YeonJoo Jung* and Scott Crossley

Stylistic alignment in natural conversation involving second language speakers

https://doi.org/10.1515/applirev-2020-0048
Received June 1, 2020; accepted September 14, 2022; published online December 21, 2022

Abstract: To date, a growing body of second language (L2) research has investigated linguistic alignment as a pedagogical intervention, focusing on L2 learners' alignment behaviors in task-based interactions (e.g., Jung, YeonJoo, YouJin Kim & John Murphy. 2017. The role of task repetition in learning word-stress patterns through auditory priming tasks. *Studies in Second Language Acquisition* 39(2). 319–346; Kim, YouJin, YeonJoo Jung & Stephen Skalicky. 2019. Linguistic alignment, learner characteristics, and the production of stranded prepositions in relative clauses: Comparing FTF and SCMC contexts. *Studies in Second Language Acquisition* 41(5). 937–969). Linguistic alignment refers to a tendency where one speaker's utterances align with particular language features of those of the other speaker in dialogue. The current study investigated how L2 speakers' alignment behaviors differ in natural dialogues between L2-L1 and L2-L2 dyads in terms of language style (i.e., stylistic alignment) and the role of non-linguistic factors in the occurrence of stylistic alignment. The study analyzed a corpus of 360 texts using a computational tool. Results showed that stylistic alignment occurred to a greater extent in the L2-L2 dyad than in the L2-L1 dyad with respect to the word range, word frequency, word imageability, and proportion of bigrams proportion produced by the interlocutors. Furthermore, findings demonstrated the degree of stylistic alignment on each of the four selected lexical features was affected by numerous factors including age, group membership, nonnative speaker status, familiarity between interlocutors, and linguistic distance between L1 and L2. The effect of each factor on stylistic alignment in conversation is discussed in detail.

Keywords: computational analysis; language style matching; linguistic alignment; stylistic alignment; TAALES

***Corresponding author: YeonJoo Jung**, Department of English Education, Pusan National University, Busandaehak-ro 63beon-gil 2, Geumjeong-gu, Busan 46241, Republic of Korea, E-mail: yjjung@pnu.edu
Scott Crossley, Department of Applied Linguistics, Georgia State University, Atlanta, GA, USA, E-mail: scrossley@gsu.edu

1 Introduction

The past decades have witnessed a proliferation of psycholinguistic research investigating the cognitive mechanisms that underlie human language use (Vinson et al. 2016). The majority of this research has investigated the processing of isolated words, sentences, and texts in monologic situations where no interaction occurs (e.g., Levelt 1989). As a result, the most common form of language use (i.e., dialogue) has been largely ignored in psycholinguistic research because it does not fit classical assumptions of individual linguistic processing (Fusaroli and Tylén 2016). In response to this, researchers have proposed psycholinguistic models of dialogic processing in order to promote our understanding of speakers' linguistic behaviors during conversational interaction (Pickering and Garrod 2004). A notable example of such models is the interactive alignment model (IAM), according to which (interactive) alignment is referred to as the development of aligned linguistic representations via an automatic psycholinguistic priming mechanism that acts on every level of linguistic representation. Because linguistic alignment is caused by an automatic perception-behavior link, the likelihood that interlocutors engage in the same linguistic behaviors automatically increases through merely perceiving each other's linguistic behavior (Pickering and Garrod 2004) and social perceptions should not affect the degree of alignment (Krauss and Pardo 2004).

However, there has been a shift in psycholinguistic research as characterized by growing interest in examining the social factors that could affect linguistic alignment (speakers' attitudes, beliefs, and status in a relationship) (e.g., Branigan et al. 2010; Branigan et al. 2011; Weatherholtz et al. 2014). Proponents of this view suggest that linguistic alignment is essentially a socio-cognitive phenomenon whereby the interlocutors align their language at various levels of linguistic components in the course of interaction to achieve their goal of successful communication (Garrod and Pickering 2007).

Despite the increasing interest in the alignment behaviors of L1 speakers and the mediating effects of across various contexts such as task-based interactions and natural conversations, it appears that little effort has been devoted to addressing this issue in L2 discourse, particularly in naturalistic settings. This may be because most psycholinguistic models of alignment tacitly assume that both interlocutors are fully competent speakers of the language (i.e., L1 speakers) in which the conversation is carried out (Garrod and Pickering 2009). However, given that dialogue involving an L2 speaker is not uncommon, research has begun to explore whether linguistic alignment takes place in conversations where one of the interlocutors uses an L2 (L2-L1 dialogue) or where both interlocutors are using

the same target language (L2-L2 dialogue), and what factors may affect the occurrence of alignment in such dialogues (Costa et al. 2008; Purmohammad 2015). Researchers have speculated that L2-L1 dialogue may differ not only from L1-L1 dialogue but from L2-L2 dialogue possibly because L2 speakers' alignment behaviors may change according to a range of variables related to the interlocutor such as L2 speaker status, perceived and actual proficiency, and relationship with the interlocutor (Costa et al. 2008). However, empirical evidence is still rare and only few studies have attempted to test the assumptions (e.g., Purmohammad 2015). In the field of second language acquisition (SLA), most alignment research has been conducted within the interactionist perspective, focusing on L2 learners' alignment behaviors in task-based interactions (e.g., Jung et al. 2017; Kim et al. 2019; McDonough et al. 2015). There is a dearth of empirical research exploring how linguistic alignment between L2 speakers differs from that between L2 and L1 speakers in natural dialogue and what factors affect the magnitude of alignment.

To this end, the current study seeks to examine how speakers' alignment behaviors differ in natural dialogues between L2-L1 and L2-L2 dyads in terms of language style (i.e., stylistic alignment). Stylistic alignment refers to nonconscious verbal coordination that occurs when two people subtly match each other's speaking or writing style in the course of conversation. In addition, we were interested in the extent to which a range of non-linguistic variables would mediate the degree of stylistic alignment occurring in natural conversations involving L2 speakers. These variables included age and proficiency difference, gender combinations within the dyad, group membership, familiarity between interlocutors, and linguistic distance between L1 and L2.

2 Literature review

2.1 The phenomenon of linguistic alignment

During dialogue, speakers mutually adapt to each other's linguistic behaviors without explicit negotiation and/or control of the language they use (McDonough and Trofimovich 2009). Such alignment is pervasive in dialogue because it may contribute to communication success. Researchers have suggested a number of different accounts to explain the mechanisms underlying linguistic alignment in dialogue. One such account is the IAM (Pickering and Garrod 2004). According to the IAM, interlocutors converge on common language patterns in the course of interaction to arrive at a common situation model or to establish common ground, which in turn promotes successful communication. The central mechanism within the IAM that brings about aligned linguistic representations is automatic priming.

Priming refers to instances in which interlocutors arrive at aligned mental states signaled through the use of similar linguistic representations (Pickering and Garrod 2004). Linguistic alignment can occur at the lexical, syntactic, and phonological level as well as at the stylistic level. This alignment becomes evident during conversation when interlocutors adopt and repeatedly use each other's language patterns.

Past research has shown that interlocutors constantly demonstrate linguistic alignment in spoken interaction. Interlocutors not only re-use each other's words (lexical alignment; Brennan and Clark 1996) and grammatical structures (structural alignment; Branigan et al. 2010) but also converge on common phonetic realizations of words (Pardo 2006), and common accent and speech rate (phonological alignment; Giles et al. 1991). Alignment behaviors have also been observed at the stylistic level (i.e., stylistic alignment), which is the focus of the current study. Language style is generally defined as how an individual speaker/writer puts words together to communicate a message. People can communicate the same message in a variety of different ways. Language style captures these differences as style is indicated by features representing the author's choice of one mode of expression from a set of equivalent modes for a given content. This may be expressed by a wide variety of possible features of a text including choice of particular words (Bayram and Ta 2018).

Stylistic alignment occurs when two people in natural conversation nonconsciously match each other's speaking or writing style. Although people may naturally match their language styles to a certain degree in everyday dialogues, stylistic alignment is rarely detectable by the interlocutors (Ireland and Pennebaker 2010; Niederhoffer and Pennebaker 2002). Stylistic alignment is operationalized as language style matching (LSM), which is a dyad-level measure of automatic verbal coordination between speakers in their word use. LSM refers to the alignment of conversational partners' linguistic behavior and captures "the degree to which two people in a conversation subtly match each other's speaking style" (Ireland et al. 2011). LSM is based on the notion that the words one speaker uses can prime the other speaker to respond in a specific way during natural conversation (Niederhoffer and Pennebaker 2002), and that the words a speaker uses covary with those chosen by his/her interlocutor on both a turn-by-turn level and on the broader conversational level (Cappella 1996). Pennebaker and colleagues argued that similarity in the speakers' use of words (especially function words) can be used to quantify speakers' nonconscious language style coordination (Gonzales et al. 2010; Ireland et al. 2011; Niederhoffer and Pennebaker 2002).

Support for LSM has been reported in two lab-based experiments and a corpus-based study (Niederhoffer and Pennebaker 2002). In the lab experiments, participants carried out dialogues in a text-based computer-mediated communication (CMC) setting. The participants and outside raters judged the quality of the interactions. In

the corpus-based study, parts of the transcripts from the Watergate tapes in which President Richard Nixon had one-on-one discussions with his aides (i.e., face-to-face [FTF] conversations) were analyzed and the quality of the interactions was rated by human judges. Results demonstrated that speakers in dyadic interactions exhibited LSM at the conversation level as well as on a turn-by-turn level in both CMC and FTF contexts. However, a clear pattern was not found between LSM and the quality of the interaction perceived by both participants and judges. Based on their findings, Niederhoffer and Pennebaker suggested that interlocutors may unconsciously match their language patterns even when they do not like each other as long as they are engaged in conversation, indicating a positive relationship between the degree of stylistic alignment and engagement.

Most work on stylistic analysis has focused on sets of content-independent features such as function words, which are considered as robust markers of various individual differences and social behaviors, ranging from leadership style to honesty. Accordingly, scholars claim that LSM theoretically reflects interpersonal alignment across the array of psychological states that function words represent. However, earlier LSM and text analysis studies suggested that not only function words but also features relevant to content words have strong associations with interaction outcomes. For example, Taylor and Thomas (2008) found a significant relationship between LSM and successful negotiations, in that successful negotiations were associated with higher levels of LSM in word count, articles and prepositions (function words), features of content words such as causation (e.g., *because*, *effect*, *hence*), insight (e.g., *think*, *know*, *consider*), and positive emotion (e.g., *happy*, *pretty*, *good*). Furthermore, Argamon et al. (2007) also demonstrated the importance of including a novel set of functional lexical features for stylistic text classification in addition to the function word baseline. Examples of such functional lexical features include appreciation (e.g., *elaborate*, *consistent*, *profound*), affect (e.g., *joyful*, *miserable*), value (e.g., *probably*, *occasionally*). Such previous work suggests that LSM analyses on features relevant to both function and content words may provide a better understanding of the nature of stylistic alignment in conversational interactions.

2.2 Linguistic alignment and social factors

While the majority of previous research has focused on the cognitive mechanisms of alignment within a psycholinguistic framework (e.g., IAM), there have been increasing efforts to investigate the interaction of cognitive processes (e.g.,linguistic knowledge, attention, semantic access) and social factors (e.g., interlocutors' relative status and familiarity, speakers' age, gender, social status) in linguistic alignment. For

example, Branigan et al. (2010) proposed that speakers may choose to use a particular expression among several options, in a process of *audience design* (italics in the original; Bell 1984) because it is considered as the most appropriate expression to use for that particular interlocutor (Clark 1996). Clark (1996) suggested that speakers select linguistic expressions based in part on their beliefs about their interlocutors' speech communities and the information that members of those communities can access (Fussell and Krauss 1992). Speakers make judgments of their interlocutors' community membership on the basis of their linguistic knowledge (or proficiency) as well as direct personal experience with their interlocutors. These judgments are not static but change in the course of interaction as speakers may be continually adjusting their beliefs about the interlocutors throughout a dialogue (Branigan et al. 2010).

The influence of social factors on linguistic alignment has been evidenced in a number of studies. For instance, Balcetis and Dale (2005) found that speakers tended to structurally align with their interlocutor to a greater extent when the interlocutor acted affably. Alignment was also enhanced when the interlocutor was unpleasant about participating in the communicative task (i.e., when the mood was negative). The speaker perceived the negative mood as a signal that the conversation was in danger of failure and therefore allocated more resources to establishing rapport via augmented levels of alignment. These findings indicated that the degree to which speakers engage in alignment may depend on social factors, such as whether their interlocutor's social behavior invites an affiliative response and how the interlocutor's attitude towards the conversation is perceived by the speaker. Lev-Ari (2015) examined the role of prestige (operationalized as intelligence), similarity (i.e., speakers being similar in their level of ability to complete the task), and liking of the interlocutor in structural alignment of Dutch speakers. Results of the study showed that structural alignment was affected by prestige of the interlocutor and how much the interlocutor was liked by the speaker, indicating that the degree of alignment is modulated by social factors. Similarly, Weatherholtz et al. (2014) examined whether the magnitude of structural alignment is mediated not only by a particular grammatical structure that speakers are recently exposed to but also by social factors such as participants' perceptions towards socially different accents (standard US English, African-American English, and Mandarin accented English) and similarity in political ideology. Their findings demonstrated that while structural alignment occurred to some extent as a result of an automatic process, independent of speakers' social perceptions, the degree of alignment was indeed influenced by their perceptions towards different accents and interpersonal similarity in terms of political orientation.

Previous LSM research also lends support to an interaction between the degree of stylistic alignment and social factors. For instance, Gonzales et al. (2010) found a

positive relationship of LSM to group cohesiveness and task performance. In their study, participants were assigned to work in small groups on an information search task in either FTF or CMC mode. Findings showed that the more speakers liked their group members, the more their function words aligned during both FTF and CMC conversations. Furthermore, as speakers aligned in their word choice to a greater extent, they were more successful in completing the task.

2.3 Mediating factors in L2 alignment

To date, a substantial body of research has investigated speakers' alignment behaviors in L1 interlocutors who are likely process language in similar ways and who share relevant linguistic and cultural background knowledge (Garrod and Pickering 2009). Researchers have posited that there may be additional cognitive and social variables affecting linguistic alignment when both or either of the speakers is an L2 speaker. Thus, the basic alignment mechanisms proposed for L1-L1 dialogue is assumed to be different for dialogues involving an L2 speaker (Costa et al. 2008; Purmohammad 2015).

For instance, Beebe (1985) argued that L2 speakers' linguistic style-shifting (i.e., linguistic alignment and divergence) was a function of the social and psychological distance between interlocutors. The social distance becomes small when the interlocutors perceive themselves as being equal in terms of their identity, speech community membership, power, etc. In this case, speakers align in their language use with the speech norms of their interlocutors to emphasize solidarity and facilitate communication. On the other hand, the distance gets large when one of the interlocutor is in a superior position to the other. In this case, speakers shift in style away from their interlocutors (diverge) in order to assert the identity of their own social group.

Some research has focused on L2 speakers' alignment behaviors in a conversation with an L1 speaker (e.g., Costa et al. 2008; Purmohammad 2015). These studies suggest that failure to align or only partially align with L1 speakers on the part of L2 speakers may be due to incomplete knowledge of the target lexical items or structures. The degree of similarity or difference between the linguistic systems of a speaker's L1 and L2 may also affect the magnitude of alignment such that effects of automatic priming may be mediated by the extent to which the L2 learner's L1 and the target language diverge from each other (Costa et al. 2008). In contrast, there can be occasions where alignment is promoted in L2-L1 dialogue. Researchers have suggested that L2 learners adjust the amount of attention they pay to language form, which results in style-shifting or variation, depending on the identity and role of their interlocutor (Tarone 1988, 2007). For example, L2 learners

tend to produce a vernacular style of speech (unmonitored style) in a conversation wherein they focus more on meaning than on form (i.e., conversations with L2 peers). The vernacular style may include nonnative-like forms or informal, colloquial linguistic variants (Tarone 1983). However, when conversing with someone who they perceive as being in a superior position to themselves in terms of language knowledge and status, (e.g., teacher, researcher), L2 speakers may produce a more careful style because they tend to pay more attention to language forms. In this regard, L2 speakers are thought to devote conscious attention to the language forms produced by their L1 interlocutors during a conversation. This attention may lead to a stronger tendency to align with L1 speakers, compared to other L2 speakers, because L1 speakers may be seen as reliable sources of the target language and L2 interlocutors may believe that talking to an L1 speaker is the most efficient way to improve their L2 speaking abilities (Gass 2003). If L2 speakers repeat new words or expressions used by an L1 speaker, those new lexical items may later be available in their lexicon as a result of the repetition process (Purmohammad 2015).

L1 speakers may also adapt their utterances according to their interlocutor's proficiency level when conversing with an L2 speaker. This adaptation may involve a simplified form of language or 'foreigner talk' (Ellis 1997), which is characterized by the use of shorter and less complex structures, more high frequency words, less idiomatic expressions, and more repetition (Gass 2003). Foreigner talk is often accompanied by a corrective feedback (i.e., recasts) which are deliberate choices by an L1 interlocutor rather than a product of automatic priming. Features of foreigner talk such as recasts may contribute to the promotion of linguistic alignment, which may, in turn, facilitate L2 learning (Costa et al. 2008).

As reviewed thus far, several factors have been proposed to have a potential effect on the speakers' alignment behaviors in conversation where one or both of the speakers interact in an L2. However, to the best of my knowledge, very little research, if not none, has been carried out to empirically examine the alignment phenomenon, particularly in natural dialogues involving L2 speakers and the role of non-linguistic variables in the occurrence of alignment in such conversations. Furthermore, virtually no research has focused on the occurrence of stylistic alignment in L2 conversations.

2.4 The current study

The purpose of the current study is two-fold: (1) to investigate which lexical features speakers in L2-L2 dyads and in L2-L1 dyads align differently during natural conversations, as measured by the language style matching (LSM) metric

(Gonzales et al. 2010; Ireland et al. 2011; Niederhoffer and Pennebaker 2002) and (2) to examine to what extent non-linguistic variables mediate the degree of stylistic alignment operationalized as LSM. This study aimed to provide information about how lexical features may align differently between L2-L2 and L2-L1 interlocutors and the relationship between stylistic alignment and a range of cognitive and social non-linguistic factors. To accomplish this, a corpus of spoken language texts was analyzed for lexical features using a computational tool. Additionally, building on previous L1 and L2 research, a range of non-linguistic variables were included in the analyses to determine the mediating effect of the variables on the magnitude of stylistic alignment. The variables include age differences, proficiency difference, gender combinations within the dyad, group membership, familiarity between interlocutors, and linguistic distance between L1 and L2. The current study was guided by the following research question:

1. To what extent do L2-L2 and L2-L1 groups differ in terms of stylistic alignment occurring in natural conversation, as measured by language style matching (LSM) metric for lexical features related to the proportion of trigrams, the semantic diversity, word range, and word imageability?
2. To what extent do non-linguistic variables related to age differences, proficiency differences, gender pairings, familiarity between interlocutors, and linguistic distance between L1 and L2 explain variance in stylistic alignment between the two groups beyond lexical features?

3 Method

3.1 Spoken corpus

As reported in Crossley et al. (2012), the corpus used in this study comprised natural conversations between two interlocutors. Dyads were formed based on the interlocutor's status of being an L1 or L2 speaker: L2-L2 and L2-L1. A total of 50 L2 speakers were included in the L2-L2 group (20 males and 30 females) and the L2-L1 group comprised 50 speakers of which 25 were L1 speakers and 25 were L2 speakers (25 males and 25 females). The mean age of the participants was 23.82 (SD = 4.46) for the L2-L2 group and 27.31 (SD = 8.67) for the L2-L1 group. All participants were students at a large university in the United Sates at the time of data collection. The L1 speakers were all matriculated students enrolled in linguistics courses, whereas the L2 speakers were either matriculated students taking linguistics courses or ESL learners studying English at an Intensive English Program (IEP) housed at the university. In the L2-L1 dyads, all L2 speakers were IEP students. In the L2-L2 group, 25 were IEP students and 25 were matriculated international students. All

the matriculated students (whether L1 or L2 speakers) conversed with an IEP student in a naturalistic environment in which no tasks or topics were given and the conversations were not being monitored, though audio-recorded.

Each pair met at least twice, but as many as four meetings occurred (M = 3.02). In the L2-L2 dyads, speakers did not share the same L1. All of the L2 participants were at the high-beginning (n = 9), low-intermediate (n = 16), high-intermediate (n = 17) or advanced (n = 33) levels of language proficiency as classified by TOEFL tests (in the case of the matriculated L2 students in the linguistics class) or internal classification assessments (in the case of the L2 students at the IEP). The internal classification assessment used by the IEP center was a combination of the ACT Compass ESL reading and writing tests and internal tests of listening and speaking proficiency. All L1 participants were American, except for one participant from the United Kingdom. They were assigned as native speaker for their English proficiency. In order to form dyads in the two groups (L2-L1 and L2-L2) according to proficiency difference, each of the L2 participants was paired up with another participant either at the same proficiency level (e.g., advanced and advanced) or at different proficiency levels (e.g., low-intermediate and native speaker, high-intermediate and high-beginning). The configuration of the dyads in each group is as shown in Table 1 below:

The L2 speakers' L1 backgrounds varied and included Arabic, Japanese, Korean, Malay, Portuguese, Sri Lankan, Thai, French, Uzbek, and Spanish. Linguistic distance between participants' L1 and L2 (English) was determined based on Chiswick and Miller (2005). Linguistic distance refers to a scalar measure of the distance between English and other languages. This measure is based on the difficulty native speakers

Table 1: Configuration of dyads according to proficiency difference.

Configuration of dyads		L2-L2 (n = 25)	L2-L1 (n = 25)
Advanced	High-beginning	3	–
	Low-intermediate	7	–
	High-intermediate	6	–
	Advanced	2	–
High-intermediate	High-beginning	2	–
	Low-intermediate	3	–
	High-intermediate	2	–
Native speaker	Advanced	–	16
	High-intermediate	–	2
	Low-intermediate	–	5
	High-beginning	–	2

of American English may have in learning other languages. The corpus was collected over a five-year time span. All the conversations were recorded and a random excerpt from each recording was transcribed. The transcriptions were closely monitored by the researchers to make sure that they were all comparable in terms of word count and L2 usage. However, the researchers did not intentionally choose a part of each transcript for LSM analyses. The following parts of excerpts are representative samples that show the types of language present in conversations of the two dyads. Due to word limit constraints, only part of a conversation is presented for each dyad:

(1) L2-L1 dyad
L1 speaker (henceforth, L1): So, how are your classes?
L2 speaker (henceforth, L2): Today? some so so. I have some talking about that um, American history.
L1: Oh, you talked about American history.
L2: He, he asked something who I found something about that and he didn't want something about that and people went and going and blah blah blah. I, um, internet some stuff. Um, searching?
L1: Yes. You went on the internet and searched the stuff.
L2: Yes. But I'm very very… not well internet search.
L1: Oh, you're not very good at searching on the internet.
L2: Some people is, most people is, ah, five and ten minutes but I am one hour and two hour

(2) L2-L2 dyad
L2A: So, will you drive to there?
L2B: Yeah, but, I cannot drive, but my friend will drive, and we already rented our car.
L2A: Ah-ha, big van or.
L2B: Also, we are maybe eleven, so one van and one mini car.
L2A: Ah-ha. Wow, wow, interesting.
L2B: Yeah, but I heard that it will take more than twelve hours.
L2A: Yeah, and you should be careful, because the police there is very mean. I do not how to say that, because if you overspeed, they will give you a ticket and check very seriously.
L2B: Ah, really?
L2A: Yeah, so you, you should tell your friend to be careful.
L2B: I think they are terrible driver.
L2A: Yeah, their speed limit is seven, seventy mile.

3.2 Data preparation

To calculate LSM, each transcript was first segmented by speaker to produce two text files for each conversation. For example, in an L2-L1 dyad, two sub-transcripts would be created: an L2-L1 text that included only the L2 output used to analyze lexical features produced by an L2 speaker speaking to an L1 interlocutor and an L1-L2 text that included only the L1 output in order to assess the lexical features of an L1 speaker's utterances when talking to an L2 speaker. The transcripts of the L2-L2 dyad were also segmented into two sub-transcripts, each of which corresponded to the individual L2 speaker involved in the discourse. A total of 360 sub-transcripts were created for the subsequent language analysis and comprised the texts in the corpus. Descriptive statistics for the transcripts is displayed in Table 2:

Texts were then analyzed with the Tool for Automatic Analysis of Lexical Sophistication (TAALES; Kyle et al. 2018; Kyle and Scott 2015). TAALES reports on over 100 classic and recently developed lexical indices that focus on a number of areas of lexical sophistication including lexical frequency and range, n-gram frequency, range and proportion, word association strengths, and psycholinguistic word properties. Detailed description about the lexical sophistication indices included in TAALES is provided in the following section.

3.3 TAALES indices

3.3.1 Word frequency and range indices

Frequency is measured based on word counts taken from representative corpora (British National Corpus [BNC] 2007; Brown 1984; SUBTLEXus [Brysbaert and New 2009]). The frequency measure determines the reference corpus frequency of each word that occurs in a target text and creates an average frequency score for a text by dividing the sum of all frequency scores by the number of words in a text. Word frequency indices have been found to be predictive of holistic quality/proficiency scores in L1 and L2 writing research (Guo et al. 2013). Lastly, TAALES calculates

Table 2: Descriptive statistics for the transcripts in the corpus.

Dyad	Number of transcripts in the corpus	Number of words in the corpus	Mean number of words per transcript
L2-L1	176	75,138	426.920
L2-L2	184	77,131	419.190

range indices for words (Kyle and Scott 2015). These indices account for how widely a word or word family is used across texts by providing a count of the number of documents in which that word occurs. For this study, only texts from SUBTLEXus and the spoken portion of the BNC were selected. Also, only indices that were reported for all words (as compared to content or function words) were selected.

3.3.2 N-gram frequency and range indices

Frequency-based and range indices are also calculated for n-grams (e.g., bigram, trigram) in the written subsections of the BNC. N-gram refers to groups of two or more words that repeatedly appear in language as fixed items more frequently than expected by chance and much more frequently than phrasal verbs and idioms (Hyland 2008). TAALES also calculates proportion scores for n-grams. Proportion scores are a variant of frequency scores that treat n-gram identification as binary, reporting the proportion of bigrams and trigrams in a sample text that are also frequently found in the representative corpora. For this study, only indices for all words were selected.

3.3.3 Psycholinguistic properties of words

Psycholinguistic properties of words have been used to account for the variance in lexical decision times (e.g., Kuperman et al. 2012), lexical proficiency (e.g., Crossely et al. 2011a), speaking proficiency (e.g., Guo et al. 2013), and holistic scores of writing quality (e.g., Kyle and Scott 2015). These properties include concreteness, familiarity, imageability, meaningfulness, and age of acquisition. TAALES indices evaluate age of acquisition based on Kuperman et al.' s (2012) Age of Acquisition list and concreteness based on Brysbaert et al. (2014). Concreteness is rated based on perceptions of how simple it is to describe the meaning of a word. Familiarity scores are based on perceptions of how familiar words are to adults and correlate with frequency counts. Imageability scores indicate how easily a word can evoke a mental image and age of acquisition ratings are based on human judgments of the age at which a particular word is learned. For this study, only indices for all words were selected.

3.3.4 Word association norms

There are a number of ways to compute word associations, primarily using the statistical regularities observed in large reference corpora. For instance, TAALES

reports on two semantic distinctiveness indices reported by Hoffman et al. (2013) and McDonald and Shillcock (2001). The first approach examines semantic diversity (Sem-D; Hoffman et al. 2013) by measuring the variety of semantic contexts in which a word occurs. For instance, a word such as *one* that occurs in a variety of semantic contexts is more semantically ambiguous and less contextually distinct than a word occurring in constrained semantic contexts (e.g., puppy). The second approach measures contextual distinctiveness (McDonald and Shillcock 2001) by calculating the probability of a given word statistically co-occurring with others words in general language usage. As an example, *today* as compared to *lone* is used within a number of lexical contexts and is thus less contextually distinct. TAALES also reports association scores for individual words based on Latent Semantic Analysis, an unsupervised learning algorithm based on Singular Value Decomposition that can indicate association strength between words within a vector space (Landauer et al. 2007). For this study, only indices for all words were selected.

3.4 Language analysis

In this study, we used the LSM metric to operationalize stylistic alignment (Gonzales et al. 2010; Ireland et al. 2011; Niederhoffer and Pennebaker 2002) because of its relevance to the analysis of linguistic alignment in natural conversation. The LSM metric is an unobtrusive measure of non-conscious verbal coordination (Gonzales et al. 2010) that estimates the degree to which two speakers in a conversation match each other's linguistic style at the lexical level (Ireland et al. 2011). The focus of LSM is on stylistic alignment, as opposed to content accommodation. The LSM metric measures alignment of linguistic style in naturally occurring conversations whereas other available alignment measures (e.g., Local Linguistic Alignment; Fusaroli et al. 2012; Hierarchical Alignment Model; Doyle et al. 2016) analyze alignment that occurs during highly structured tasks designed to yield alignment by providing prime and target items. Also, unlike other alignment measures, LSM can be independently calculated for each linguistic feature. This is important because empirical evidence has suggested that different linguistic markers may have distinctively different alignments (e.g., Ireland et al. 2011). For example, usage of a second person pronoun such as *you* can will not demonstrate alignment since a speaker's *you* is actually the interlocutor's *me* in most instances (Doyle et al. 2016).

In the present study, separate LSM scores were calculated for the lexical indices of interest reported by TAALES. In the LSM formula, the absolute value of the difference between two speakers was divided by the total for each category. The resultant LSM

score was between 0 and 1, with scores closer to 1 reflecting high degrees of style matching. For example, in the case of the all word range index reported by SUBTLEXus (i.e., *SUBTLEXus Range AW*) between speaker 1 and 2, the calculation was as follows:

$$LSM_{\text{SUBTLEXus Range AW}} = 1 - \left[(|SUBTLEXus\ Range\ AW_1 - SUBTLEXus\ Range\ AW_2|)/ (SUBTLEXus\ Range\ AW_1 + SUBTLEXus\ Range\ AW_2 + 0.0001) \right]$$

SUBTLEXus Range AW_1 is the range score for all the words in a text based on the SUBTLEXus corpus in the speaker's transcripts and *SUBTLEXus Range* AW_2 is the score of the same index in his/her conversation partner's transcript. In the denominator, 0.0001 is added to prevent null values. Each LSM score is bounded by 0 and 1, with higher number representing stronger alignment in the use of particular lexical feature of the two interlocutors.

3.5 Statistical analysis

We first conducted a one-way multivariate analysis of variance (MANOVA) to examine which of our 20 selected lexical indices from our four lexical categories (word frequency and range, N-gram frequency, range, and proportion, psycholinguistic properties of words, and word association norms) demonstrated differences between L2-L1 conditions and L2-L2 dyads. This analysis was used to answer our first research question. The dependent variables for the MANOVA were LSM scores and the independent variable was group membership: L2-L1 or L2-L2. A Bonferroni correction was applied to control for multiple comparisons. The index with the highest effect size from each category was selected for analysis. Prior to analyses, data were screened to ensure that statistical assumptions were not violated. First, visual inspection of the data was performed to identify and remove indices that were not normally distributed. All variables demonstrated normal distributions. Multicollinearity was also checked to ensure that none of the preliminary LSM scores for the 20 lexical indices (henceforth, LSM variables) correlated at $|r| \geq 0.90$. If two or more indices correlated at this rate with each other, only one index was retained based on theoretical and statistical considerations for the subsequent statistical analysis and the other indices were removed. Five variables were removed after the multicollinearity check. After checking the statistical assumptions, a total of 15 indices remained.

To answer the second research question, we employed LME models with the lexical features selected from the MANOVA. Unlike traditional repeated measures analyses of variance or regression models, an LME analysis uses a combination of

fixed and random effects to predict a dependent variable. Fixed effects are treated as predictor variables (i.e., independent variables) much like they would in regression analyses. However, LMEs also include random effects in order to account for variation attributed to individual differences amongst participants and items. As such, LMEs allow for more accurate interpretations of the influence of specific effects when attempting to measure the influence of a predictor variable on an outcome variable (Baayen et al. 2008). We used R (R Development Core Team 2015) with the *lme4* (Bates et al.2015) and *lmerTest* (Kuznetsova et al. 2015) packages in order to build our LME models. Additionally, we used the *MuMIn* package (Nakagawa and Schielzeth 2013) to calculate effect sizes based on marginal and conditional R^2. Marginal R^2 is the variance explained by the fixed factors alone (i.e. predictor variables), whereas conditional R^2 is the variance explained by the fixed and random factors. The outcome variables were LSM variables that significantly differentiated between L2-L1 and L2-L2 groups in the first analysis. The fixed effects included non-linguistic factors that we interpreted as social including group membership (L2 and L1 speaker combinations in the dyad), and familiarity operationalized as order of meetings (with the first meeting as the baseline) and non-social factors including language proficiency differences between the speakers in the dyad (e.g., an advanced speaker interacting with a beginning speaker), age differences between the speakers within the dyad, and gender combinations within the dyad (with the female-female dyad as the baseline).

4 Results

The MANOVA results showed that there was a significant effect of group membership on the magnitude of stylistic alignment, Pillai's trace = 0.178, $F(15, 151) = 2.178$, $p = 010$, partial $\eta^2 = 0.178$. The significant main effect for group membership observed in the MANOVA was followed up with subsequent ANOVA analyses with Bonferroni corrections in order to identify the variables for which group differences existed. Results showed that the source of the significance for group membership was the difference between the L2-L1 and L2-L2 groups in their LSM scores for the four lexical markers. The four indices include semantic diversity (Sem-D), the percentage of trigrams in the spoken portion of the BNC (BNC Spoken Trigram Proportion), word range based on the SUBTLEXus corpus (SUBTLEXus Range), and word imageability based on the MRC database (MRC Imageability). Table 3 displays descriptive statistics for the selected 4 indices in the two dyads (i.e., L2-L2, L2-L1). LSM scores of the indices are listed for each dyad in Table 4 and Table 5 presents a summary of the individual comparisons for each index. The results indicate that greater style matching was found in the L2-L2 dyads than in the L2-L1 dyads.

Table 3: Descriptive statistics for the selected indices.

Dyad	L2-L2		L2-L1	
Index	***M***	**SD**	***M***	**SD**
Sem-D	2.038	0.044	2.036	0.068
BNC Spoken Trigram Proportion	0.163	0.054	0.177	0.068
SUBTLEXus Range	6,322.989	429.779	6,391.862	448.210
MRC Imageability	347.423	16.334	339.988	19.466

Table 4: Average LSM scores for the selected indices.

Dyad	L2-L2		L2-L1	
Index	***M***	**SD**	***M***	**SD**
Sem-D	0.989	0.009	0.982	0.016
BNC Spoken Trigram Proportion	0.822	0.141	0.744	0.170
SUBTLEXus Range	0.974	0.023	0.960	0.034
MRC Imageability	0.979	0.016	0.969	0.024

Table 5: Results of ANOVA analyses.

Index	*F*	*p*	partial η^2
Sem-D	12.653	<0.001	0.071
BNC Spoken Trigram Proportion	10.35	0.002	0.059
SUBTLEXus Range	9.079	0.003	0.052
MRC Imageability	8.949	0.003	0.051

To answer our second research question, we constructed four LME models with the dependent variable in each model the LSM score that was significant in our MANOVA. The LME tested whether the degree of stylistic alignment on each of the selected four lexical features was affected by non-linguistic factors such as age differences, gender combinations, familiarity between interlocutors (i.e., meeting order), proficiency differences, and group membership (L2-L2 or L2-L1). For each model, we entered each LSM score as the dependent variable, all of the factors listed above as fixed effects, and participants as a random effect.

Table 6: LSM on semantic diversity index.

Fixed Effect (baseline)	Coefficient	St. Error	*t*	*p*
(Intercept)	0.989	0.004	235.748	0.000
Group (L2-L1)	−0.007	0.003	−2.560	0.014

The level in parentheses was chosen a baseline, and all levels are compared against that baseline.

The LME model testing whether stylistic alignment in the semantic diversity index (Sem-D) was affected by the selected variables found that only group membership had a significant effect, with the L2-L2 group having a higher LSM score (i.e., greater alignment). A final model using only this significant predictor was computed. The model reported a marginal R^2 of 0.068 and a conditional R^2 of 0.358. Table 6 reports the intercept and fixed effect, along with their coefficients, standard error, *t* values, and *p* values. The results indicate that L2-L2 dyads reported greater alignment scores for the semantic diversity index.

An LME model testing whether stylistic alignment in the trigram proportion index (BNC Spoken Trigram Proportion) was influenced by the selected variables found significant effects for proficiency differences and effects that approached significance for group membership. A second model was calculated with only proficiency differences and group membership. In the second model, both proficiency differences and group membership reported significant effects indicating that dyads with lower differences in proficiencies and L2-L2 dyads reported greater alignment scores for the trigram proportion index. The final model reported a marginal R^2 of 0.173 and a conditional R^2 of 0.281. Table 7 reports the intercept and fixed effects, along with their coefficients, standard error, *t* values, and *p* values.

An LME model testing which factors affected the degree of stylistic alignment in the word range index (SUBTLEXus Freq AW) found that group membership, proficiency differences, and meeting order had significant effects on the LSM score. A final model was calculated using only these variables. This model reported a marginal R^2 of 0.163 and a conditional R^2 of 0.374. Table 8 reports the intercept and fixed effect, along with their coefficients, standard error, *t* values, and *p*

Table 7: LSM on BNC spoken trigram proportion index.

Fixed Effect (baseline)	Coefficient	St. Error	*t*	*p*
(Intercept)	0.903	0.027	34.034	< 0.001
Group (L2-L1)	−0.058	0.027	−2.165	0.036
Proficiency differences	−0.054	0.013	−4.205	< 0.001

The level in parentheses was chosen a baseline, and all levels are compared against that baseline.

Table 8: LSM on SUBTLEXus range index.

Fixed Effect (baseline)	Coefficient	St. Error	*t*	*p*
(Intercept)	0.979	0.006	155.443	0.000
Group (L2-L1)	–0.010	0.006	–1.799	0.079
Proficiency differences	–0.008	0.003	–3.142	0.003
Meeting 3 (Meeting 1)	0.011	0.005	2.073	0.040
Meeting 4 (Meeting 1)	0.016	0.006	2.838	0.005
Meeting 4 (Meeting 2)	0.025	0.006	2.286	0.024

The level in parentheses was chosen a baseline, and all levels are compared against that baseline.

values. The results indicated that L2-L2 dyads reported greater linguistic alignment as did dyads that reported less difference in proficiency in terms of the range of word produced. In addition, as speakers became more familiar with their partners, their alignment increased. Post-hoc analysis of the familiarity variable indices significant differences between the first and third meetings, the first and fourth meeting and the second and fourth meeting.

An LME model testing whether the degree of stylistic alignment in the word imageability index (MRC Imageability AW) found significant effects for group membership, proficiency difference, and familiarity. A second model was ran using only these variables. In the second model, the group variable only approached significance. As a result, a third model with only proficiency differences and familiarity was calculated. This model reported a marginal R^2 of 0.136 and a conditional R^2 of 0.370. Table 9 reports the intercept and fixed effect, along with their coefficients, standard error, *t* values, and *p* values. The results indicated that dyads with smaller differences in language proficiency reported greater alignment in terms of word imageability. In addition, as speakers became more familiar with their partners, their imageability alignment increased. Post-hoc analysis of the familiarity variable indicated significant differences between the first and third meetings, the first and fourth meeting, and the second and fourth meeting.

Table 9: LSM on the MRC imageability index.

Fixed Effect (baseline)	Coefficient	St. Error	*t*	*p*
(Intercept)	0.975	0.006	160.328	0.000
Proficiency differences	–0.009	0.003	–3.410	0.001
Meeting 3 (Meeting 1)	0.011	0.005	2.129	0.035
Meeting 4 (Meeting 1)	0.016	0.006	2.856	0.005
Meeting 4 (Meeting 2)	0.013	0.006	2.300	0.023

The level in parentheses was chosen a baseline, and all levels are compared against that baseline.

5 Discussion

The purpose of the current study was two-fold: (1) to investigate the degree to which interlocutors align their language styles during natural conversations; and (2) to explore the mediating effects of a range of non-linguistic factors on the magnitude of stylistic alignment. Operationalized as LSM, stylistic alignment was quantified using the LSM metric. LSM scores on various lexical indices were analyzed in order to determine whether the lexical indices were aligned differently between the L2-L1 and L2-L2 dyads during natural conversations and to examine to what extent non-linguistic factors mediated the magnitude of stylistic alignment.

First, this study examined whether the degree of stylistic alignment differed according to group membership (L2-L2 and L2-L1) in naturally occurring conversations. The findings for lexical indices related to semantic diversity, word range, trigram proportion, and word imageability demonstrated differences in stylistic alignment between the groups. As shown in Table 4, the LSM scores for the four indices were higher for the L2-L2 group than for the L2-L1 group, demonstrating that L2 speakers were more likely to stylistically align with their L2 interlocutors than with L1 interlocutors in their choice of words related to semantic diversity, word range, trigram proportion, and word imageability. The different degree of stylistic alignment between L2-L1 and L2-L2 dyads may be attributed to the language style that L2 learners produced during conversation. Learners' interlanguage can be conceptualized as a set of speech styles including the vernacular style and the careful style (Tarone 1983). It is likely that when a learner is focused on meaning in a natural conversation, the vernacular style is produced without the learner's conscious awareness of language form whereas the careful style is elicited when the learner shifts attention from meaning to form to complete a task, for example, in a classroom setting (Adamson and Regan 1991; Tarone 2007). Given that in this study conversations were made between participants in an informal, naturalistic setting, we assume that the L2 speakers were focused more on meaning than form, aligning in the vernacular style of speech in their interlanguage. Furthermore, greater alignment in L2-L2 dyads may also indicate the L2 speakers' desire to identify with the communicative norms of the learner group rather than those of the L1 group (Giles et al. 1991).

Two of the four lexical features were related to word range and n-gram proportion (SUBTLEXus Range and BNC Spoken Trigram Proportion). These indices measure word frequency either based on the range of texts within a corpus that a word occurs (SUBTLEXus Range) or based on the proportion of common n-grams in a reference corpus (BNC Spoken Trigram Proportion). Another significant index was related to semantic diversity (SEM-D) and the final variable was

related to word imageability (MRC Imageability). The results indicated that L2 speakers, when conversing with an L1 speaker, tended to produce words with a wider range, a higher proportion of trigrams commonly used in spoken language, more semantically distinct words, and less imageable words as compared to when interacting with an L2 interlocutor (see Table 3). Previous research has demonstrated that lexical sophistication has a positive relationship with a higher percentage of common multiword units (Kyle and Scott 2015) and semantic diversity (Berger et al. 2017) but a negative relationship with word imageability (Crossely et al. 2011a, 2011b; Kyle and Scott 2015). With this in mind, the current findings may indicate that L2 learners tend to produce more standardized terms in relation to trigram proportion, range, and semantic diversity that indicate a movement toward expected language distributions. However, they produced more sophisticated words in relation to imageability. This provides a picture of L2-L1 interactions as leading to language that follows the expected patterns of language, but also exhibits trends toward more sophisticated language. More specifically, in L2-L1 conversations, L2 speakers may attend to their L1 interlocutors' language use and consciously or unconsciously pick up particular styles (i.e., lexically sophisticated language) that they may align in their later utterances. In contrast, L2-L2 interactions distance themselves from expected language distributions and produce less sophisticated lexical items.

Different language use and alignment behaviors as reported between L2-L1 dyads and L2-L2 dyads may represent a monitoring process in which speakers monitor their addressees' level of understanding during interaction for successful communication (Costa et al. 2008). The difference in the degree of alignment between the L2-L1 and L2-L2 dyads may also be related to different levels of attentional resources required between the interaction with L2 speakers and L1 speakers. For example, when conversing with an L1 interlocutor, L2 speakers may assume that the L1 speaker has complete knowledge of the language and is therefore able to understand all acceptable language use. When addressing an L2 interlocutor, however, L2 speakers may need to monitor their language to a greater degree to make sure that the L2 interlocutor understands the messages correctly due to the interlocutor's incomplete knowledge of the target language. In this regard, it may be plausible that L2 participants modified their language to be lexically less sophisticated in order to become more comprehensible to their L2 interlocutors.

The differing degree of stylistic alignment between L2-L1 and L2-L2 dyads may also be attributed to the effect of social factors (i.e., the interlocutor). According to Giles et al. (1991), alignment in conversation can be interpreted as the interlocutors' desire (whether overt or tacit) for social integration and identification whereas an absence of alignment might reflect divergence in order to maintain social distance, identity, or integrity (Chartrand and Dalton 2008; Meltzoff 2005;

Tomasello and Carpenter 2007). In other words, speakers tend to converge on common language features when they want to increase solidarity with their interlocutors while language divergence occurs when speakers keep their distance from their conversation partners (Branigan et al. 2011). In this sense, greater alignment in L2-L2 dyads may indicate the L2 speakers' desire to identify with the communicative norms of the L2 learner group rather than those of the L1 speaker group.

It is also possible that L1 interlocutors modified their input to accommodate L2 speakers and that L2 speakers showed increasing lexical sophistication that corresponds to this input modification (Crossley et al., 2014, 2016). Crossley et al. (2014, 2016) demonstrated that L1 interlocutors modify their lexical expressions to match those of the L2 learner they are conversing with and the L2 learners show developments over the course of such interactions insomuch as their vocabulary becomes more sophisticated.

Beyond linguistic features, we also investigated if non-linguistic factors mediated the degree of alignment for the four lexical indices across the two groups by examining the relationship between LSM scores for the four indices and a range of non-linguistic factors (i.e., age differences, proficiency differences, gender combinations within the dyad, group membership, familiarity between interlocutors, and linguistic distance between L1 and L2). Overall, the findings support previous L1 studies that have demonstrated that certain properties assigned to speakers and their conversation partners may affect the degree of linguistic alignment. For instance, previous L1 research has found that linguistic alignment has a positive relationship with various social dynamics including similarity and liking between interlocutors (Ireland 2011), suggesting that similarity in status and affinity between two speakers lead to greater linguistic alignment over the course of interaction. In this study, the LME analyses predicting the four LSM indices were each associated with a different set of non-linguistic factors. For example, group membership (L2-L2 or L2-L1 dyad) was a significant predictor of LSM scores for word range, trigram proportion, semantic diversity, and imageability index.

In addition, another social factor, familiarity between interlocutors, was a significant predictor of LSM scores for word range and imageability indices, suggesting that as interlocutors become more familiar with each other, they may align their lexical choice to a greater extent. This finding supports previous studies which have demonstrated that speakers choose the most appropriate expressions for a particular interlocutor they are speaking with based on direct personal experience in order to facilitate successful communication (Clark 1996). For example, linguistic alignment can take place when speakers believe that their partners may understand specific expressions they use because they have been

previously produced during conversation. This belief is subject to change in the course of interaction as speakers may be continually adjusting their belief and perception about their partners throughout a dialogue (Branigan et al. 2010). Thus, it is likely that as the participants in this study interacted to a greater extent with their partners they aligned their lexical choices with their interlocutors to a greater extent, particularly in terms of word range and imageability.

Finally, a non-social factor, proficiency differences between interlocutors, was also a significant factor affecting stylistic alignment. Results showed that the proficiency differences factor was a negative predictor of LSM score for the word range, trigram proportion, and word imageability indices. Specifically, pairs with lower differences in proficiencies were more likely to align their lexical choice with their partners in terms of the range of word, proportion of common trigrams, and degree of imageability of word produced during a dialogue. This finding is consistent with that of previous research (e.g., Costa et al. 2008; Geeslin 2020; Purmohammad 2015), which suggested that speakers' judgments about the interlocutor's proficiency could affect their linguistic behaviors such as lexical or grammatical choice in such a way that they align linguistically with their interlocutors in order to facilitate communication. For example, when an L1 speaker interacts with a more proficient L2 speaker (i.e., an interlocutor with smaller gap in proficiency), automatic linguistic alignment is promoted to a greater extent because the L1 speaker may find that the conscious monitoring of linguistic decisions, which can interfere with alignment, is not necessary (Purmohammad 2015). Likewise, similarities in proficiency between L2 speakers may enhance linguistic alignment because L2 speakers that have similar knowledge of the L2 are more likely to align on a set of representations (lexical, structural, etc.) that they know are shared between them (Costa et al., 2008). Geeslin (2020) also emphasized the role of the interlocutor's characteristics in the language input that an L2 learner receives as well as the output that the learner produces. She proposed that a speaker may modify his/her speech in response to the characteristics of the interlocutor (i.e., proficiency) throughout the interaction.

In terms of methods, most past LSM research (if not all) has analyzed only function words. In contrast, the current study focused on lexical characteristics related to both content and function words in order to examine whether alignment at the level of lexical features occurs exclusively on function words only. Previous LSM research has demonstrated that a speaker's language style can be defined by function words (e.g., pronouns, prepositions, articles) because they tend to be fairly consistent across different utterances (i.e., context independent; Ireland, 2011). Researchers have also argued that although interlocutors tend to match each other's use of both function and content words (e.g., nouns, verbs), stylistic convergence may occur predominantly in function words that are produced and

processed through automatic processes (Garrod and Pickering 2007) rather than in content words which require conscious processing and convey the conceptual meaning of an utterance (Ireland 2011). Interestingly, all of the four lexical features demonstrating differences in stylistic alignment between the L2-L1 and L2-L2 groups were found to be related to all types of words including content and function words. This result extends our understanding of stylistic alignment, indicating that stylistic alignment takes place at the level of lexical characteristics related to not only function words but also content words.

6 Conclusion

This study demonstrated that stylistic alignment behaviors differed between L2-L2 and L2-L1 groups in naturally occurring conversations, with regard to the speakers' lexical choice. In addition, stylistic alignment was affected by non-linguistic factors. The present findings indicate that L2-L2 dyad exhibited stronger alignment across four indices related to semantic diversity, word range, imageability, and the proportion of common trigrams. Importantly, L2-L1 dyads results in more sophisticated lexical production which may lead to greater lexical acquisition while L2-L2 dyads resulted in less sophisticated language that may lead to greater comprehension. Furthermore, factors such as group membership, familiarity between interlocutors, and proficiency differences were found to be significant predictors of stylistic alignment.

While the findings of this study are significant, they are subject to several limitations, which we hope will stimulate future studies. First, the present study endeavored to understand linguistic alignment using the LSM formula, which was appropriate for the nature of the corpus (natural conversation) analyzed in this study. Although the LSM metric has been widely used to estimate linguistic alignment in the literature, some researchers (Doyle et al. 2016) have suggested that, in some cases, the LSM metric quantifies the homophily of a dyad rather than linguistic alignment. This research has introduced an alternative measure (Hierarchical Alignment Model; Doyle et al. 2016) which should be investigated in future research. Second, this study found a difference in the degree of alignment between the L2-L1 and L2-L2 dyads. Although we assumed that the difference may be attributed to differing levels of attentional resources needed in the conversation with L2 speakers and L1 speakers as well as speakers' modification of language, examining this possibility was beyond the scope of the current study. Follow-up studies are necessary to address this issue by investigating how both L1 and L2 interlocutors modify their input while speaking with L2 speakers and how L2

speakers' lexical choice changes in terms of the level of lexical sophistication as a function of their interlocutors' input modifications.

Furthermore, the current findings showed that L2 learners, when talking to L1 speakers, used words that are less contextually variable, more likely to be phrases in reference corpus, and of a greater range, indicating that conversation with L1 speakers may lead to the use of more standardized language. At the same time, the language used is more sophisticated in terms of imageability. Further investigation is required to determine whether conversation with L1 speakers benefits L2 speakers in terms of promoting the use of other distributional norms (i.e., statistical norms or pronunciation norms).

Future research may also consider exploring the interaction between group membership and other factors, such as familiarity between interlocutors, with respect to the occurrence and degree of stylistic alignment. Previous research has suggested that familiarity with the interlocutor's speaking style may influence the extent to which speakers align on their choice of linguistic features (Costa et al. 2008). For instance, compared to two L1 interlocutors, an L1 and L2 speakers in conversation showed reduced phonetic alignment due to the L1 speaker's unfamiliarity with the nonnative-like accent of the L2 speaker (Kim et al. 2011). To our knowledge, the relationship between group membership and familiarity between interlocutors has yet to be examined in the literature on stylistic alignment. Accordingly, future research may benefit from investigating how L2 learners' alignment behaviors differ at the stylistic level between L2-L1 and L2-L2 dyads and the extent to which the degree of stylistic alignment is associated with familiarity between interlocutors. In addition, as the anonymous reviewer pointed out, the current findings could have been conditioned by the configuration of the dyads to a certain extent because it was not possible to balance the number of dyads that fell into each of the seven (L2-L2) and four (L2-L1) possible combinations of proficiency levels. Follow-up research is needed to explore the extent to which L2 proficiency difference may mediate the degree of stylistic alignment in conversations involving L2 learners by controlling the number of dyads in each combination for L2 proficiency difference (e.g., advanced and high-beginning; native speaker and low-intermediate).

Finally, it remains an open question whether the stylistic alignment in natural conversation reported in this study was a product of automatic priming (i.e., automatic linguistic alignment). While researchers have posited that LSM is an automatic process that strongly correlates with human behaviors (Ireland et al. 2011), it is possible that linguistic alignment might have automatically occurred, especially since the L2 participants in this study were all intermediate or advanced learners. In addition, the possibility of conscious, non-automatic alignment may have occurred, particularly in L2-L1 interaction because L2 learners may have aligned with L1 speakers in order to achieve their goal of learning the target language (Costa et al. 2008). In doing so, they

may have repeated the L1 speaker's previous words and phrases or mimicked the L1 interlocutor's language use patterns, which would have increased the amount of non-automatic linguistic alignment driven by the deliberate lexical choice of the L2 learner. This issue will need to be addressed in future research that examines which words and phrases were actually repeated by two interlocutors engaged in dialogue. Additionally, future research is called for to investigate L2 speakers' perception of whether and how their language use and alignment behaviors were affected by their L2 speaker status and the type of interlocutor (L1 or L2 speaker) they interacted with in natural conversations.

References

Adamson, Hugh Douglas & Vera M. Regan. 1991. The acquisition of community speech norms by Asian immigrants learning English as a second language: A preliminary study. *Studies in Second Language Acquisition* 13(1). 1–22.

Argamon, Shlomo, Whitelaw Casey, Paul Chase, Sobhan Raj Hota, Navendu Garg & Shlomo Levitan. 2007. Stylistic text classification using functional lexical features. *Journal of the American Society for Information Science and Technology* 58(6). 802–822.

Baayen, Rolf Harald, Doug J. Davidson & Douglas M. Bates. 2008. Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language* 59(4). 390–412.

Balcetis, Emily E. & Rick Dale. 2005. An exploration of social modulation of syntactic priming. In *Proceedings of the annual meeting of the cognitive science society*, vol. 27. https://escholarship.org/uc/item/0kn388rj (accessed 17 December 2019).

Bates, Douglas M., Martin Maechler, Ben Bolker & Steven Walker. 2015. lme4: Linear mixed-effects models using Eigen and S4_. R package version 1.1-9. https://CRAN.R-project.org/package=lme4 (accessed 5 September 2019).

Bayram, Burcu & Vivian P. Ta. 2018. Diplomatic chameleons: Language style matching and agreement in international diplomatic negotiations. *Negotiation and Conflict Management Research* 12(1). 23–40.

Beebe, Leslie. 1985. Input: Choosing the right stuff. In Susan M. Gass & Carolyn G. Madden (eds.), *Input in second language acquisition*, 104–144. Rowley, MA: Newbury House.

Bell, Allan. 1984. Language style as audience design. *Language in Society* 13(2). 145–204.

Branigan, Holly P., Martin J. Pickering, Jamie Pearson & Janet F. McLean. 2010. Linguistic alignment between people and computers. *Journal of Pragmatics* 42(9). 2355–2368.

Berger, Cynthia M., Scott A. Crossley & Kristopher Kyle. 2017. Using novel word context measures to predict human ratings of lexical proficiency. *Journal of Educational Technology & Society* 20(2). 201–212.

Branigan, Holly P., Martin J. Pickering, Jamie Pearson & Ash Brown. 2011. The role of beliefs in lexical alignment: Evidence from dialogs with humans and computers. *Cognition* 121(1). 41–57.

Brennan, Susan E. & Herbert H. Clark. 1996. Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 22(6). 1482–1493.

British National Corpus. 2007. version 3 (BNC XML ed.). Available at: http://www.natcorp.ox.ac.uk.

Brown, Gordon D. A. 1984. A frequency count of 190, 000 words in the London-Lund Corpus of English Conversation. *Behavior Research Methods, Instrumentation and Computers* 16(6). 502–532.

Brysbaert, Marc & Boris New. 2009. Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods* 41(4). 977–990.

Brysbaert, Marc, Amy B. Warriner & Victor Kuperman. 2014. Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods* 46(3). 904–911.

Cappella, Joseph N. 1996. Dynamic coordination of vocal and kinesic behavior in dyadic interaction: Methods, problems, and interpersonal outcomes. In James H. Watt & C. Arthur VanLear (eds.), *Dynamic patterns in communication processes*, 353–386. Thousand Oaks, CA: Sage Publications.

Chartrand, Tanya L. & Amy N. Dalton. 2008. Mimicry: Its ubiquity, importance, and functionality. In Ezequiel Morsella, John A. Bargh & Peter M. Gollwitzer (eds.), *Oxford handbook of human action*, 458–483. Oxford, UK: Oxford University Press.

Chiswick, Barry R. & Paul W. Miller. 2005. Linguistic distance: A quantitative measure of the distance between English and other languages. *Journal of Multilingual and Multicultural Development* 26(1). 1–11.

Clark, Herbert H. 1996. *Using language*. Cambridge, UK: Cambridge University Press.

Costa, Albert, Martin J. Pickering & Antonella Sorace. 2008. Alignment in second language dialogue. *Language & Cognitive Processes* 23(4). 528–556.

Crossely, Scott A., Tom Salsbury, Danielle S. McNamara & Scott Jarvis. 2011a. What is lexical proficiency? Some answers from computational models of speech data. *TESOL Quarterly* 45(1). 182–193.

Crossley, Scott, Amanda Clevinger & YouJin Kim. 2014. The role of lexical properties and cohesive devices in text integration and their effect on human ratings of speaking proficiency. *Language Assessment Quarterly* 11(3). 250–270.

Crossley, Scott, Kristopher Kyle & Thomas Salsbury. 2016. A usage-based investigation of L2 lexical acquisition: The role of input and output. *The Modern Language Journal* 100(3). 702–715.

Crossley, Scott A., Tom Salsbury, Danielle S. McNamara & Scott Jarvis. 2011b. Predicting lexical proficiency in language learner texts using computational indices. *Language Testing* 28(4). 561–580.

Crossley, Scott A. & Danielle S. McNamara. 2012. Interlanguage talk: A computational analysis of non-native speakers' lexical production and exposure. In Philip M. McCarthy & Chutima Boonthum-Denecke (eds.), *Applied natural language processing: Identification, investigation and resolution*, 425–437. Hershey, PA: IGI Global. https://doi.org/10.4018/978-1-60960-741-8.ch025.

Crossley, Scott A., Stephen Skalicky, Mihai Dascalu, Danielle S. McNamara & Kristopher Kyle. 2017. Predicting text comprehension, processing, and familiarity in adult readers: New approaches to readability formulas. *Discourse Processes* 54(5–6). 340–359.

Doyle, Gabriel, Yurovsky Dan & Michael C. Frank. 2016. A robust framework for estimating linguistic alignment in social media conversations. *Proceedings of the International Conference on World Wide Web* 25. 637–648.

Ellis, Rod. 1997. *Second language acquisition*. Oxford, UK: Oxford University Press.

Fusaroli, Riccardo, Bahador Bahrami, Karsten Olsen, Andreas Roepstorff, Geraint Rees, Chris Frith & Kristian Tylén. 2012. Coming to terms quantifying the benefits of linguistic coordination. *Psychological Science* 23(8). 931–939.

Fusaroli, Riccardo & Kristian Tylén. 2016. Investigating conversational dynamics: Interactive alignment, Interpersonal synergy, and collective task performance. *Cognitive Science* 40(1). 145–171.

Fussell, Susan R. & Robert M. Krauss. 1992. Coordination of knowledge in communication: Effects of speakers' assumptions about others' knowledge. *Journal of Personality and Social Psychology* 62(3). 378–391.

Garrod, Simon & Martin J. Pickering. 2007. Automaticity of language production in monologue and dialogue. In Antje S. Meyer, Linda R. Wheeldon & Andrea Krott (eds.), *Automaticity and control in language processing, advances in behavioral brain science*, 1–21. New York, NY: Psychology Press.

Garrod, Simon & Martin J. Pickering. 2009. Joint action, interactive alignment, and dialog. *Topics in Cognitive Science* 1(2). 292–304.

Gass, Susan. 2003. Input and interaction. In Catherine Doughty & Michael Long (eds.), *Handbook of second language acquisition*, 224–255. Oxford, UK: Blackwell.

Geeslin, Kimberly L. 2020. Variationist perspective(s) on interlocutor individual differences. In Laura Gurzynski-Weiss (ed.), *Cross-theoretical explorations of interlocutors and their individual differences*, 127–157. Amsterdam: John Benjamins.

Giles, Howard, Justin Coupland & Nikolas Coupland. 1991. Accommodation theory: Communication, context, and consequence. In Howard Giles, Justin Coupland & Nikolas Coupland (eds.), *Contexts of accommodation: Developments in applied sociolinguistics*, 1–68. New York, NY: Cambridge University Press.

Gonzales, Amy L., Jeffrey T. Hancock & James W. Pennebaker. 2010. Language style matching as a predictor of social dynamics in small groups. *Communication Research* 37(1). 3–19.

Guo, Liang, Scott A. Crossley & Danielle S. McNamara. 2013. Predicting human judgments of essay quality in both integrated and independent second language writing samples: A comparison study. *Assessing Writing*, 18(3). 218–238.

Hoffman, Paul, Matthew A. Lambon Ralph & Timothy T. Rogers. 2013. Semantic diversity: A measure of semantic ambiguity based on variability in the contextual usage of words. *Behavior Research Methods* 45(3). 718–730.

Hyland, Ken. 2008. As can be seen: Lexical bundles and disciplinary variation. *English for Specific Purposes* 27(1). 4–21.

Ireland, Molly E. 2011. *Three explanations for the link between language style matching and liking*. Austin, TX: The University of Texas at Austin dissertation.

Ireland, Molly E. & James W. Pennebaker. 2010. Language style matching in writing: Synchrony in essays, correspondence, and poetry. *Journal of Personality and Social Psychology* 99(3). 549–571.

Ireland, Molly E., Richard B. Slatcher, Paul W. Eastwick, Lauren E. Scissors, Eli J. Finkel & James W. Pennebaker. 2011. Language style matching predicts relationship initiation and stability. *Psychological Science* 22(1). 39–44.

Jung, YeonJoo, Kim YouJin & Murphy John. 2017. The role of task repetition in learning word-stress patterns through auditory priming tasks. *Studies in Second Language Acquisition* 39(2). 319–346.

Kim, Midam, William S. Horton & Ann R. Bradlow. 2011. Phonetic convergence in spontaneous conversations as a function of interlocutor language distance. *Laboratory Phonology* 2(1). 125–156.

Kim, YouJin, Jung YeonJoo & Skalicky. Stephen. 2019. Linguistic alignment, learner characteristics, and the production of stranded prepositions in relative clauses: Comparing FTF and SCMC contexts. *tudies in Second Language Acquisition* 41(5). 937–969.

Krauss, Robert M. & Jennifer S. Pardo. 2004. Is alignment always the result of automatic priming? *Behavioral and Brain Sciences* 2(2). 203–204.

Kuperman, Victor, Hans Stadthagen-Gonzales & Marc Brysbaert. 2012. Age-of-acquisition ratings for 30 thousand English words. *Behavioral Research Methods* 44(4). 978–990.

Kuznetsova, Alexandra, Per Bruun Brockhoff & Rune Haubo Bojesen Christensen. 2015. lmerTest: Tests in linear mixed effects models. R package version 2.0-29. http://CRAN.R-project.org/package=lmerTest (accessed 17 December 2019).

Kyle, Kristopher & A. Crossley Scott. 2015. Automatically assessing lexical sophistication: Indices, tools, findings, and application. *TESOL Quarterly* 49(4). 757–786.

Kyle, Kristopher, A. Crossley Scott & M. Berger Cynthia. 2018. The tool for the automatic analysis of lexical sophistication (TAALES): version 2.0. *Behavior Research Methods* 50(13). 1030–1046.

Landauer, Thomas, Danielle S. McNamara, Simon Dennis & Walter Kintsch. 2007. *LSA: A road to meaning*. Mahwah, NJ: Lawrence Erlbaum Associates.

Levelt, Willem J. M. 1989. *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

Lev-Ari, Shiri. 2015. Comprehending non-native speakers: Theory and evidence for adjustment in manner of processing. *Frontiers in Psychology* 5. 1546.

McDonald, Scott A. & Richard C. Shillcock. 2001. Rethinking the word frequency effect: The neglected role of distributional information in lexical processing. *Language and Speech* 44(3). 295–322.

McDonough, Kim, Heike Neumann & Pavel Trofimovich. 2015. Eliciting production of L2 target structures through collaborative priming activities. *Canadian Modern Language Review* 71(1). 75–95.

McDonough, Kim & Pavel Trofimovich. 2009. *Using priming methods in second language research*. New York, NY: Routledge.

Meltzoff, Andrew N. 2005. Imitation and other minds: The "like me" hypothesis. *Perspectives on Imitation: From Neuroscience to Social Science* 2. 55–77.

Nakagawa, Shinichi & Holger Schielzeth. 2013. A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution* 4(2). 133–142.

Niederhoffer, Kate G. & James W. Pennebaker. 2002. Linguistic style matching in social interaction. *Journal of Language and Social Psychology* 21(4). 337–360.

Pickering, Martin J. & Simon Garrod. 2004. The interactive-alignment model: Developments and refinements. *Behavioral and Brain Sciences* 27(2). 212–225.

Purmohammad, Mehdi. 2015. Linguistic alignment in L1–L2 dialogue. *Languages and Dialogues* 5(2). 312–333.

R Development Core Team. 2015. *R: A language and environment for statistical computing*. (version 3.1.3) [software]. Vienna: R Foundation for Statistical Computing.

Tarone, Elaine. 1983. On the variability of interlanguage systems. *Applied Linguistics* 4(2). 142–164.

Tarone, Elaine. 1988. *Variation in interlanguage*. London, UK: Edward Arnold.

Tarone, Elaine. 2007. Sociolinguistic approaches to second language acquisition research-1997–2007. *The Modern Language Journal* 91(1). 837–848.

Tomasello, Michael & Malinda Carpenter. 2007. Shared intentionality. *Developmental Science* 10(1). 121–125.
Vinson, David W., Maryam Tabatabaeian Rick Dale & Nicholas D. Duran. 2016. Seeing and believing: Social influences on language processing. In Ramesh Kumar Mishra, Narayanan Srinivasan & Huettig Falk (eds.), *Attention and vision in language processing*, 197–213. Berlin, Germany: Springer.
Weatherholtz, Kodi, Kathryn Campbell-Kibler & T. Florian Jaeger. 2014. Socially-mediated syntactic alignment. *Language Variation and Change* 26(3). 387–420.

Bionotes

YeonJoo Jung
Department of English Education, Pusan National University, Busan, Republic of Korea
yjjung@pnu.edu

YeonJoo Jung is an assistant professor in the Department of English Education at Pusan National University in South Korea. Her interests include the application of experimental techniques from psychology to second language (L2) processing and acquisition research as well as the use of natural language processing tools in assessing linguistic features of written and spoken data produced by L2 learners.

Scott Crossley
Department of Applied Linguistics, Georgia State University, Atlanta, GA, USA
scrossley@gsu.edu

Scott Crossley is a Professor at Georgia State University. His interests include the application and development of natural language processing tools in educational technology. He has published articles on the use of natural language processing tool to examine lexical acquisition, writing proficiency, reading comprehension, discourse processing, language assessment, and automatic feedback in intelligent tutoring systems.