from each other on multiple linguistic levels during collaborative problem-solving tasks can provide valuable insights into the relevance of linguistic alignment theories for a younger age group.

In this paper, we investigate linguistic alignment of young students with themselves (self-alignment) and with their collaboration partners (partner alignment) and the implications of self and partner alignment on peer satisfaction ratings following the collaborative problem-solving task. To do this, we use the transcribed spoken dialogues of middle school (seventh graders, aged 12–13) student pairs working on collaborative computer programming tasks. We focus on the syntactic and lexical levels of linguistic alignment by modeling the repetition of syntax structures and words within a certain dialogue window (10 utterances) from speakers themselves and from the collaboration partners.

We investigate the following research questions (RQs):

- RQ1: Do middle school students *syntactically* align with their partners and with themselves during collaborative problem-solving?

- RQ2: Do middle school students *lexically* align with their partners and with themselves during collaborative problem-solving?
- RQ3: How are syntactic and lexical alignments associated with the *collaboration outcome* of partner satisfaction?

Our study draws on prior linguistic alignment (e.g., Rasenberg et al., 2020; Reitter & Moore, 2014) and learning sciences research (Dillenbourg et al., 2016; Mitchell et al., 2012), suggesting that dialogue partners adopt each other's linguistic attributes. We anticipate this phenomenon will also be observed in students who engage in collaborative problem-solving tasks, with evidence of positive alignment on both syntactic and lexical levels in their utterances. In addition, prior collaborative learning research has linked linguistic alignment to positive outcomes such as task success and interpersonal rapport (Friedberg et al., 2012; Hirschberg et al., 2008; Madaio et al., 2018; Sinha & Cassell, 2015). Therefore, we hypothesize that the alignment on both syntactic and lexical levels will have a positive relationship with the outcome of partner satisfaction.

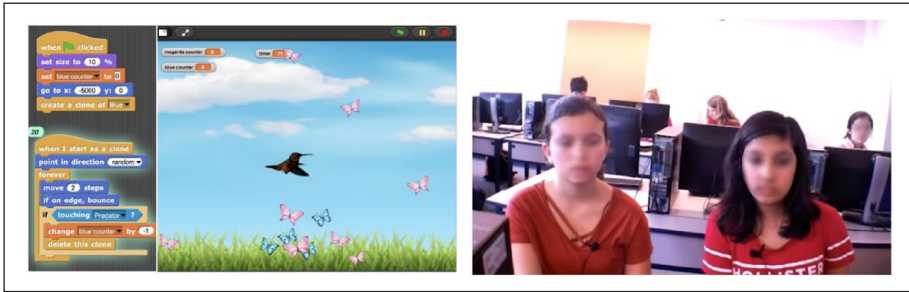
Our empirical findings reveal that students align more with themselves than their collaborative partners. Furthermore, we found that divergent dialogue on a subset of task-relevant words may benefit collaborative learning. Notably, students' self-alignment on both syntactic and lexical levels emerges as a significant predictor of the collaborative outcome of partner satisfaction. These findings highlight the importance of investigating self-alignment, divergent dialogue, and their impact on collaborative outcomes among K–12 students. Furthermore, our results contribute to the growing body of literature on linguistic alignment and offer insights into how this phenomenon operates in a developmental context, with implications for instructional practices that promote productive and satisfying collaborative interactions among young learners.

In the remainder of this article, we begin by describing the classroom computer science–learning study (middle school classroom pair-programming study), the calculation of linguistic alignment (linguistic alignment calculation), and the modeling method for linguistic alignment (modeling linguistic alignment and collaboration outcomes). The results and analyses are reported in Section 3. We conclude with a discussion of the theoretical and practical implications of the role of linguistic alignment in collaborative learning in Section 4.

## 2 Method

### 2.1 Middle school classroom pair-programming study

This study investigates the role of linguistic alignment during pair-programming tasks among middle school students. Pair programming is a popular collaborative learning approach in computer science education that involves two students working together on a programming task (Bowman et al., 2020; Campe et al., 2020; Denner et al., 2014). As collaboration is a critical component of pair programming, this context is ideal for investigating the role of linguistic alignment in collaborative problem-solving. The study was conducted in a science classroom at a public middle school (grades 6–8, aged 11–13) in the southeastern United States more than three consecutive semesters beginning Fall 2018. Students learned programming concepts such as variables, conditions, loops, and object-oriented programming, and then, they created programs collaboratively based on various science topics (e.g., light waves and evolution) with a block-based programming environment called *Snap!* (Figure 1). During study sessions, the researchers were present in person and facilitated classroom activities. The students were paired randomly by the teacher. During the pair-programming sessions, students were assigned a role to be either the “driver” (who types code) or the



**Figure 1.** Sample evolution activity created with the block-based programming environment Snap! by middle school students (left). A pair of students doing activity in class (right).

“navigator” (who observes and suggests) and were instructed to switch roles every 15 min. The average study session time was 31.22 min (standard deviation [ $SD$ ] = 4.36).

The study enrolled 75 middle school students (seventh grade, average age = 12.1). Participants included 46 girls (61.3%), 28 boys (37.3%), and 1 unspecified (1.3%). Race/ethnicities reported were White (46%), Hispanic (19%), Asian (16%), Multiracial (14%), Black (4%), and Other (1%).

**2.1.1 Corpus.** During the study session, both video and audio recordings were collected. The audio recordings were manually transcribed, following a detailed transcription protocol. Disfluencies such as filled and unfilled pauses, repetitions, self-corrections, interjections, and laughter were specifically noted during the transcription process. Other types of non-silence disfluencies such as sneezes, coughs, and lip smacks were not transcribed, as they were less pertinent to our study on interaction dynamics within collaboration pairs. Since the data were collected in a noisy classroom environment, some sessions were inaudible or had poor audio quality and were excluded. In addition, one group of students was highly uncooperative and disruptive. Because their session violates the collaboration assumption of the task, we exclude their data as well. Our analysis focuses on 24 pairs (48 speakers) of collaborative programming dialogues. The study was conducted across different semesters, and there were four students who were present in both the first and second rounds of pairings, each time with different partners and working on different tasks. In our analysis, we treat the repeated students in the second round of pairings as independent from their previous appearances, thus considering all speakers in such study sessions as independent ( $n = 48$ ). While this might risk the independence assumption of statistical tests, we believe the key interest of this study, linguistic alignment, is only meaningful in a social, interactive situation. It is not merely an individual behavior; rather, it is highly dependent on who the student is partnered with, and the specific learning tasks.<sup>1</sup>

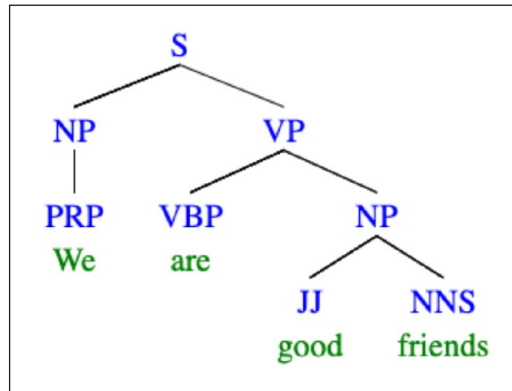
In addition, in a noisy classroom, students sometimes ask questions or chat with others nearby; thus, the original transcript contained utterances from the instructor and neighboring students. To investigate only the alignment between the collaborating pairs, we manually removed the dialogue excerpts that involved third-party speech and confirmed accuracy with video recordings. Our utterance unit in the corpus is a conversational turn, which is defined as speech by a single speaker without interruption by the interlocutor (Nakajima & Allen, 1993). After data cleaning, the corpus contained 5,148 turns, with a mean of 210.68 turns per conversation ( $SD = 99.29$ , median = 188, minimum = 83, and maximum = 456). The mean of word count in each turn is 9.78 ( $SD = 10.99$ , median = 7, minimum = 1, and maximum = 150).

**2.1.2 Partner satisfaction survey data.** After the pair-programming task, the participants completed a survey that assessed their satisfaction with their partners. To minimize the reciprocal effect,

Speaker Utterance	
S1	Wait.
S2	Okay. You do, you do it. We have to do a if.
S1	Yeah, yeah, yeah. I got it.
S2	Put the if inside the forever. Yeah.
S1	Uh. Let us see.
S2	There is another if and we have another if to get another if. Get in there. Okay, say two. Wait. Amplitude less than ... Okay, if amplitude is greater than 30, 30 say two bright for sunlight. Okay, If uh amplitude is less than 30.
S1	Amplitude is less than 30.
S2	Maybe.
S1	(laughs)
S2	Amplitude is less than 30.
S1	Okay
S2	Say two dark for 2 seconds. Okay, this is right now.
S1	Oops.
S2	Too dark.
S1	Too dark.
S2	Okay. Now my turn.
S1	Okay.
S2	Okay. And then if the amplitude is, is wait. If less than 80
S1	Wait we need one of these and one's.
S2	Yes
S1	Wait we need to put the if's and stuff. The other pieces too.
S2	And put this in ... There, right? And ... Is ... And ... And if amplitude is ... I'll like guess. Amplitude. Save. Save. Perfect.

**Figure 2.** A dialogue excerpt of students working on the task. *Target*, *Prime (self)*, and *Prime (partner)* turns are color-coded as dark gray, light gray, and white, respectively. S1 and S2 are the two speakers in the current dialogue.

students took the survey without sitting together with their partner. The survey contained questions such as “My partner was comfortable asking me questions” and “My partner did his or her fair share today.” Survey responses consisted of a 5-point Likert-type scale where 1 represented “strongly disagree” and 5 represented “strongly agree.” We computed an aggregated satisfaction score as the average across all items. Because a conversation is bi-directional, one person’s linguistic cues can also influence how their interlocutor perceives the person. Therefore, we considered both the student’s rating of their partner (satisfaction to partner) and the partner’s rating of the student respondent (satisfaction from partner). Of the 48 speakers, 46 completed the satisfaction



**Figure 3.** Syntax tree and sub-trees of an example sentence (“We are good friends.”).

survey. On average, students rated their partners with a satisfaction score of 4.15, with a *SD* of 0.69 and a median of 4.25.

## 2.2 Linguistic alignment calculation

To model the relationship between linguistic alignment and partner satisfaction, our next step is to quantify linguistic alignment. There are various methods to calculate linguistic alignment (Xu & Reitter, 2015), including the probability of co-occurrence of linguistic features (Dubey et al., 2005; Mukherjee & Liu, 2012), document similarity (Rahimi et al., 2017; Wang et al., 2014), or decay of repetition probability (Reitter et al., 2006). Fusaroli et al. (2012) proposed local linguistic alignment measurement (LLA), which measures how much a linguistic feature in the current conversational turn (called *target*) is repeated within a certain distance of the most recent preceding turns (called *prime*). Later Wang et al. (2014) implemented LLA on both the syntactic level (*SILLA*, *syntactic indiscriminate local linguistic alignment*) and lexical level (*LILLA*, *lexical indiscriminate local linguistic alignment*). In this paper, we use the same formula as illustrated by Wang et al. (2014), which calculated the number of words or syntactic rules that appear in both *prime* and *target*, normalized by the product of the number of syntax sub-trees or words in both *prime* and *target*. The construction of *target* and *prime* in our study dialogue is shown in Figure 2. We define *target* as the speaker’s current turn, for which its corresponding *prime* is the 10 most recent turns from either interlocutor (Xu & Reitter, 2015, 2018). Previous studies have specified the source of *prime* as comprehension–production prime (producing based on comprehension of partner’s speech) and production–production prime (producing based on speaker’s previously produced speech) (Reitter et al., 2010). Similarly, we divide *prime* into two sources: *prime(self)* and *prime(partner)*. *Prime(self)* is the 10 most recent turns from the speaker themselves, and *prime(partner)* is the 10 most recent turns from the interlocutor. We calculate self-alignment and partner alignment separately based on the two different *prime* sources. For simplicity of demonstration, in the following sections, we refer to the two sources together as *prime*.

## 2.3 Calculation of local syntactic alignment (SILLA) and lexical alignment (LILLA)

For the computation of SILLA, we first use the Berkeley Neural Parser (Kitaev et al., 2019; Kitaev & Klein, 2018) to construct the full syntax trees for each sentence in a conversational



turn (see Figure 3 for an example syntax tree). Then, we split the parsed syntax trees into sub-trees. For instance, if a turn were “We are good friends,” the full syntax tree would be (S (NP (PRP We)) (VP (VBP are) (NP (JJ good) (NNS friends))))).<sup>2</sup> The sub-trees are “S → NP + VP,” “NP → PRP,” “VP → VBP + NP,” and “NP → JJ + NNS.” We use these syntax sub-trees from *prime* and *target* to calculate SILLA. The formulae for calculating SILLA are shown in Equation (1).

To compute LILLA, we processed the corpus using standard text preprocessing techniques, including tokenization, lowercasing, stemming, and removing punctuation and stopwords. Given the programming task context, we did not remove programming-related stopwords (i.e., “if” and “for,” which were the two logical words involved in this particular task). Similar to the calculation of SILLA, instead of syntax sub-trees, we use individual words from *prime* and *target* to calculate LILLA. The formulae of LILLA are shown in Equation (2).

For each (*target and prime*) pair, we found the total number of overlapped sub-trees (or words) between *prime* and *target* and used it as the numerator. Then, we took the product of the number of syntax sub-trees or words in *prime* and *target* as the denominator. This yields an alignment score for each *target* turn, ranging from zero (no alignment) to one (perfect alignment). In this study, all alignment scores are less than one: a perfect alignment would not be possible given the large window size of *prime* we adopt.

Inspired by Fusaroli et al.’s (2012) observation that alignment on different lexical sets (all words or selective words) can lead to different outcomes, we calculated LILLA more than three categories: (1) *LILLA general*, the alignment on all words; (2) *LILLA task*, the alignment on task-relevant words; and (3) *LILLA non-task*, the alignment on words excluding task-relevant words. We selected the task-relevant word set manually. These words are directly related to the computer programming task (e.g., “wavelength,” “bright,” “dim,” “blue,” “visible,” “add,” “multiply,” “smaller,” and “greater”). We identified 153 distinct words that could reasonably be considered task-relevant. The task-relevant words are listed in Appendix task-relevant word list.

We compute the turn-level scores for the SILLA and three types of LILLA and then average them across conversation for a speaker. The descriptive statistics of these variables are shown in Table 7. These variables are used to construct model 3 to answer RQ3.

$$\begin{aligned} \text{SILLA}(target, prime) &= \frac{p(target | prime)}{p(target)} \\ &= \frac{\sum_{subtree_i \in target} \delta(subtree_i)}{\text{length}(prime) * \text{length}(target)} \end{aligned} \quad (1)$$

$$\delta(subtree_i) = \begin{cases} 1 & \text{if } subtree_i \in prime \\ 0 & \text{otherwise} \end{cases}$$

$$\begin{aligned} \text{LILLA}(target, prime) &= \frac{p(target | prime)}{p(target)} \\ &= \frac{\sum_{word_i \in target} \delta(word_i)}{\text{length}(prime) * \text{length}(target)} \end{aligned} \quad (2)$$

$$\delta(word_i) = \begin{cases} 1 & \text{if } word_i \in prime \\ 0 & \text{otherwise} \end{cases}$$

## 2.4 Modeling linguistic alignment and collaboration outcomes using Bayesian mixed-effect models

**2.4.1 Linguistic alignment.** To investigate the effect of self-alignment and partner alignment (RQ1 and RQ2), we build regression models to predict how likely a word (or syntax rule) produced in the *target* turn is a repetition from *prime* utterances. We adopt mixed-effect regression models to account for phenomena that occur on different levels. Because of the complex structure of linguistic data, Bayesian analysis is a popular approach over frequentist statistics (such as  $p$  and confidence intervals), as the latter either may not converge or result in uninterpretable results (Bates et al., 2015; Hoekstra et al., 2014; Morey et al., 2016). Thus, given the complex nature of our data and our RQs, we adopt Bayesian mixed-effects regression, fitted using the *brms* package (Bürkner, 2017) in *R* (R Core Team, 2017).

We build three regression models, each corresponding to one linguistic feature (syntax rules, general words, or task-relevant words). Because non-task and general word sets are highly overlapped, we only select general words between the two for the regression model. The selection of lexical features (general words and task-relevant words) is consistent with a prior study by Fusaroli et al. (2012). In these regression models, the unit of analysis is a unique linguistic item (a syntax rule or a specific word) that occurs in a turn spoken by one speaker. We refer to these linguistic items as a speaker-turn-word (or speaker-turn-syntax) rule. We counted how many times a unique linguistic item is present in *target* (*count target*) as the response variable. Because the linguistic item needs to be present at least once in *target* to be counted, the minimum value of *count target* is 1. Similarly, the two independent variables are constructed as how many times the word type or syntax rule that occurred in the *target* turn is present in the *prime self* and *prime partner* turn sets, which we refer to as *count prime self* and *count prime partner*. A word type that occurred in *target* does not necessarily exist in *prime*; thus, the minimum value for the two independent variables is 0.

We model the number of items (words or syntax rules) in the target turn (*count target*) as a function of the number of the same items in the prime (*count prime self* and *count prime partner*). Following standard practice in behavioral studies, we accounted for the idiosyncrasy of each participant by including *speaker* and conversational *turn* as random effects. Our random effects also include the conversational *group*, to account for the within-group dynamics, as well as the study *term* (the study semester), to account for the variations of classroom dynamics and different tasks we offered in different semesters. In addition, we include random slopes for the *group* and *speaker* random effects. In contrast to a model with only random intercepts—which allows different groups (such as participants, sessions, or experimental conditions) to have different baseline responses—a model with random slopes permits the effect of a predictor variable to vary across different levels of a random effects variable (Barr et al., 2013). This adds flexibility to the model, allowing it to potentially provide a better fit to our data by capturing additional sources of variability. Because the response variable *count target* is count-based data and its variability is relatively similar to its mean, a Poisson distribution is well-suited for the response variable. However, since *count target* has a range from 1 to 20, the Poisson distribution needs to be truncated to exclude the value of zero. Thus, we fit the model with a zero-truncated Poisson distribution. We used the default priors provided by *brms* in *R* (R Core Team, 2017). These are weakly informative priors that have minimal impact on the posterior estimates while still regularizing them (Bürkner, 2021). The mixed-effects regression structure is shown below in the syntax of *brms*.

$$\begin{aligned} \text{Count target} \mid \text{Trunc}(lb = 1) \sim & \text{Count prime self} + \text{Count prime partner} + \\ & (1 + \text{Count prime self} + \text{Count prime partner} \mid \text{Group}) + \\ & (1 + \text{Count prime self} + \text{Count prime partner} \mid \text{Speaker}) + \\ & (1 \mid \text{Word}) + (1 \mid \text{Term}) + (1 \mid \text{Turn ID})^3 \end{aligned}$$

To determine the importance of self-alignment in the collaborative learning context, we compare the models with full variable set and the model without self-alignment (*Count prime self*). The models are compared using the leave-one-out information criterion (LOOIC) and Watanabe–Akaike information criterion (WAIC) scores to measure the goodness of fit, which are common measures in Bayesian regression analysis (Vehtari et al., 2017). We found dropping self-alignment did not significantly improve the model fit for all three models; thus, we report the results from the full model (shown above) for all three linguistic features.

**2.4.2 Collaboration outcomes.** To investigate whether different types of linguistic alignment are associated with the outcome of partner satisfaction (RQ3), we adopt Bayesian mixed-effects regression, fitted using the *brms* package (Bürkner, 2017) in *R* (R Core Team, 2017) to estimate the impact of different types of linguistic alignment on the satisfaction outcomes of students.

We build two regression models, one for the speaker’s rating of the partner, *satisfaction to partner* (model 3–1), and one for the partner’s rating of the speaker, *satisfaction from partner* (model 3–2). In these regression models, our unit of analysis is the individual student. The independent variable is the self-alignment and partner-alignment scores on the syntactic level and lexical level of each student, quantified as the average SILLA, task LILLA, and non-task LILLA scores across all utterances by one speaker. Each model contains three fixed-effect predictor sets, each considering one linguistic feature: SILLA (*SILLA self and SILLA partner*), task LILLA (*task LILLA self and task LILLA partner*), and non-task LILLA (*non-task LILLA self and non-task LILLA partner*). Following standard practice in regression modeling, all the independent variables are *z*-score normalized (Baayen, 2008, Section 3.3), which allows us to compare the relative strength of the predictors directly. In addition to these fixed effects, we include by-term and by-group random intercepts to take into account task variability and within-group dynamics. In addition, we include random slopes for the *group* random effect.

Similar to the models for RQ1 and RQ2, we added random slopes into our model. The mixed-effects regression structure is shown below in the syntax of *brms* in *R* (R Core Team, 2017).

$$\begin{aligned} \text{Satisfaction} \sim & \text{SILLA own} + \text{SILLA partner} + \text{non-task LILLA own} + \\ & \text{non-task LILLA partner} + \text{task LILLA own} + \text{task LILLA partner} + \\ & (1 + \text{SILLA own} + \text{SILLA partner} + \text{non-task LILLA own} + \\ & \text{non-task LILLA partner} + \text{task LILLA own} + \text{task LILLA partner} | \text{Group}) + (1 | \text{Term})^4 \end{aligned}$$

## 3 Results

In this section, we report the regression results for the three RQs, whether students align with their partner and with themselves on a syntactic level and lexical level, and how do syntactic and lexical alignments correlated with partner satisfaction outcome.

### 3.1 RQ1: Do students align on the syntactic level?

The first RQ we asked is whether students align with their partner and with themselves on a syntactic level during the collaborative problem-solving. The data set consists of 40,774 speaker-turn-syntax rule observations from 5,148 utterances from 48 speakers. The description of the variables is shown in Table 1, and the model result is shown in Table 2.

Initially, we faced a challenge with the syntax model failing to converge. After examining possible issues with the model, we noted the predictors were highly skewed. Consequently, we implemented square root transformations on the predictors, *prime (self)* and *prime (partner)*. We favored

**Table 1.** Descriptive Statistics for Model 1: Syntactic Alignment Model. Each Row Is a Unique Syntax Rule in an Utterance,  $n = 40,774$ .

	<i>M</i>	<i>SD</i>	Median
Count target	1.246	0.742	1
Count prime self	2.427	4.079	1
Count prime partner	2.151	3.848	0

*M*: mean; *SD*: standard deviation.

**Table 2.** Bayesian Regression Model Result for Model 1: Syntactic Alignment Model.

	$\beta$	Estimated error	$CI_{l-95\%}$	$CI_{u-95\%}$	$\hat{R}$	PD (%)
Intercept	-4.035	0.177	-4.371	-3.682	1.001	100.0
Count prime self	0.152	0.020	0.112	0.191	1.000	100.0
Count prime partner	0.051	0.019	0.014	0.090	1.000	99.6

$\beta$  : coefficient;  $CI_{l-95\%}$  and  $CI_{u-95\%}$  : 95% credible interval;  $\hat{R}$  : Rhat; PD: probability of direction. All bulk and tail ESS (Effective Sample Size) values  $\geq 5,040$ .

the square root transformation over the log transformation, particularly because our predictors comprised zeroes, rendering them unsuitable for log transformation. It is important to note that, while square root transformations can lead to inconsistent treatment of values between 0 and 1 and those greater than 1 (Osborne, 2002), our predictors were integers and thus did not possess values between 0 and 1, so these concerns do not pertain to our variables. The square root transformation resulted in a notable reduction in skewness of the predictors. *Prime (self)* decreased from 2.42 to 1.02 and *prime (partner)* decreased from 2.15 to 0.929. After these adjustments, the models successfully converged.

The result shows a positive relationship between both predictors to the outcome. In our Bayesian analysis, we determine the strength of evidence using the probability of direction (PD), which varies between 50% and 100%. This measure, derived from the parameter's posterior distribution, indicates the probability of the parameter being either positive or negative (Makowski et al., 2019). We interpret a PD above 95% as strong evidence, between 75% and 94% as moderate evidence, and below 75% as weak evidence for the relationship. Syntactic structures produced in *prime (self)* are positively associated with the syntactic structures produced in *target* ( $\beta = 0.152$ , 95% credible interval set = 0.112–0.191, PD = 100.00%). Similarly, syntactic structures in *prime (partner)* are also positively associated with the same ones produced in *target* ( $\beta = 0.051$ , 95% credible interval set = 0.014–0.090, PD = 99.6%). The findings provide strong evidence that speakers converge syntactically both to themselves and to their partners as indicated by the PD, which is the proportion of the posterior distribution that shares the same sign as its median and approximately corresponds to the frequentist  $p$  (Makowski et al., 2019). The effect of self-alignment is higher than the effect of partner alignment.

### 3.2 RQ2: Do students align on the lexical level?

RQ2 asks whether middle school students align with themselves and with their partner on a lexical level. We divided this question into two subparts: lexical alignment for general words (model 2–1) and lexical alignment for task-relevant words (model 2–2).

**Table 3.** Descriptive Statistics for Model 2–1: General Lexical Alignment Model. Each row is a unique word type in an utterance,  $n = 20,966$ .

	<i>M</i>	<i>SD</i>	Median
Count target	1.161	0.536	1
Count prime self	0.886	1.592	0
Count prime partner	0.667	1.339	0

*M*: mean; *SD*: standard deviation.

**Table 4.** Bayesian Regression Model Result for Model 2–1: General Word Alignment Model.

	$\beta$	Estimated error	$CI_{l-95\%}$	$CI_{u-95\%}$	$\hat{R}$	PD (%)
Intercept	-2.356	0.198	-2.747	-1.877	1.002	100.00
Count prime self	0.081	0.014	0.054	0.108	1.000	100.00
Count prime partner	0.008	0.020	-0.033	0.045	1.000	66.7

$\beta$  : coefficient;  $CI_{l-95\%}$  and  $CI_{u-95\%}$  : 95% credible interval;  $\hat{R}$  : Rhat; PD: probability of direction.

\*All bulk and tail ESS values  $\geq 1,626$ .

**Table 5.** Descriptive Statistics for Model 2–2: Task-Relevant Lexical Alignment Model. Each Row Is a Unique Task-Relevant Word Type in an Utterance,  $n = 6,621$ .

	<i>M</i>	<i>SD</i>	Median
Count target	1.220	0.618	1
Count prime self	1.107	1.752	0
Count prime partner	0.880	1.513	0

*M*: mean; *SD*: standard deviation.

For the general word alignment model 2–1, the data set consists of 20,966 speaker-turn-word observations. The description of the variables is shown in Table 3. Table 4 shows the fixed effects of the Bayesian regression model. The result indicates there is a positive relationship between the number of general word types in the *prime (self)* utterance set and the count of the word being repeated in *target* ( $\beta = 0.081$ , 95% credible interval set = 0.054–0.108, PD = 100.0%). There is also a positive relationship between the number of general word types in the *prime (partner)* utterance set and one being repeated in *target*, although the evidence for the relationship is weak ( $\beta = 0.008$ , 95% credible interval set = -0.033 to 0.045, PD = 66.7%).

For the task-relevant word alignment model 2–2, the data set consists of 6,621 speaker-turn-word observations. This data set is a subset of data set for model 2–1, which uses only the task-relevant words. Descriptive statistics for the variables are shown in Table 5. Table 6 shows the fixed effects of the Bayesian regression model. The result indicates there is strong evidence for a positive relationship between the number of task-relevant word types in the *prime (self)* utterance set and the count of the word being repeated in *target* ( $\beta = 0.042$ , 95% credible interval set = 0.004–0.079, PD = 98.6%). Interestingly, the repeated task-relevant word count in *target* has moderate evidence of a negative relationship with the word count in *prime (partner)* ( $\beta = -0.027$ , 95% credible interval set = -0.087 to 0.027, PD = 82.8%).

**Table 6.** Bayesian Regression Model Result for Model 2–2: Task-Relevant Lexical Alignment Model.

	$\beta$	Estimated error	$CI_{l-95\%}$	$CI_{u-95\%}$	$\hat{R}$	PD (%)
Intercept	-1.799	0.247	-2.290	-1.208	1.002	100.0
Count prime self	0.042	0.019	0.004	0.079	1.001	98.6
Count prime partner	-0.027	0.029	-0.087	0.027	1.000	82.8

$\beta$  : coefficient;  $CI_{l-95\%}$  and  $CI_{u-95\%}$  : 95% credible interval;  $\hat{R}$  : Rhat; PD: probability of direction.  
All bulk and tail ESS values  $\geq 2,310$ .

**Table 7.** Descriptive Statistics of the Predictors for Model 3, Which Includes Mean, Standard Deviation, and Median of Participants' Averaged Alignment Scores.

	Partner alignment			Self-alignment		
	M	SD	Median	M	SD	Median
SILLA	7.61	2.16	7.31	8.25	3.13	7.12
LILLA task	37	14.17	35.75	37.6	11.57	33.72
LILLA non-task	8.14	1.99	7.9	9.4	3.33	8.5
LILLA general	6.42	1.47	6.15	7.35	2.57	6.44

M: mean; SD: standard deviation; SILLA: syntactic indiscriminate local linguistic alignment; LILLA: lexical indiscriminate local linguistic alignment.

For the purpose of display, each number is artificially multiplied by  $10^3$  (e.g., the actual mean value of SILLA partner alignment is 0.00761).

### 3.3 RQ3: How are syntactic and lexical alignments associated with collaboration outcomes regarding partner satisfaction?

General LILLA (lexical alignment on all words) is not included in the model due to its dependency with task LILLA and non-task LILLA variables, which are calculated based on a subset of all words. The descriptive statistics table is shown in Table 7 in *Calculation of local syntactic alignment (SILLA) and lexical alignment (LILLA)*. Following standard practice in regression modeling, all the independent variables are z-score normalized (Baayen, 2008, Section 3.3), which allows us to compare the relative strength of the predictors directly.

For model 3–1, predicting the outcome of *satisfaction to partner*, none of the fixed-effect predictors had strong evidence for having any relationship with the *satisfaction to partner* outcome. Therefore, we have discarded the model from further analysis.

However, for model 3–2, predicting the outcome of *satisfaction from partner*, the results in Table 8 provide strong evidence for a negative relationship with syntactic self-alignment ( $\beta = -0.507$ , 95% credible interval set =  $-1.115$  to  $0.129$ , PD = 94.7%) and a positive relationship with non-task lexical self-alignment ( $\beta = 0.547$ , 95% credible interval set =  $-0.052$  to  $1.126$ , PD = 96.5%).

## 4 Discussion

Our analyses reveal that the participants positively aligned with themselves. This finding is statistically significant on the syntactic level and the lexical level for both general and task-related words. In addition, the participants positively aligned with their partners at the syntactic level but not as

**Table 8.** Bayesian Regression Model Result for Model 3–2: Satisfaction From Partner Model.

	$\beta$	Estimated error	$CI_{l-95\%}$	$CI_{u-95\%}$	$\hat{R}$	PD (%)
Intercept	4.194	0.574	3.007	5.369	1.002	100.0
SILLA self	-0.507	0.313	-1.115	0.129	1.001	94.7
SILLA partner	0.081	0.220	-0.353	0.514	1.001	64.9
LILLA non-task self	0.547	0.294	-0.052	1.126	1.001	96.5
LILLA non-task partner	-0.008	0.228	-0.457	0.444	1.001	51.2
LILLA task self	0.021	0.159	-0.295	0.338	1.000	55.5
LILLA task partner	0.007	0.152	-0.293	0.304	1.000	52.4

$\beta$  : coefficient;  $CI_{l-95\%}$  and  $CI_{u-95\%}$  : 95% credible interval;  $\hat{R}$  : Rhat; PD: probability of direction; SILLA: *syntactic indiscriminate local linguistic alignment*; LILLA: *lexical indiscriminate local linguistic alignment*.  
All bulk and tail ESS values  $\geq 8,778$ .

much as they aligned with themselves. There was a tendency for participants to negatively align, or diverge, from their partner when measured at a task-specific lexical level.

It is not surprising to find that in this study, the participants positively aligned with themselves, and the degree of self-alignment was higher than the partner alignment. This result is consistent with the cognitive science literature: in fact, many studies have used self-alignment as a control baseline for analyzing the effect of partner alignment (Fusaroli & Tylén, 2016; Gweon et al., 2013; Healey et al., 2014; Reitter et al., 2010). Other literature has suggested that speakers self-align more than they partner-align in their task-oriented dialogue as a way of managing cognitive load (Dillenbourg et al., 2016; Schober, 1993). As Schober (1993) explains, because speakers know they can always correct misunderstandings with their partner after they emerge, they invest less effort in aligning their speech with their partner than they might in a less task-oriented setting. This allows the speakers to recruit more cognitive resources for the task at hand, rather than for producing a perfectly comprehensible utterance.

Another finding of this analysis differed from previous work: speakers tend to diverge on the task-specific lexical level, although this divergence is not statistically significant. While widespread findings of divergence in task-specific discourse are not evident, there have been some arguments around the functionality of divergence in different settings. In the general conversational context, Healey et al. (2014) found speakers syntactically diverged from their partner. They argue that linguistic divergence may have a conversation-advancement function in discourse. Successful communication may depend on selective repetition of certain words in different syntactic contexts to produce “contrasts, elaborations, and corrections that move a conversation forward” (Healey et al., 2014, p. 2). Similarly, in task-oriented dialogue, Fusaroli et al. (2014) proposed the notion of “dialog as interpersonal synergy” which emphasizes a good dialogue should afford complementary dynamics beyond simple alignment behaviors. Here, *complementary dynamics* means that speakers use alignment as an interactive resource, but do or say things differently to help the team reach coherence as whole (Fusaroli et al., 2014; Mills, 2014). In collaborative learning settings, divergence of ideas (such as raising new topics and asking questions) has been shown to significantly influence collaborative interactions (Puntambekar, 2006; Weinberger et al., 2007). A divergent perspective between students might require more collective information synthesis and negotiation of group understanding (Borge et al., 2018), which ultimately lead to collaborative knowledge construction (Puntambekar, 2006).

To illustrate this finding of partner divergence in our corpus, we present two dialogue excerpts to show how lexical divergence on task-related words could be associated with different group

**Table 9.** A Dialogue Excerpt of a High-Satisfaction Pair of Students Who “Diverged” With Each Other on Task-Relevant Words.

Speaker	Utterance
S1	No, let is switch. Like, on the other <b>variable</b> .
S2	What is the third part? Okay. This is what the <b>light difference</b> when it. wave <b>light vibrations second</b> party. At the <b>red</b> of the <b>spectrum to 380</b> .
S1	Should be the main screen. Let is switch this. Okay. Wait, this alone is. All right. Now, this time.
S2	I think so. So, like, if the <b>ratio</b> between those <b>two</b> . Then, the <b>color</b> should be <b>vibrant</b> .
S1	Choose the <b>wavelength</b> of each <b>color</b> . <b>Disconnect</b> the when. <b>Big, block</b> in the previous code.
S2	Like, is the <b>wave color</b> supposed to change? Is the <b>wave color</b> supposed to change?
S1	Um, I think it is supposed to be the <b>wave form</b> .

Task-relevant words are highlighted in bold. S1 and S2 are two speakers in the current dialogue.

dynamics and social outcomes. To contextualize the pair satisfaction score, we categorize them as “high” or “low” based on the average satisfaction score. Pairs with an average score above the overall mean of 4.15 are classified as “high” satisfaction pairs, while those below this threshold are considered “low” satisfaction pairs. Table 9 shows a dialogue excerpt where a “high” satisfaction pair (average score was 4.17, S1 received 4.5, and S2 received 3.83) began in a flow of collaboration but diverged on their task-relevant word choices. In this example, S1 focuses on the code implementation, such as switching screen or disconnect the “when” block, while S2 concentrates on the task description and parameter setting, such as setting wave light and color. This contrast indicates each party’s focus on different roles during a task. Both students seem to agree on their role, and they are completing the task efficiently.

While divergence could be a result of mutual agreement on the supplemental roles and sufficient situational understanding of the task, it could also result from unequal contributions between partners or disengagement from the task. In the following dialogue excerpt (Table 10), S2 focused on sorting out all the different colors, while S1 requested control over the mouse and later on argued with S2 about the color being “violet.” Given this apparent disagreement and divergence on the usage of task-related words, it is not surprising to see that this pair of students both reported low satisfaction with their partner (average score was 3.5, S1 received 3.33, and S2 received 3.67). These examples highlight the differences in the use of task-related words and their social outcomes. Further study is needed to continue to examine the dynamics of linguistic divergence on task-relevant words and its role in collaborative learning.

To address RQ3, which asks how linguistic alignment is related to satisfaction outcomes: the only significant predictor of a student’s partner’s satisfaction was the student’s linguistic alignment with *themselves*. We did not observe a significant effect of linguistic alignment with a partner on satisfaction outcomes. This finding contrasts the results in the work by Fusaroli and Tylén (2016), who found that interactive alignment (partner alignment) has a stronger effect on predicting coordinative performance than the baseline self-alignment. Considering the small sample size ( $n = 48$ ) of our study, it is possible that partner alignment still contributes to partner satisfaction, but the effect of self-alignment is so strong that it masks the effect of partner alignment. Below, we provide possible explanations on why self-alignment might potentially influence partner’s satisfaction.

First, to reiterate, self-alignment on non-task words is positively correlated with a partner’s satisfaction. Repetitive words might be encoded and processed more quickly by the listener.



**Table 10.** A Dialogue Excerpt of a Low-Satisfaction Pair of Students Who “Diverged” With Each Other On Task-Relevant Words.

Speaker	Utterance
S1	Can I have the <b>mouse</b> back?
S2	If control . . . gotta put the <b>forever</b> sign. How many are there? There are one . . . <b>one, two, three, four, five, six.</b> We need <b>six</b> more of these things.
S1	Woah! Could I have the <b>mouse</b> ?
S2	<b>Three, four, five. Five</b> right? How many <b>colors</b> are in the <b>rainbow</b> ? <b>Red, orange, green, yellow, blue, . Six.</b>
S1	<b>Gray . . .</b>
S2	G . . . G wait what?
S1	Wait, no, it is <b>seven!</b>
S2	Forget it. It does not even matter. If . . . if . . . okay now you can
S1	Okay.
S2	What the heck is wrong with you?
S1	Wait, no. Is not that supposed to be <b>violet</b> ?
S2	No this is too <b>dark.</b> Then next is <b>violet.</b> Then is <b>blue,</b> then is <b>orange,</b> then is . . . no, then is <b>green,</b> then is <b>yellow,</b> then is <b>orange,</b> then is <b>red.</b>
S1	Hold on. No, it is . . . <b>three hundred fifty</b> is <b>violet.</b>
S2	Yeah yeah but, here it says too <b>dark. Dark.</b> The <b>color</b> will be <b>dark.</b>
S1	No. <b>Three hundred eighty</b> to . . . So if it is <b>larger than seven hundred fifty</b> and <b>bigger . . .</b> So if it is <b>smaller . . .</b>
S2	Oh it is not say. It is we. It is will. The <b>color</b> will be <b>darker.</b> It is not say. I am so stupid. Think <b>change.</b>
S1	<b>Change</b> ghost effect . . . . . <b>color! Color color color.</b>
S2	Wait. Wait wait wait.

Task-relevant words are highlighted in bold. S1 and S2 are two speakers in the current dialogue.

Psycholinguistics research on lexical access reveals that locating a particular word may require activating a particular pathway in memory. Lexical access pertains to the mechanism by which words stored in our memory are recognized and produced during the act of language comprehension and production (Levett et al., 1999). On the contrary, semantic access deals with the retrieval of meaning associated with words, which is essential for understanding sentences and discourse. If the words are repeated, the searches down that path to the word might occur more rapidly than a new word (Levett et al., 1999; Pickering & Gambi, 2018). High self-alignment on non-task words could result in stronger priming for the listener. This priming, given the intertwined nature of lexical and semantic access, means that repeated or self-aligned words not only facilitate word recognition but also the retrieval of associated meanings. Consequently, it might require less time for the listener to process the word in their memory and lead to a faster response time, reducing the listener’s cognitive load and making the listener more satisfied with their self-aligned partner on the collaboration task.

The negative relationship between syntactic self-alignment and partner’s satisfaction possibly indicates that the processing benefits of repetition may differ for lexical versus syntactic structures. Alignment theories based on *grounding* (Holler & Wilkin, 2011) suggest that there is a flexible relationship between behavioral alignment (on various linguistic levels) and alignment of conceptual representations, in contrast to alignment theories based on *priming*, which argue that alignment percolates across multiple linguistic levels (Garrod & Pickering, 2004). Thus, alignment of

syntactic structures does not necessarily represent situated understanding in collaborative learning. In this age group, students who are more syntactically self-aligned are likely to be repeating very similarly structured sentences over time without responding to what their partner is saying, which may lead to low partner satisfaction. Another possible explanation for this negative relationship between syntactic self-alignment and partner's satisfaction could be that in collaborative dialogue, speakers frequently adjust their way of referring to an object: in other words, paraphrasing. Paraphrasing facilitates comprehension (Escudero et al., 2019). One particular paraphrasing method is to alternate the syntactic structure but keep the same semantic meaning (e.g., "Is there . . . Is that right?" uttered by one student during the task, then reiterated in a later turn by the same student as, "Is there c? I feel like that is not right."). Our finding indicates that a partner might favor a speaker who paraphrases more frequently and applies a more divergent set of syntactic structures. Further research needs to pursue a deeper understanding of the role of paraphrasing, as well as this relationship between syntactic self-alignment and social outcomes.

In our study, the distinct role of self-alignment might be attributed to adolescent linguistic development. While not always considered, as Nippold (2000) explains, adolescence is a pivotal phase in language development even though it may appear more subtle than linguistic development during the toddler or early school ages. Nippold et al. (2014) showed that during development, "adolescents can be expected to demonstrate considerable diversity on the use of complex syntax in spoken discourse." Middle school students straddle childhood and adulthood leading to a combination of linguistic experimentation and the reinforcement of learned patterns.

Transitioning from the linguistic development of adolescents, it is also important to consider the impact of external factors such as sociolinguistic and pragmatic development from an early age. For instance, Fish and Pinkerman (2003) demonstrated that children from low socioeconomic backgrounds displayed lower language skills compared with their peers from a control group as early as their entry into kindergarten. This shows how environment and socialization play a critical role in language development from a young age.

The self-alignment among our participants might reflect an effort to maintain a consistent linguistic identity. In contrast, adults, having established their linguistic styles, might be more inclined to align with their conversation partners, as demonstrated in the study by Fusaroli and Tylén (2016). In addition, as adolescents leave the home and family environment, they seek to form identities aligning with the wider social order leading to linguistic change and variation (Eckert, 2014). The notable self-awareness of identity, coupled with the quest for societal alignment, might elevate the role of self-alignment in our corpus.

Drawing this developmental perspective into our analysis, the observed linguistic behaviors in our corpus might not only be a product of the task at hand but also reflective of broader cognitive and sociolinguistic characteristics unique to adolescence.

This study has some limitations that need to be addressed. First, our sample size is relatively small, with only 48 students. Consequently, nonsignificant results regarding partner alignment on satisfaction should be interpreted with caution. Meanwhile, with our focus on syntactic and lexical alignments, we did not examine alignment at other linguistic levels (e.g., acoustic, phonetic, and semantic), and we did not investigate how alignment changes as conversation unfolds over time. Another point to consider is that we manually removed instances of third-party interventions, like those from teachers, which might have influenced the timing and content of linguistic input between pairs. Although this was accounted for in our statistical models, the potential variability introduced by such interruptions was not comprehensively studied. Future research should extend the scope of inquiry to understand linguistic alignment across multiple levels, study the dynamics of alignment over the course of conversations, and investigate the impact of collaborative outcomes further.

## 5 Implication and conclusion

Linguistic alignment holds great promise for understanding team dynamics and assessing team outcomes. Computer science tasks are cognitively demanding because they involve logically formulating and solving problems (Wing, 2014). There is a need to better understand the role of linguistic alignment and its relationship with collaborative outcomes in such complex tasks. In this study, we examine whether middle school students linguistically align with themselves and with their partners during a collaborative pair-programming task. We also investigate how the alignment is correlated with one collaboration outcome: partner satisfaction.

The findings of this article hold both theoretical and practical implications. Theoretical research on linguistic alignment and various collaboration outcomes is primarily conducted among adults. As we observed in our corpus, young learners exhibit different social dynamics than adult learners during collaborative learning tasks. Our findings suggest a stronger self-alignment than partner alignment, which is consistent with prior literature. While a majority of prior work has focused on the effect of partner alignment, the stronger effect of self-alignment on partner's satisfaction suggests a seemingly important functionality of self-alignment and undesirable partner alignment in this context, which warrants more exploration in the research community.

From a practical standpoint, understanding the level of linguistic alignment between youth learners and their collaborative partners can inform the design of artificial intelligence (AI)-enabled learning environments. Automated measures of linguistic alignment can provide real-time assessments of social presence and relationships among student groups, allowing instructors to adjust collaborative scripts accordingly. In addition, agents can leverage proper linguistic alignment strategies to enhance interactions with learners and potentially foster rapport. Further research is needed to explore the potential of linguistic alignment in collaborative learning contexts.

### Acknowledgements

The authors would like to thank the editor and the anonymous reviewers for their constructive feedback. They also thank Anna Sophia Stein and Akhilesh Kakolu Ramarao for their technical support with running the statistical models and Lydia Pezzullo for copy editing the manuscript.

### CRedit Authorship Contribution Statement

The authors follow the CRediT taxonomy (<https://credit.niso.org/>).

**Xiaoyi Tian:** Project administration, Conceptualization, Formal analysis, Investigation, Methodology, Visualization, Writing—original draft, and Writing—review & editing.

**Amanda E. Griffith:** Conceptualization, Data curation, Formal analysis, Methodology, Software, Writing—original draft, and Writing—review & editing.

**Zane Price:** Conceptualization, Data curation, Investigation, and Writing—original draft.

**Kristy Elizabeth Boyer:** Conceptualization, Resources, Supervision, and Writing—review & editing

**Kevin Tang:** Conceptualization, Formal Analysis, Investigation, Methodology, Supervision, Visualization, Software, and Writing—review & editing.

### Data Availability Statement


The data and code of the analyses can be found on an Open Science Framework repository: <http://doi.org/10.17605/OSF.IO/97SAK>.

### Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

## ORCID iDs

Xiaoyi Tian  <https://orcid.org/0000-0002-5045-0136>

Kevin Tang  <https://orcid.org/0000-0001-7382-9344>

## Notes

1. The results reported in this paper's "Results" section are qualitatively the same after excluding the four students who participated in both rounds.
2. S: sentence; NP: noun phrase; PRP: personal pronoun; VP: verb phrase; VBP: verb; JJ: adjective; NNS: plural noun. The full tagset is available at Penn Treebank (Santorini, 1990).
3. "Trunc(lb = 1)" means the response variable is zero-truncated. "(1 + Count prime self + Count prime partner | Group)" means that we included group as a random effect, with a random intercept and random slopes for the effects of Count prime self and Count prime partner. The other terms with the "(1 | X)" structure are included as random intercepts in the model.
4. "(1 + SILLA own + SILLA partner + non-task LILLA own + non-task LILLA partner + task LILLA own + task LILLA partner | Group)" means that we included group as a random effect, with a random intercept and random slopes for the effects of SILLA own, SILLA partner, non-task LILLA own, non-task LILLA partner, task LILLA own, and task LILLA partner. The other terms with the "(1 | X)" structure are included as random intercepts in the model.

## References

- Abel, J., & Babel, M. (2017). Cognitive load reduces perceived linguistic convergence between dyads. *Language and Speech, 60*(3), 479–502. <https://doi.org/10.1177/0023830916665652>
- Angus, D., Watson, B., Smith, A., Gallois, C., & Wiles, J. (2012). Visualising conversation structure across time: Insights into effective doctor–patient consultations. *PLOS ONE, 7*(6), Article e38014.
- Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge University Press.
- Babel, M. (2012). Evidence for phonetic and social selectivity in spontaneous phonetic imitation. *Journal of Phonetics, 40*(1), 177–189.
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*(3), 255–278. <https://doi.org/10.1016/j.jml.2012.11.001>
- Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). Parsimonious mixed models. *arXiv*.
- Bock, J. (1986). Syntactic persistence in language production. *Cognitive Psychology, 18*(3), 355–387. [https://doi.org/10.1016/0010-0285\(86\)90004-6](https://doi.org/10.1016/0010-0285(86)90004-6)
- Bock, K., Dell, G. S., Chang, F., & Onishi, K. H. (2007). Persistent structural priming from language comprehension to language production. *Cognition, 104*(3), 437–458.
- Borge, M., Ong, Y. S., & Rosé, C. P. (2018). Learning to monitor and regulate collective thinking processes. *International Journal of Computer-Supported Collaborative Learning, 13*(1), 61–92. <https://doi.org/10.1007/s11412-018-9270-5>
- Bowman, N. A., Jarratt, L., Culver, K., & Segre, A. M. (2020). Pair programming in perspective: Effects on persistence, achievement, and equity in computer science. *Journal of Research on Educational Effectiveness, 13*(4), 731–758.
- Branigan, H. P., Pickering, M. J., Stewart, A. J., & McLean, J. F. (2000). Syntactic priming in spoken production: Linguistic and temporal interference. *Memory & Cognition, 28*(8), 1297–1302.
- Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*(6), 1482–1493. <https://doi.org/10.1037/0278-7393.22.6.1482>
- Bürkner, P.-C. (2017). Brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software, 80*(1), 1–28. <https://doi.org/10.18637/jss.v080.i01>

- Bürkner, P.-C. (2021). Bayesian item response modeling in R with brms and Stan. *Journal of Statistical Software*, 100(5), 1–54. <https://doi.org/10.18637/jss.v100.i05>
- Campe, S., Denner, J., Green, E., & Torres, D. (2020). Pair programming in middle school: Variations in interactions and behaviors. *Computer Science Education*, 30(1), 22–46.
- Chang, F., Dell, G. S., Bock, K., & Griffin, Z. M. (2000). Structural priming as implicit learning: A comparison of models of sentence production. *Journal of Psycholinguistic Research*, 29, 217–230.
- Cook-Gumperz, J., & Corsaro, W. A. (1977). Social-ecological constraints on children's communicative strategies. *Sociology*, 11(3), 411–434. <https://doi.org/10.1177/003803857701100301>
- Denner, J., Werner, L., Campe, S., & Ortiz, E. (2014). Pair programming: Under what conditions is it advantageous for middle school students? *Journal of Research on Technology in Education*, 46(3), 277–296.
- Dillenbourg, P., Lemaignan, S., Sangin, M., Nova, N., & Molinari, G. (2016). The symmetry of partner modelling. *International Journal of Computer-Supported Collaborative Learning*, 11(2), 227–253. <https://doi.org/10.1007/s11412-016-9235-5>
- Dubey, A., Sturt, P., & Keller, F. (2005). Parallelism in coordination as an instance of syntactic priming: Evidence from corpus-based modeling. In *Proceedings of human language technology conference and conference on empirical methods in natural language processing* (pp. 827–834). Association for Computational Linguistics.
- Duran, N. D., Paige, A., & D'Mello, S. K. (2024). Multi-level linguistic alignment in a dynamic collaborative problem-solving task. *Cognitive Science*, 48(1), e13398.
- Duran, N. D., Paxton, A., & Fusaroli, R. (2019). ALIGN: Analyzing linguistic interactions with generalizable techniques—A Python library. *Psychological Methods*, 24(4), 419–438. <https://doi.org/10.1037/met0000206>
- Eckert, P. (2014). *Language and gender in adolescence. The handbook of language, gender, and sexuality*. Wiley.
- Escudero, I., Fuertes, N., & López, L. (2019). Paraphrasing strategy in EFL Ecuadorian B1 students and implications on reading comprehension. *English Language Teaching*, 12(1), 56–66. <https://doi.org/10.5539/elt.v12n1p56>
- Fish, M., & Pinkerman, B. (2003). Language skills in low-SES rural Appalachian children: Normative development and individual differences, infancy to preschool. *Journal of Applied Developmental Psychology*, 23, 539–565. [https://doi.org/10.1016/S0193-3973\(02\)00141-7](https://doi.org/10.1016/S0193-3973(02)00141-7)
- Friedberg, H., Litman, D., & Paletz, S. B. F. (2012). Lexical entrainment and success in student engineering groups. In *2012 IEEE Spoken Language Technology Workshop (SLT)* (pp. 404–409). <https://doi.org/10.1109/SLT.2012.6424258>
- Fusaroli, R., Bahrami, B., Olsen, K., Roepstorff, A., Rees, G., Frith, C., & Tylén, K. (2012). Coming to terms: Quantifying the benefits of linguistic coordination. *Psychological Science*, 23(8), 931–939. <https://doi.org/10.1177/0956797612436816>
- Fusaroli, R., Rączaszek-Leonardi, J., & Tylén, K. (2014). Dialog as interpersonal synergy. *New Ideas in Psychology*, 32, 147–157. <https://doi.org/10.1016/j.newideapsych.2013.03.005>
- Fusaroli, R., & Tylén, K. (2016). Investigating conversational dynamics: Interactive alignment, interpersonal synergy, and collective task performance. *Cognitive Science*, 40(1), 145–171. <https://doi.org/10.1111/cogs.12251>
- Gallois, C., Ogay, T., & Giles, H. (2005). Communication accommodation theory: A look back and a look ahead. In Gudykunst W. B. (ed.), *Theorizing about intercultural communication* (pp. 121–148). SAGE.
- Garrod, S., & Pickering, M. J. (2004). Why is conversation so easy? *Trends in Cognitive Sciences*, 8(1), 8–11. <https://doi.org/10.1016/j.tics.2003.10.016>
- Giles, H. (2016). *Communication accommodation theory: Negotiating personal relationships and social identities across contexts*. Cambridge University Press.
- Giles, H., Coupland, N., & Coupland, J. (1991). Accommodation theory: Communication, context, and consequence. In Giles H. (ed.), *Contexts of accommodation: Developments in applied sociolinguistics* (pp. 1–68). Cambridge University Press.

- Giles, H., & Ogay, T. (2007). Communication accommodation theory. In Bryan B. W., & Wendy S. (Eds.) *Explaining communication: Contemporary theories and exemplars* (pp. 293–310). Lawrence Erlbaum Associates Publishers.
- Gweon, G., Jain, M., McDonough, J., Raj, B., & Rosé, C. P. (2013). Measuring prevalence of other-oriented transactive contributions using an automated measure of speech style accommodation. *International Journal of Computer-Supported Collaborative Learning*, 8(2), 245–265. <https://doi.org/10.1007/s11412-013-9172-5>
- Hartsuiker, R. J., Bernolet, S., Schoonbaert, S., Speybroeck, S., & Vandereerst, D. (2008). Syntactic priming persists while the lexical boost decays: Evidence from written and spoken dialogue. *Journal of Memory and Language*, 58(2), 214–238.
- Healey, P. G. T., Purver, M., & Howes, C. (2014). Divergence in dialogue. *PLOS ONE*, 9(6), 1–6. <https://doi.org/10.1371/journal.pone.0098598>
- Hirschberg, J. B., Nenkova, A., & Gravano, A. (2008). High-frequency word entrainment in spoken dialogue. In *Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics on Human Language Technologies* (pp. 169–172). Association for Computational Linguistics.
- Hoadley, C. M., & Enyedy, N. (1999). *Between information and communication: Middle spaces in computer media for learning*. <https://repository.isls.org/bitstream/1/4324/1/330-343.pdf>
- Hoekstra, R., Morey, R. D., Rouder, J. N., & Wagenmakers, E.-J. (2014). Robust misinterpretation of confidence intervals. *Psychonomic Bulletin & Review*, 21(5), 1157–1164. <https://doi.org/10.3758/s13423-013-0572-3>
- Holler, J., & Wilkin, K. (2011). Co-speech gesture mimicry in the process of collaborative referring during face-to-face dialogue. *Journal of Nonverbal Behavior*, 35(2), 133–153. <https://doi.org/10.1007/s10919-011-0105-6>
- Jung, M., Chong, J., & Leifer, L. (2012). Group hedonic balance and pair programming performance: Affective interaction dynamics as indicators of performance. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 829–838). Association for Computing Machinery.
- Kitaev, N., Cao, S., & Klein, D. (2019, July). Multilingual constituency parsing with self-attention and pre-training. In *Proceedings of the 57th annual meeting of the Association for Computational Linguistics* (pp. 3499–3505). Association for Computational Linguistics. <https://doi.org/10.18653/v1/P19-1340>
- Kitaev, N., & Klein, D. (2018, July). Constituency parsing with a self-attentive encoder. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Vol. 1: Long papers)* (pp. 2676–2686). Association for Computational Linguistics. <https://doi.org/10.18653/v1/P18-1249>
- Ku, H.-Y., Tseng, H. W., & Akarasriworn, C. (2013). Collaboration factors, teamwork satisfaction, and student attitudes toward online collaborative learning. *Computers in Human Behavior*, 29(3), 922–929.
- Levelt, W. J., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22(1), 1–38.
- Lubold, N., Walker, E., & Pon-Barry, H. (2021). Effects of adapting to user pitch on rapport perception, behavior, and state with a social robotic learning companion. *User Modeling and User-Adapted Interaction*, 31(1), 35–73. <https://doi.org/10.1007/s11257-020-09267-3>
- Ludvigsen, S., Cress, U., Law, N., Stahl, G., & Rosé, C. P. (2018). Multiple forms of regulation and coordination across levels in educational settings. *International Journal of Computer-Supported Collaborative Learning*, 13(1), 1–6.
- Madaio, M., Peng, K., Ogan, A., & Cassell, J. (2018). *A climate of support: A process-oriented analysis of the impact of rapport on peer tutoring*. International Society of the Learning Sciences. <https://repository.isls.org/handle/1/909>
- Makowski, D., Ben-Shachar, M. S., Chen, S. A., & Lüdtke, D. (2019). Indices of effect existence and significance in the Bayesian framework. *Frontiers in Psychology*, 10, Article 2767.
- Mills, G. J. (2014). Dialogue in joint activity: Complementarity, convergence and conventionalization. *New Ideas in Psychology*, 32, 158–173. <https://doi.org/10.1016/j.newideapsych.2013.03.006>
- Mitchell, C., Boyer, K., & Lester, J. (2012). From strangers to partners: Examining convergence within a longitudinal study of task-oriented dialogue. In *Proceedings of the 13th annual meeting of the special interest group on discourse and dialogue* (pp. 94–98). Association for Computational Linguistics.

- Morey, R. D., Hoekstra, R., Rouder, J. N., Lee, M. D., & Wagenmakers, E.-J. (2016). The fallacy of placing confidence in confidence intervals. *Psychonomic Bulletin & Review*, 23(1), 103–123. <https://doi.org/10.3758/s13423-015-0947-8>
- Mukherjee, A., & Liu, B. (2012, December). Analysis of linguistic style accommodation in online debates. In *Proceedings of COLING 2012* (pp. 1831–1846). The COLING 2012 Organizing Committee. <https://aclanthology.org/C12-1112>
- Nakajima, S., & Allen, J. F. (1993). A study on prosody and discourse structure in cooperative dialogues. *Phonetica*, 50(3), 197–210. <https://doi.org/10.1159/000261940>
- Nguyen, D., Dođruöz, A. S., Rosé, C. P., & De Jong, F. (2016). Computational sociolinguistics: A survey. *Computational Linguistics*, 42(3), 537–593.
- Nippold, M. A. (2000). Language development during the adolescent years aspects of pragmatics, syntax, and semantics. *Topics in Language Disorders*, 20, 15–28. <https://doi.org/10.1097/00011363-200020020-00004>
- Nippold, M. A., Frantz-Kaspar, M. W., Cramond, P. W., Kirk, C., Hayward-Mayhew, C., & Mackinnon, M. (2014). Conversational and narrative speaking in adolescents: Examining the use of complex syntax. *Journal of Speech Language and Hearing Research*, 57, 876–886. [https://doi.org/10.1044/1092-4388\(2013/13-0097\)](https://doi.org/10.1044/1092-4388(2013/13-0097))
- Osborne, J. (2002). Notes on the use of data transformations. *Practical Assessment, Research, and Evaluation*, 8(1), 6.
- Ostrand, R., & Chodroff, E. (2021). It's alignment all the way down, but not all the way up: Speakers align on some features but not others within a dialogue. *Journal of Phonetics*, 88, 101074.
- Pickering, M. J., & Ferreira, V. S. (2008). Structural priming: A critical review. *Psychological Bulletin*, 134(3), 427–459.
- Pickering, M. J., & Gambi, C. (2018). Predicting while comprehending language: A theory and review. *Psychological Bulletin*, 144(10), 1002–1044.
- Puntambekar, S. (2006). Analyzing collaborative interactions: Divergence, shared understanding and construction of knowledge. *Computers & Education*, 47(3), 332–351.
- Rahimi, Z., Kumar, A., Litman, D. J., Paletz, S., & Yu, M. (2017). Entrainment in multi-party spoken dialogues at multiple linguistic levels. In *Proceedings of interspeech 2017* (pp. 1696–1700). <https://doi.org/10.21437/Interspeech.2017-1568>
- Rasenberg, M., Özyürek, A., & Dingemanse, M. (2020). Alignment in multimodal interaction: An integrative framework. *Cognitive Science*, 44(11), e12911. <https://doi.org/10.1111/cogs.12911>
- R Core Team. (2017). *R: A language and environment for statistical computing* [Computer software manual]. <https://www.R-project.org/>
- Reitter, D., Keller, F., & Moore, J. D. (2006). Computational modelling of structural priming in dialogue. In *Proceedings of the human language technology conference of the NAACL, companion volume: Short papers* (pp. 121–124). Association for Computational Linguistics.
- Reitter, D., Keller, F., & Moore, J. D. (2011). A computational cognitive model of syntactic priming. *Cognitive Science*, 35(4), 587–637.
- Reitter, D., & Moore, J. D. (2007). *Predicting success in dialogue*. Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics, pp. 808–815, Prague, Czech Republic, June 2007. c 2007 Association for Computational Linguistics.
- Reitter, D., & Moore, J. D. (2014). Alignment and task success in spoken dialogue. *Journal of Memory and Language*, 76, 29–46. <https://doi.org/10.1016/j.jml.2014.05.008>
- Reitter, D., Moore, J. D., & Keller, F. (2010). Priming of syntactic rules in task-oriented dialogue and spontaneous conversation. In *Proceedings of the 28th annual conference of the cognitive science society*.
- Sagi, E., & Diermeier, D. (2017). Language use and coalition formation in multiparty negotiations. *Cognitive Science*, 41(1), 259–271.
- Santorini, B. (1990). *Part-of-speech tagging guidelines for the Penn treebank project* (Tech. Report, MS-CIS-90-47). Department of Computer and Information Science, University of Pennsylvania.
- Schegloff, E. A. (2007). *Sequence organization in interaction: A primer in conversation analysis* (Vol. 1). Cambridge University Press.

- Schober, M. F. (1993). Spatial perspective-taking in conversation. *Cognition*, 47(1), 1–24. [https://doi.org/10.1016/0010-0277\(93\)90060-9](https://doi.org/10.1016/0010-0277(93)90060-9)
- Schoot, L., Heyselaar, E., Hagoort, P., & Segaert, K. (2016). Does syntactic alignment effectively influence how speakers are perceived by their conversation partner? *PLOS ONE*, 11(4), Article e0153521.
- Sinha, T., & Cassell, J. (2015). We click, we align, we learn: Impact of influence and convergence processes on student learning and rapport building. In *Proceedings of the 1st workshop on modeling interpersonal synchrony and influence* (pp. 13–20). Association for Computing Machinery.
- So, H.-J., & Brush, T. A. (2008). Student perceptions of collaborative learning, social presence and satisfaction in a blended learning environment: Relationships and critical factors. *Computers & Education*, 51(1), 318–336.
- Vehtari, A., Gelman, A., & Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, 27(5), 1413–1432. <https://doi.org/10.1007/s11222-016-9696-4>
- Visser, M., Kraemer, E., & Swerts, M. (2014). Children’s expression of uncertainty in collaborative and competitive contexts. *Language and Speech*, 57(1), 86–107. <https://doi.org/10.1177/0023830913479117>
- Wang, Y., Reitter, D., & Yen, J. (2014). Linguistic adaptation in conversation threads: Analyzing alignment in online health communities. In *Proceedings of the 2014 ACL workshop on cognitive modeling and computational linguistics* (pp. 55–62). Association for Computational Linguistics.
- Weinberger, A., Stegmann, K., & Fischer, F. (2007). Knowledge convergence in collaborative learning: Concepts and assessment. *Learning and Instruction*, 17(4), 416–426. <https://doi.org/10.1016/j.learninstruc.2007.03.007>
- Wing, J. M. (2014). Computational thinking benefits society. In *40th anniversary blog of social issues in computing* (p. 26).
- Xu, Y., & Reitter, D. (2015). An evaluation and comparison of linguistic alignment measures. In *Proceedings of the 6th workshop on cognitive modeling and computational linguistics* (pp. 58–67). Association for Computational Linguistics.
- Xu, Y., & Reitter, D. (2018). Information density converges in dialogue: Towards an information-theoretic model. *Cognition*, 170, 147–163. <https://doi.org/10.1016/j.cognition.2017.09.018>

## Appendix

### Task-relevant word list

30 15000 80 740 380 450 495 570 590 620 750 2 3 8 5 0 1 ten twenty thirty forty fifty sixty seventy eighty ninety hundred zero one two three four six seven eight nine ten eleven twelve thirteen fourteen fifteen sixteen seventeen eighteen nineteen double duplicate increase multiple none negative positive several single triple twice sprite code variable letter uppercase run program slider click left block value green flag if and loop first second random image copy keyboard mouse by forever change timer create blank arrow control down up entry than link save cloud big set edit add start move remove red stop never decrease counter merge combine x y pick button tab option name show wave main intensity color light amplitude speed frequency wavelength bright perfect dark visible violet blue green yellow orange red disconnect dim less larger between smaller greater too bird centimeters live reproduce break short die seconds long clone identical parent average size generation survive background length increase location less origin grow exponentially generate minus negative faster too old legacy fly children evolution different original spawned disappear simulation.