Research Article

Speech Entrainment in Adolescent Conversations: A Developmental Perspective

Camille J. Wynn,[a] (iD) **Tyson S. Barrett,**[b] (iD) **Visar Berisha,**[c,d] (iD) **Julie M. Liss,**[c] (iD) **and Stephanie A. Borrie**[b] (iD)

[a] Department of Communication Sciences and Disorders, University of Houston, TX [b] Department of Communicative Disorders and Deaf Education, Utah State University, Logan [c] Department of Speech and Hearing Science, Arizona State University, Tempe [d] School of Electrical, Computer and Energy Engineering, Arizona State University, Tempe

ABSTRACT

Purpose: Defined as the similarity of speech behaviors between interlocutors, speech entrainment plays an important role in successful adult conversations. According to theoretical models of entrainment and research on motoric, cognitive, and social developmental milestones, the ability to entrain should develop throughout adolescence. However, little is known about the specific developmental trajectory or the role of speech entrainment in conversational outcomes of this age group. The purpose of this study is to characterize speech entrainment patterns in the conversations of neurotypical early adolescents.

Method: This study utilized a corpus of 96 task-based conversations between adolescents between the ages of 9 and 14 years and a comparison corpus of 32 task-based conversations between adults. For each conversational turn, two speech entrainment scores were calculated for 429 acoustic features across rhythmic, articulatory, and phonatory dimensions. Predictive modeling was used to evaluate the degree of entrainment and relationship between entrainment and two metrics of conversational success.

Results: Speech entrainment increased throughout early adolescence but did not reach the level exhibited in conversations between adults. Additionally, speech entrainment was predictive of both conversational quality and conversational efficiency. Furthermore, models that included all acoustic features and both entrainment types performed better than models that only included individual acoustic feature sets or one type of entrainment.

Conclusions: Our findings show that speech entrainment skills are largely developed during early adolescence with continued development possibly occurring across later adolescence. Additionally, results highlight the role of speech entrainment in successful conversation in this population, suggesting the import of continued exploration of this phenomenon in both neurotypical and neurodivergent adolescents. We also provide evidence of the value of using holistic measures that capture the multidimensionality of speech entrainment and provide a validated methodology for investigating entrainment across multiple acoustic features and entrainment types.

Defined as the similarity of speech behaviors between interlocutors, speech entrainment has been documented extensively in neurotypical adult conversation. Adults frequently modify the rhythmic (e.g., speech rate:

Manson et al., 2013; Wynn & Borrie, 2020), articulatory (e.g., articulatory precision: Borrie, Wynn, et al., 2020; Lubold et al., 2019), and phonatory (e.g., pitch properties: Borrie et al., 2015; Lubold & Pon-Barry, 2014) behaviors of their speech to more closely align with the behaviors of their conversation partner. This similarity of behavior can be conceptualized and measured in different ways (see the study of Wynn & Borrie, 2022, for classification framework; see also the study of Rasenberg et al., 2020, for additional measurement considerations). For example,

some studies have focused on proximity (e.g., Fusaroli & Tylen, 2016; Willi et al., 2018), measuring the similarity of speech features between two interlocutors. Other studies have focused on synchrony, measuring the similarity in movement (i.e., direction and magnitude of change) of speech features between two interlocutors, regardless of the actual raw feature values (e.g., De Looze et al., 2014; Schweitzer & Lewandowski, 2013). Beyond evidence of speech entrainment as a robust phenomenon in the conversations of adults, a large body of research has indicated that speech entrainment is generally predictive of functional measures of conversational success (although see the study of Dideriksen et al., 2022, for differing and more nuanced findings regarding linguistic aspects of entrainment). Across different types of conversations (e.g., transactional versus social) with different goals (e.g., accuracy vs. efficiency), high levels of speech entrainment are correlated with greater conversational quality (e.g., Gregory et al., 1997; Wynn et al., 2022) and efficiency (e.g., Borrie & Delfino, 2017; Borrie et al., 2015; Reichel et al., 2018). Furthermore, people who entrain well are rated as being more likeable, competent, friendly, and cooperative than those that do not (e.g., Michalsky & Schoormann, 2017; Polyanskaya et al., 2019; Schweitzer et al., 2017).

The fact that entrainment occurs so frequently in adult conversations does not negate its complexity. Speech entrainment requires the coordination of speech perception and production processes while considering social and contextual factors and simultaneously monitoring the other aspects of conversation. Consequently, there are numerous components that must be in place for speech entrainment to occur (see Figure 1 for an overview). In their model of speech entrainment, Lewandowski and Jilka (2019; see also the study of Lewandowski, 2012) suggest that the speech entrainment process can be broken into a number of different steps, each requiring certain underlying abilities. For instance, the entrainment process begins with the acquisition and encoding of the acoustic details from the speech of an individual's conversation partner. Accordingly, an individual must have adequate attention skills to detect and recognize the acoustic details within their partner's speech patterns and working memory skills

to process and store this information. Indeed, previous research has found both attention (e.g., Yu et al., 2013) and working memory (e.g., Petrone et al., 2021) skills to be tied to higher levels of speech entrainment. Once information has been stored, an individual must then be able to retrieve this information and integrate it into their own speech patterns. This demands the cognitive flexibility and processing speed to make rapid, on-the-spot adaptations as well as sufficient motor control and coordination to integrate the speech patterns of their partner into their own productions. In addition, an individual must have the rhythmic abilities to perceive the speech rhythms of their partner and integrate them into their own speech patterns (Phillips-Silver et al., 2010; Todd et al., 2002; Wynn et al., 2022). Beyond the ability to entrain, Lewandowski and Jilka (2019; Lewandowski, 2012) also note factors that affect an individual's motivation to entrain. This motivation can be driven by both internal (i.e., characteristics of interlocutor) and external (i.e., characteristics of conversational partner and/or environment) factors but often stems from a desire for affiliation with an individual's conversation partner and the need for social approval (e.g., Aguilar et al., 2016; Giles et al., 1991; Natale, 1975).

While speech entrainment has been widely documented in adult conversations, little is known about when these conversational patterns emerge in childhood or adolescence. Theoretically, we would anticipate speech entrainment to emerge after the prerequisite abilities and motivational factors are sufficiently established. From a developmental perspective, early adolescence is a period characterized by rapid and extensive changes across physical, cognitive, and social domains (Feldman & Elliott, 1990; Hill, 1980; Steinberg, 2020). During this time, the speech motor system continues to mature, with acquisition of adultlike control occurring after the age of 14 years (Smith & Zelaznik, 2004; Walsh & Smith, 2002). In theory, with increased speech motor control comes a greater ability to make the fine-grained adjustments necessary to more closely align with the speech of one's interlocutor. Early adolescence is also a time of rapid cognitive development. Consequently, as skills such as attention (e.g., Karns et al., 2015; Memmert, 2014; Mizuno et al., 2011),

Figure 1. Visual representation of some of the components necessary for entrainment to occur.
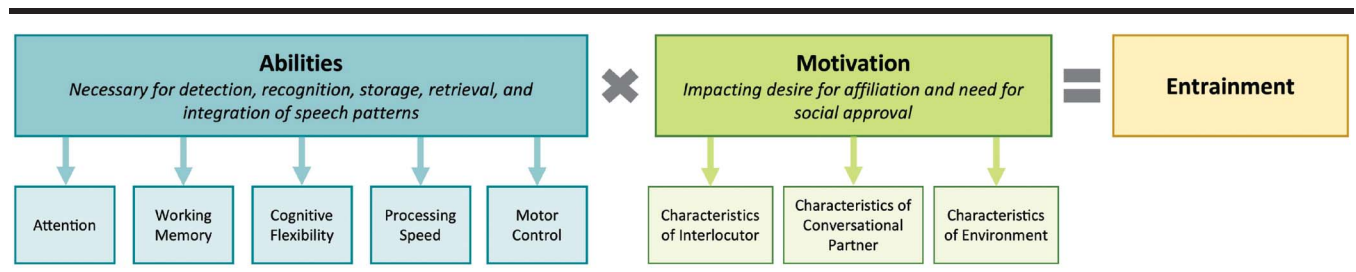
working memory (e.g., Carriedo et al., 2016; Ferguson et al., 2021; Mizuno et al., 2011), cognitive flexibility (Luna et al., 2004; Rubia et al., 2006; Williams et al., 1999), and processing speed (Demetriou et al., 2002; Kail & Ferrer, 2007; Luna et al., 2004) become more adultlike, so too should the capacity to entrain. Beyond the development of underlying abilities necessary for entrainment, changes in the social domain may affect an early adolescent's motivation to entrain as well. Research has shown that adolescence is a time of heightened social sensitivity (Dreyfuss et al., 2014; Somerville, 2013; Somerville et al., 2011), and adolescents show higher levels of self-consciousness (Somerville et al., 2013) and greater emotional responses to social rejection (Platt et al., 2013; Sebastian et al., 2010; Stroud et al., 2009) than children or adults. During early adolescence, the nature of social interactions also changes with an increased emphasis on peer interactions. To start, the sheer amount of time spent with peers dramatically increases during this time period (Lam et al., 2014; Larson & Richards, 1991). With this increase comes changes in the importance of these relationships as well. As individuals transition from childhood to adolescence, they are increasingly more likely to identify friends as part of the network of people most close and important to them (Levitt et al., 1993). Furthermore, early adolescents report a more positive affect when they are with their friends than with family members or alone (Larson & Richards, 1991; Raffaelli & Duckett, 1989). Accordingly, an amplified need for social approval and a desire for affiliation may lead to increased motivation to entrain, particularly in peer interactions. This may be especially true for adolescent girls who spend more time with peers (Lam et al., 2014) and rely more on friends for intimacy and support (Buhrmester & Furman, 1987) than adolescent boys. Taken together, this body of literature leads us to hypothesize that speech entrainment is largely developed during early adolescents, increasing with age, and that, across early adolescence, robust patterns of entrainment will increase more rapidly in the peer conversations of girls than boys.

Should entrainment indeed emerge during early adolescence, another important question regards the role of entrainment in early adolescent conversations. Early adolescence is a time when friendships are incredibly important for emotional health and well-being. For example, early adolescents who report closer friendships also report a more positive self-concept, higher self-esteem, less loneliness, and lower levels of depression (Levitt et al., 1993; Lodder et al., 2017; Pachucki et al., 2015). Research has shown that positive peer interactions constitute some of the most important contributions to adolescent quality of life (Helseth & Misvaer, 2010). Not only is friendship more important, but conversation begins to take a more prominent role in these interactions as friendships begin to rely less on "play" and more on conversation, intimacy, and self-disclosure (La Gaipa, 1979; Larson, 2001; Raffaelli & Duckett, 1989). Raffaelli and Duckett (1989) found that the amount of time that girls are engaged in conversation doubles between the fifth and ninth grade (i.e., ages 10–15 years), with ninth-grade girls spending 16 hr a week "just talking" (i.e., talking in the absence of any other activities). Although less dramatic, there is also an increase in the amount of time boys spend in conversation, with ninth-grade boys reporting conversation as their primary activity for an average of 8 hr a week. Given the importance of peer interactions and the role of conversation within these interactions, understanding the role of speech entrainment in fostering successful conversations is particularly crucial in this age group. The robust body of literature showing a relationship between speech entrainment and conversational success in adult conversations (e.g., Borrie & Delfino, 2017; Polyanskaya et al., 2019; Wynn et al., 2022) leads us to hypothesize that a similar relationship will occur in adolescent conversations. However, more research is needed to validate this hypothesis.

While there are a few studies examining speech entrainment in childhood and/or adolescence generally, and early adolescents specifically, key methodological decisions limit the conclusions that can be drawn. For instance, of the few studies in this area, the majority have examined entrainment in highly structured environments, using shadowing or quasiconversational paradigms with prerecorded stimuli (e.g., Oviatt et al., 2004; Wynn et al., 2018). Although these studies have provided important foundational information in a tightly controlled setting, entrainment in these types of settings may not be indicative of entrainment in naturalistic and embodied conversations (Lewandowski & Jilka, 2019; Pardo et al., 2018). Additionally, studies that have investigated entrainment within more natural contexts have evaluated conversational dyads consisting of one adolescent and one adult (Lehnert-LeHouillier et al., 2020). No study, to our knowledge, has explored the speech entrainment patterns between adolescent–adolescent dyads engaged in conversation with each other. Given the importance of peer interaction in this age group and the documented differences in adolescent interactions with adults versus peers (Larson, 1983; Raffaelli & Duckett, 1989), research exploring speech entrainment in naturalistic conversations between adolescent peers is needed. Beyond the type of interaction, the vast majority of studies have only focused on entrainment across one or a couple of speech features (e.g., speech rate: Wynn et al., 2019; voice onset time: Schertz & Johnson, 2022) and have only explored one type of entrainment (generally proximity; e.g., Lehnert-LeHouillier et al., 2020). Consequently, gaining a more holistic representation

of the speech entrainment patterns in this age group requires investigation of entrainment across multiple speech features and different types of entrainment. Finally, while these studies have explored the presence of speech entrainment within adolescent conversations, we know of no study that has examined the developmental trajectory of entrainment nor the consequences of entrainment.

Purpose

The purpose of this study is to comprehensively examine speech entrainment patterns in the conversations of neurotypical early adolescents. To do this, we use a corpus of 96 task-based dyadic conversations between early adolescent peers (Hazan et al., 2016). In our analysis, we examine entrainment of many different acoustic features across rhythmic, articulatory, and phonatory dimensions of speech rather than focusing on one or a couple of features. This multidimensional approach was selected because there is currently no strong theoretical or empirical rationale for selecting entrainment of one speech feature over another (Pardo, 2013) and research has shown that entrainment of one feature is not necessarily indicative of entrainment of other features (Ostrand & Chodroff, 2021). Accordingly, as the overall aim of this article was to gain a broad and holistic understanding of entrainment in this age group, we did not want to make assumptions about speech entrainment generally based on a couple of self-selected features. In a similar vein, we also examine two different types of entrainment (i.e., proximity and synchrony). The overall purpose of this study can be divided into two primary objectives. In the first objective, we investigate the developmental trajectory of acoustic–prosodic entrainment patterns of early adolescents. Specifically, we ask the following question: (a) How do speech entrainment patterns differ across age group and gender in the peer conversations of early adolescents? We also seek to delineate a more detailed understanding of the entrainment patterns of this age group by asking the following question: (b) How do speech entrainment patterns differ across acoustic feature sets and entrainment types in the peer conversations of early adolescents? In the second objective, we explore the relationship between speech entrainment and conversational success in this same age group. Specifically, we ask the following question: (c) To what degree does entrainment predict metrics of conversational success (i.e., conversational quality and conversational efficiency) in the peer conversations of early adolescents? Again, in order to gain more detailed understanding of the relationship between speech entrainment and conversational success, we also ask the following question: (d) How does the relationship between speech entrainment and conversational success differ across speech dimension and entrainment type? In the process, we weave together techniques and methods from several previous studies (e.g., Borrie et al., 2019; Ostrand & Chodroff, 2021; Reichel et al., 2018) to create and validate a new methodology for capturing speech entrainment.

Method

Conversational Corpus

This study relied on an existing corpus collected and made available to the research community by Hazan et al. (2016; see also Bradlow, n.d.). Participants consisted of 96 neurotypical individuals (46 boys, 50 girls) between the ages of 9 and 14 years inclusive[1] (M = 12 years; 2 months; SD = 21 months). Participants were divided into three broad age groups: 9–10 years, 11–12 years, and 13–14 years. All participants were native speakers of Southern British English with no reported history of hearing or language impairments. Additionally, all participants passed a hearing screening at 25 dB at octave frequencies between 250 and 8000 Hz in both the left and right ears. Participants completed the conversational task in dyads with another individual with whom they were friends. All dyads consisted of participants who were the same gender and fell within the same age group. In total, 48 dyads completed two conversations each,[2] yielding a total of 96 conversations for analysis in this study.

Procedure

An overview of the methodological process for this study is illustrated in Figure 2. All conversations for this corpus were elicited using the Diapix task (Baker & Hazan, 2011; Van Engen et al., 2010), a task-based dialog elicitation procedure commonly used in speech entrainment research (e.g., Borrie et al., 2015, 2019). The Diapix task is a collaborative "spot-the-difference" task in which dyads must verbally work together to identify differences between sets of pictures. In this task, each partner is given one of a pair of pictures. Pictures are virtually identical but have 12 differences in details between the two (e.g., green vs. blue garbage can).

To familiarize participants with the task, dyads completed a practice trial prior to the first recorded conversation. If participants struggled to understand the task

---

[1]One child was 15;1 (years;months).
[2]Although the original corpus contains additional conversations taken in different conditions, we only use the conversations from the normal transmission condition in this study. All conversations for this condition were collected prior to conversations for other conditions.

Figure 2. Overview of methodological process for this study. Spoken dialogs are divided into individual speaking turns. Moreover, 429 acoustic features (divided into five acoustic feature sets) are extracted from each speaking turn in every conversation. Proximity and synchrony scores are calculated for each acoustic feature, yielding 858 entrainment scores per speaking turn. Predictive modeling is used to evaluate the degree of entrainment (i.e., degree to which entrainment scores could be used to distinguish real and sham conversational turns) and the relationship between entrainment and conversational success (i.e., degree to which entrainment scores could be used to predict conversational efficiency and quality scores). EMS = envelope modulation spectrum; LTAS = long-term average spectrum; MFCC = mel-frequency cepstrum coefficient; VR = voice report.
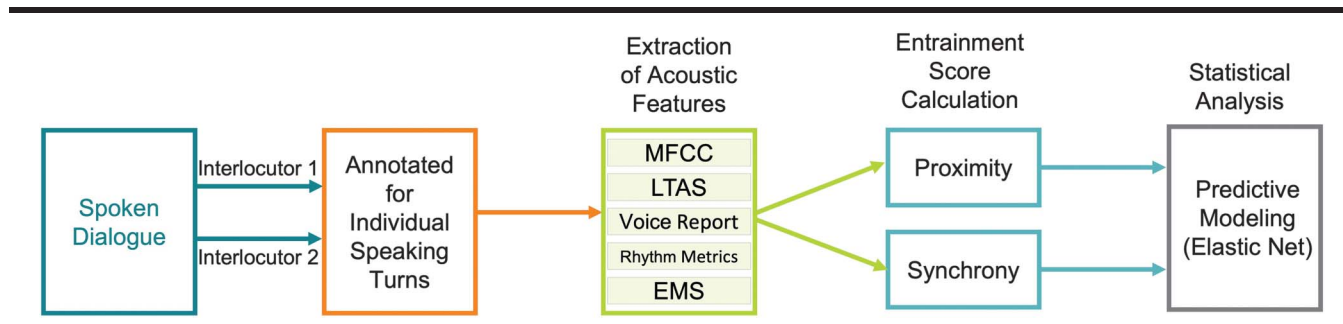
during practice, the experimenter gave clues, and after the dyad had found several differences, they were allowed to look at each other's pictures and continue comparing them. After the practice, participants sat in different rooms from each other and communicated via head-mounted condenser cardioid microphones (Beyerdynamic DT297). Dialogue was audio-recorded in separate channels at a sampling rate of 44100 Hz (16 bits) using an EMU 0404 USB audio interface and Adobe Audition software. Participants were told they would be given 10 min to find as many differences as they could from one set of pictures. For each conversation, one child was designated as the "leader" (i.e., the person responsible for leading the conversation) and one child was designated as the "follower" (i.e., the person responsible to ask questions and make suggestions during the interaction). As dyads participated in two conversations together, each participant played the role of leader and follower in one conversation. For each conversation, the recording was stopped when participants found all 12 differences or after approximately 10 min had elapsed.

Comparison Corpus

Thirty-two adult conversations were also analyzed in order to compare adolescent entrainment patterns to the patterns of adults. Conversations came from 16 randomly selected dyads of 32 adults (16 men, 16 women) from an additional corpus collected by Baker and Hazan (2011; see also Bradlow, n.d.). All participants were native speakers of Southern British English with no reported history of hearing or language impairments. Additionally, all participants passed a hearing screening at 25 dB at octave frequencies between 250 and 8000 Hz in both the left and right ears. As with the adolescent corpus, adult dyads consisted of individuals of the same gender who were friends

prior to the experiment. Participants completed the same conversational task and were recorded in the same conditions as the adolescent participants.

Acoustic Analysis

We used a procedure validated by Borrie et al. (2019; Borrie, Barrett, et al., 2020) to extract acoustic information from the audio-recorded conversations. Trained research assistants manually coded each audio file, annotating individual conversational turns (defined as units of speech by a single interlocutor that were free from pauses greater than 50 ms) by conversational partner using the Praat textgrid function (Boersma & Weenink, 2020). We then utilized a previously used (Borrie et al., 2019; Borrie, Barrett, et al., 2020) software package implemented in MATLAB (Version 9.9.0), to extract 429 acoustic features that can be broadly divided into five acoustic feature sets for each interlocutor in each conversational turn. While nuanced and complex, these acoustic feature sets can be roughly divided into three dimensions of the speech signal: rhythmic, articulatory, and phonatory.[3] Each acoustic feature set is described briefly below. For comprehensive

---

[3]While our division of acoustic measures into three broad dimensions provides simplicity and interpretability, it is important to note that these dimensions are complex and largely overlapping. For example, rhythmic features technically represent temporal aspects of articulation. Phonation also represents a specific type of passive articulation at the level of the larynx. Therefore, our measures of articulation can best be thought of as a reflection of active articulation at the level of the supralaryngeal articulators. Furthermore, given the complexity of the signal, specific acoustic measures will often pick up information representing multiple dimensions of speech. Accordingly, here, we categorize acoustic measures by the dimension with which they are most closely associated.

details of feature calculation, please refer to Supplemental Material S1.

Mel-Frequency Cepstrum Coefficients

Mel-frequency cepstrum coefficients (MFCCs) are coefficients that capture the short-term power spectrum of a speech segment and most closely represent the articulatory dimension of the speech signal (Davis & Mermelstein, 1990). In this analysis, the speech signal is filtered into 39 frequency bands distributed approximately evenly along the Mel scale. For each of the 39 signals, the data are framed using a 20-ms window with 10-ms frame increment from which log energy is calculated. Log energy values are then decorrelated by using an inverse discrete cosine transform. Within each of the 39 MFCCs, six different statistics are computed, resulting in a 234-dimensional feature vector.

Long-Term Average Spectrum

Long-term average spectrum (LTAS) is an analysis of the average energy distribution across frequency over an utterance and most closely represents the articulatory dimension of the speech signal. In this analysis, the speech signal is passed through an octave filter breaking it into nine bands with center frequencies at 30, 60, 120, 240, 480, 960, 1920, 3840, and 7680 Hz. Data for each of the 10 band signals (i.e., the nine octave bands and the full signal) are framed using a 20-ms rectangular window with no overlap. Ten features are then extracted from each of the 10 signals, resulting in a 99-dimensional[4] speech vector.

Voice Report

The voice report is based on a set of features such as fundamental frequency, jitter, shimmer, and harmonics-to-noise ratio and most closely represents the phonatory dimension of speech. Using a custom Praat script, several features are extracted using a 5-ms time step. Default parameters for pitch floor, pitch ceiling, silence threshold, and voicing threshold were used for adults. Default parameters for pitch floor, silence threshold, and voicing threshold were also used for adolescents. However, in order to minimize tracking errors that often occur with higher pitches, parameters for pitch ceiling were adjusted from the default (600 Hz) to 510 Hz.[5] Additional phonatory features and measures of central tendency and variation are also included in the feature set, resulting in a 24-dimensional feature vector.

Rhythm Metrics

Rhythm metrics are analyses of voice timing based on voiced (i.e., vocalic) and voiceless (i.e., intervocalic) interval durations and most closely represent the rhythmic dimension of speech. In this analysis, a Praat script is used to partition the speech signal into vocalic and intervocalic intervals on a frame-by-frame basis using the periodicity detection algorithm outlined by Boersma (1993). The duration of the vocalic and intervocalic segments and a series of related metrics are then extracted, resulting in a 12-dimensional speech feature vector.

Envelope Modulation Spectrum

Envelope modulation spectrum (EMS) is a representation of the slow amplitude modulations in the speech signal and most closely represents rhythmic dimensions of speech (Liss et al., 2010). In this analysis, speech recordings are filtered into nine bands with center frequencies at 30, 60, 120, 240, 480, 960, 1920, 3840, and 7680 Hz. Amplitude envelopes are taken for each of the 10 band signals (i.e., the nine octave bands and the full signal). The mean is removed, and the power spectrum for each of the bands is calculated. Six EMS metrics are computed for each of the 10 power spectra, resulting in a 60-dimensional feature vector.

Entrainment Measures

Two entrainment scores (i.e., proximity and synchrony) were calculated for each of the 429 feature values generated for each conversational turn, resulting in 858 entrainment scores per turn. These measures were selected based on substantial documentation of their occurrence in adult conversations (e.g., Borrie et al., 2015; Reichel et al., 2018).

Proximity

We calculated a proximity score (specifically static local proximity; see the study of Wynn & Borrie, 2022), which accounts for the degree of similarity of speech features between two interlocutors across adjacent turns. If two interlocutors have similar features values, their conversation would be characterized by high levels of proximity, whereas conversations with dissimilar feature values would be characterized by low levels of proximity. Here, we compute a proximity score for each adjacent turn in each conversation by calculating the absolute value of the difference between the feature value of the interlocutor and the feature value of their partner's preceding turn. Thus, the formula for proximity scores can be expressed as follows:

$$\text{Proximity} = |\ Y_t - Y_{t-1}|, \qquad (1)$$

where Y is an acoustic feature value, t is the current conversational turn, and $t-1$ is the conversational partner's

---

[4]Root-mean-square energy is not applicable to analysis of the 10th band signal (i.e., the full signal) and was thus not extracted. Thus, the resulting vector is 99, rather than 100 dimensions.

[5]This value was determined because it was 2 *SD*s above the mean fundamental frequency for the participant with the highest mean fundamental frequency. We opted to keep the same parameters for all participants (regardless of age and gender) as frequency changes due to puberty occur at different times for each adolescent.

previous turn. An example of proximity can be viewed in Figure 3.

Synchrony

We also calculated a synchrony score (specifically positive static local synchrony), which accounts for the degree of similarity in movement (i.e., direction and magnitude of change) of speech features across adjacent turns. For example, two interlocutors may have very different feature values, but on a turn-by-turn basis as they converse, they adjust their speech in the same direction and to the same degree as their partner. Here, we compute a difference score for each turn by subtracting an interlocutor's feature value from their mean value across the entire conversation. A synchrony score is then generated by calculating the absolute difference between difference scores of adjacent turns. Thus, the formula for synchrony scores can be expressed as follows:

$$\text{Synchrony} = |(Y_t - \bar{Y}_{\text{speaker of } t}) - (Y_{t-1} - \bar{Y}_{\text{speaker of } t-1})|, \quad (2)$$

where Y is an acoustic feature value, t is the current conversational turn, and $t-1$ is the conversational partner's previous turn. The $\bar{Y}_{\text{speaker of } t}$ represents the average of the feature across all turns by the individual whose turn was time t. Synchrony-related distance is thus low if both interlocutors realize a feature to approximately the same degree either above or below their respective means. An example of synchrony can be viewed in Figure 3.

Sham Corpus Construction

Using procedures commonly used in entrainment research (e.g., Fusaroli & Tylen, 2016; Ostrand & Chodroff, 2021; Wynn et al., 2022), we generated a comparative sham corpus in order to assess the degree to which the speech of interlocutors is aligned above chance. To do this, sham conversations were created by pairing the feature values of each interlocutor in each conversation with the feature values of a conversation partner with whom they did not actually converse. In order to account for natural differences in acoustic feature values across age and genders, sham partners were assigned from conversations within the same age group and gender. Additionally, the order of conversational turns was maintained across sham conversations so that the feature values of the first interlocutor's first turn were paired with the values from the second interlocutor's first turn as well. Thus, sham conversations have all of the interdependent behavior of entrainment removed from the conversation, leading to measures that represent a null (or no relationship) distribution, while maintaining the other aspects of conversation (e.g., acoustic speech signals of two individuals across a conversation). Measures of synchrony and proximity for each acoustic feature for each conversational turn in each sham conversation were calculated the same way as was done for real conversations.

Conversational Success

Scores for two different metrics of conversational success (which are commonly used in conversational research; see the studies of Baker & Hazan, 2011; Gregory et al., 1997; Van Engen et al., 2010; Wynn & Borrie, 2022) were calculated for each conversation. The first score focused on conversational efficiency, and the second score focused on conversational quality. Details for each score are described below.

Conversational Efficiency

The Diapix task completed in each conversation grants us an objective measure of conversational efficiency. Recall that the Diapix task required dyads to work

Figure 3. Schematic illustrating two types of entrainment evaluated within this study. Proximity represents similarity in the speech features between two interlocutors. Synchrony represents similarity in the movement of speech features between two interlocutors.
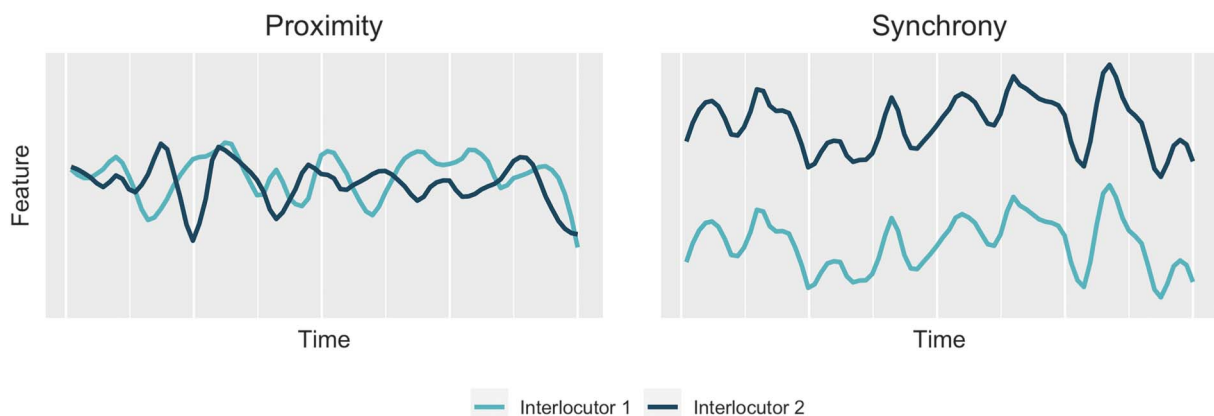
together to find differences in sets of pictures. Thus, the efficiency score is essentially an evaluation of joint task performance, accounting for how effectively the dyad used speech communication to collaboratively work through the demands of the task. In this study, while not every dyad found all 12 differences, every dyad found at least eight. Accordingly, conversational efficiency scores were obtained by determining the length of time (in seconds) required by dyads to find eight of the 12 differences.

Conversational Quality

In order to evaluate the quality of each conversation, we collected ratings from certified school-based speech-language pathologists (SLPs). SLPs were selected as raters because of their expertise in making judgments about conversational quality and high interrater reliability on similar sorts of conversational evaluation in previous studies (Borrie, Barrett, et al., 2020; Borrie et al., 2019). Because interlocutors often take the first minute of the conversation to adjust to their partner and the conversational task, SLPs were asked to listen to audio recordings of the second minute of the conversation, as is frequently done in studies employing interactional ratings (e.g., Balaam et al., 2011; Bernieri et al., 1994; Ingham et al., 2001). Using the 7-point Likert-type rating scale with options from *strongly agree* to *neutral* to *strongly disagree*, SLPs assessed conversational recordings according to the extent that they agree with the following statement: "This conversation seems to flow well (i.e., the conversation feels natural and both participants seem actively engaged)." In order to ensure better attention and engagement during the task, raters were told they were free to take breaks at any time. Additionally, two attentional "catch trials" were placed after one third and two thirds of the audio recordings. In these trials, the audio recordings contained a man's voice asking participants to select a specific answer. Accuracy on catch trials across all raters was 100%. The order of presentation of audio clips was randomized across SLP raters. However, each of the SLPs rated each conversation. Conversational quality scores were calculated assigning a numeric value to each rating (7 = *strongly disagree,* 4 = *neutral,* 1 = *strongly agree*) and subsequently averaging the five SLPs' ratings for each conversation. Thus, for both conversational efficiency and conversational quality, lower scores were representative of more successful conversations.

Statistical Analysis

Our first objective focused on characterizing the developmental trajectory of speech entrainment patterns of early adolescents. To do this, statistical analysis relied on predictive modeling, similar to what has been done in previous studies (Ostrand & Chodroff, 2021; Willi et al.,

2018). Here, predictive models were used to determine the accuracy with which entrainment scores (both proximity and synchrony) for each acoustic feature (the predictor variables) could be used to classify a conversational turn as belonging to a real or sham conversation (the outcome variable). Accuracy values above 50% would indicate some level of entrainment (i.e., models are able to differentiate entrainment in real vs. sham conversational turns above the level of chance), with higher predictive accuracy values representing higher levels of entrainment. For these predictive models, we relied on elastic net (or Lasso) using a logit link with a binomial distribution. Elastic net is a popular predictive approach built on (generalized) linear models that handles high multicollinearity naturally and is commonly used in human interaction literature (e.g., Borrie, Barrett, et al., 2020; Borrie et al., 2019). Model-specific parameters were selected and evaluated based on 10-fold cross-validation. Additionally, each cross-validation performance was run 10 times, and mean results from all 10 folds of all 10 repetitions are reported. This was done to reduce the noise in estimates of model performance, providing more reliable predictive accuracy values. Each model was assessed in R (Version 4.1.2) using the "caret" package (Version 6.0.90, Kuhn, 2017). Predictive models were run separately for different age and gender groups, and we examined the predictive accuracy patterns across groups. Additionally, separate models for each of the five acoustic feature sets (i.e., MFCC, LTAS, voice report, rhythm metrics, and EMS) and the two types of entrainment (i.e., synchrony and proximity) were analyzed, and differences in predictive accuracy were examined across feature set and entrainment type.

Our second objective focused on understanding the relationship between speech entrainment and conversational success. For this aim, because all age and gender groups were considered together within the same models, all raw feature values were scaled by age and gender groups prior to calculating entrainment scores, meaning that the range of possible entrainment scores was standardized along the same scale for each age and gender group. This was done to account for differences in entrainment scores that are a natural result of differences in acoustic feature variation across age and gender. Additionally, because there was only one conversational efficiency score and one conversational quality score per conversation, entrainment scores for each conversational turn were averaged together to create one synchrony score and one proximity score for each acoustic feature per participant for each conversation. After this was accomplished, statistical analysis again relied on elastic net regression using an identity link and distribution with 10-fold cross-validation with 10 repetitions. For this objective, because conversational efficiency and quality scores represent

continuous variables, predictive accuracy was evaluated using root-mean-square error (RMSE) and R^2 values (from the cross-validated models). Here, models were used to determine the degree to which entrainment scores (the predictor variables) were predictive of conversational quality and conversational efficiency (the outcome variables) across all adolescent participants. Additionally, separate models for each of the five acoustic feature sets (i.e., MFCC, LTAS, voice report, rhythm metrics, and EMS) and the two types of entrainment (i.e., synchrony and proximity) were analyzed, and differences in RMSE and R^2 values were examined across feature set and entrainment type. Analysis code associated with this work is provided in Supplemental Material S2.

Results

Summary Statistics for Conversations

Table 1 shows the descriptions of conversations for every age group of adolescents and for the adult group. Linear mixed models (with a random intercept for dyad) indicated no significant difference between number of turns per conversation, total conversation length, or average turn length for any age group of adolescents. Additionally, there was no significant difference between adults and any age group of adolescents for most of these factors. The only exception was that the 9- to 10-year group had significantly longer conversations with longer turn durations than adults.

Characterization of Entrainment Patterns in Early Adolescence

Entrainment Across Age Group

Our first analysis focused on the percentage of accuracy with which entrainment scores could be used to distinguish between real and sham conversations in predictive models of each age group. Findings are illustrated in Figure 4 (see also Table 1 in Supplemental Material S3). In this analysis, predictive accuracy was higher than chance (50%) for all groups, meaning some degree of entrainment was present across all age groups. The degree of predictive

accuracy (using cross-validation with untrained data) in the 9- to 10-year group was 59% (standard error [SE] of the average prediction accuracy = 2%), the 11- to 12-year group was 61% (SE = 2%), the 13- to 14-year group was 64% (SE = 2%), and the adult group was 66% (SE = 2%). Here, accuracy can be seen as a proxy measure for the degree of entrainment occurring in each age group. Thus, collective incremental changes between adjacent age groups resulted in a substantive increase in entrainment from the 9- to 10-year group to the adult group.

Entrainment Across Gender

Our next analysis focused on the accuracy of predictive models of each gender for each adolescent age group. In the 9- to 10-year group, predictive accuracy was 61% (SE = 3%) for girls and 59% (SE = 3%) for boys. In the 11- to 12-year group, accuracy was 64% (SE = 3%) for girls and 63% (SE = 4%) for boys, and in the 13- to 14-year group, accuracy was 68% (SE = 2%) for girls and 65% (SE = 2%) for boys. Further details regarding this analysis can be found in Table 2 in Supplemental Material S3.
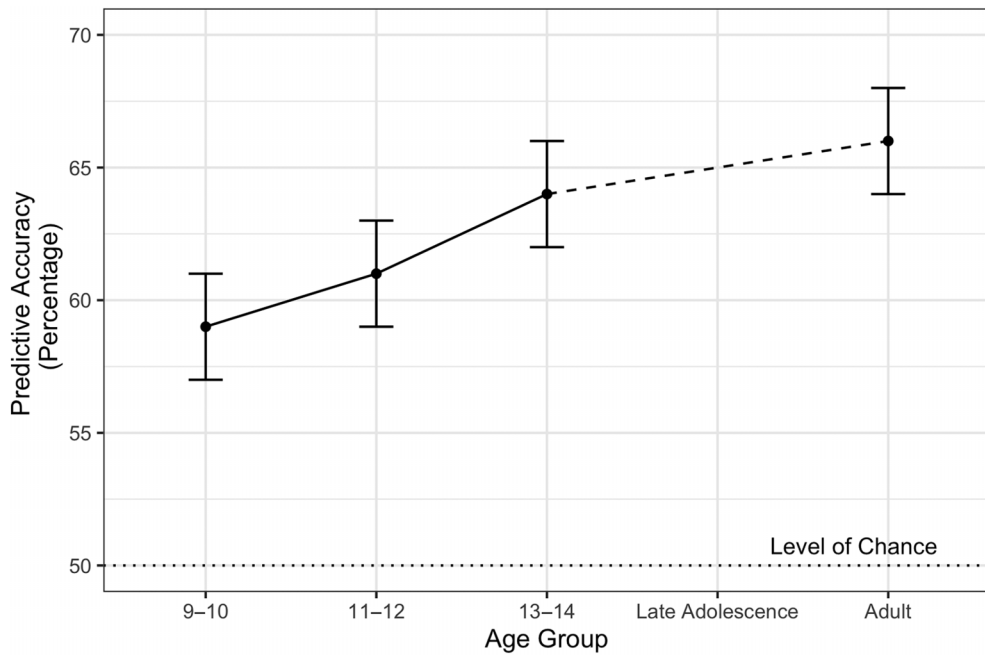
Entrainment Across Acoustic Feature Sets

Next, we examined the accuracy of predictive models for each of the five acoustic feature sets described above (i.e., MFCCs, LTAS, voice report, rhythm metrics, EMS) stratified by age and gender groups. Findings are illustrated in Figure 5 (see also Table 3 in Supplemental Material S3). To summarize, across age and gender groups, predictive accuracy scores were generally highest for LTAS and MFCC feature sets (i.e., feature sets that most closely represent articulation). This was followed by predictive accuracy of the voice report (i.e., feature set that most closely represents phonation). Finally, this was followed by predictive accuracy of the rhythm metrics and EMS feature sets (i.e., feature sets that most closely represent rhythm). Here, we note that predictive accuracy of different acoustic features sets follows a similar trajectory with older adolescents generally showing a higher degree of predictive accuracy than younger adolescents and girls showing a higher degree of predictive accuracy more often than boys. We also note that in nearly every age and gender group, the predictive accuracy of the full set of acoustic features (across all feature sets) was higher than or

Table 1. Descriptive statistics of conversations by age group.

Variable	Age group			
	9–10, *M (SD)*	11–12, *M (SD)*	13–14, *M (SD)*	Adult, *M (SD)*
Number of turns	127.1 (25.9)	133.8 (39.2)	146.5 (51.1)	128.0 (33.5)
Duration of conversation (in seconds)	529.4 (86.8)	501.3 (125.5)	508.4 (105.7)	447.2 (136.3)
Turn duration (in seconds)	2.6 (0.3)	2.4 (0.4)	2.4 (0.4)	2.3 (0.3)

Figure 4. Predictive accuracy of entrainment models by age group. Here, error bars represent standard error. The solid line represents the trajectory of entrainment development as represented by data analyzed within the study. Although no data were collected for a late adolescence group, the dotted line represents a possible continued trajectory for entrainment development across this time period.

Figure 5. Predictive accuracy of entrainment models by acoustic feature set. Full represents models containing all acoustic feature sets EMS = envelope modulation spectrum; LTAS = long-term average spectrum; MFCC = mel-frequency cepstrum coefficient; VR = voice report.
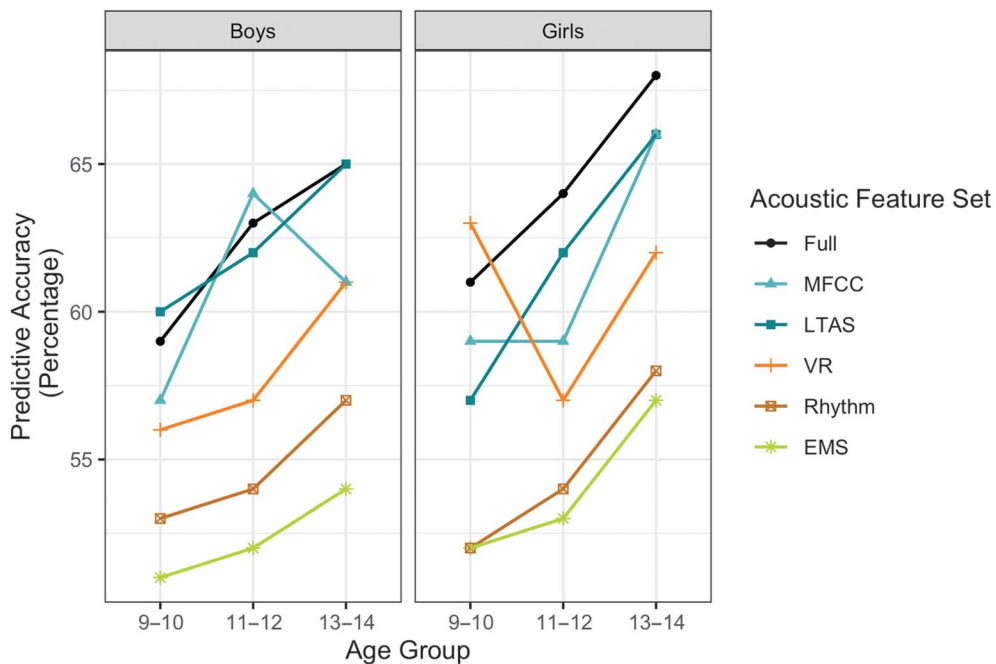
equal to the accuracy of any single dimension. The exceptions to this are in the 9- to 10-year group of boys where the predictive accuracy for LTAS was 60% and the predictive accuracy for the full set was 59%, the 11- to 12-year group of boys where predictive accuracy for the MFCC feature set was 64% and predictive accuracy for the full set was 63%, and the 9- to 10-year group of girls where the predictive accuracy for the voice report feature set was 63% and the predictive accuracy for the full set was 61%.

Entrainment Across Entrainment Type

Our last analysis for this objective focused on the accuracy of predictive models for each of the two entrainment types described above (i.e., proximity and synchrony) divided by age and gender groups. Findings are illustrated in Figure 6 (see also Table 4 in Supplemental Material S3). In sum, across all age and gender groups, predictive accuracy for proximity scores was higher than synchrony scores. As with acoustic feature sets, predictive accuracy of different entrainment types follows a similar trajectory with older children showing a higher degree of predictive accuracy than younger children and girls generally showing higher degree of predictive accuracy more often than boys. Finally, we note that across every age and gender group, full models with both proximity and synchrony together show higher predictive accuracy than separate models with either proximity or synchrony on their own. Additional analyses examining entrainment type across acoustic feature sets is included in Table 5 and Figure 10 in Supplemental Material S3.

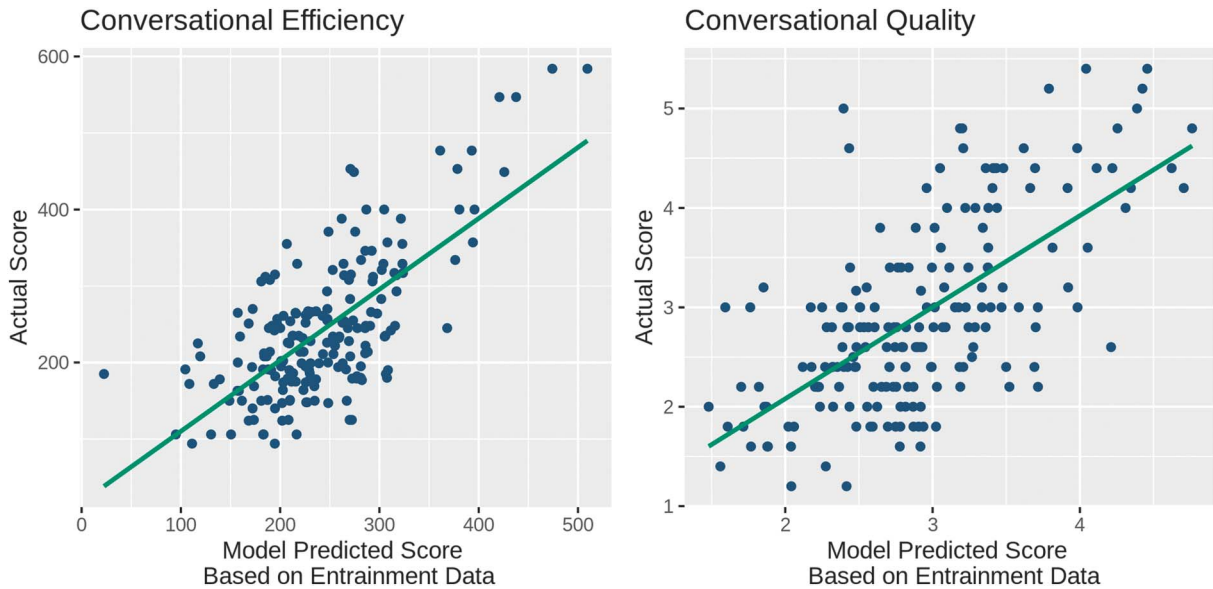### Entrainment and Conversational Success

Predictive Models of Conversational Success

In these analyses, we focused on two measures of conversational success—conversational efficiency and conversational quality. Linear mixed models (with a random intercept for dyad) showed a small (but significant) relationship between conversational efficiency and conversational quality (β [standardized coefficient] = .22, p = .04), indicating that these metrics were largely measuring different constructs. Our first analysis for this objective focused on the predictive relationship between entrainment scores (both synchrony and proximity across all acoustic features) and conversational efficiency. Findings are illustrated in Figure 7 (see also Table 6 in Supplemental Material S3). First, looking at conversational efficiency, our R^2 value was .44 (SE = .19), and the RMSE value was 69 s (SE = 10). Second, looking at conversational quality, our R^2 value was .40 (SE = .14), and the RMSE was .74 units (SE = .12). This indicates that 44% of variance in conversational efficiency scores and 40% of variance in conversational quality scores can be explained by entrainment scores. Furthermore, the prediction error (i.e., the average distance between the predicted value and real value) for conversational efficiency was approximately a minute

Figure 6. Predictive accuracy of entrainment models by entrainment type. Full represents models containing both entrainment types.
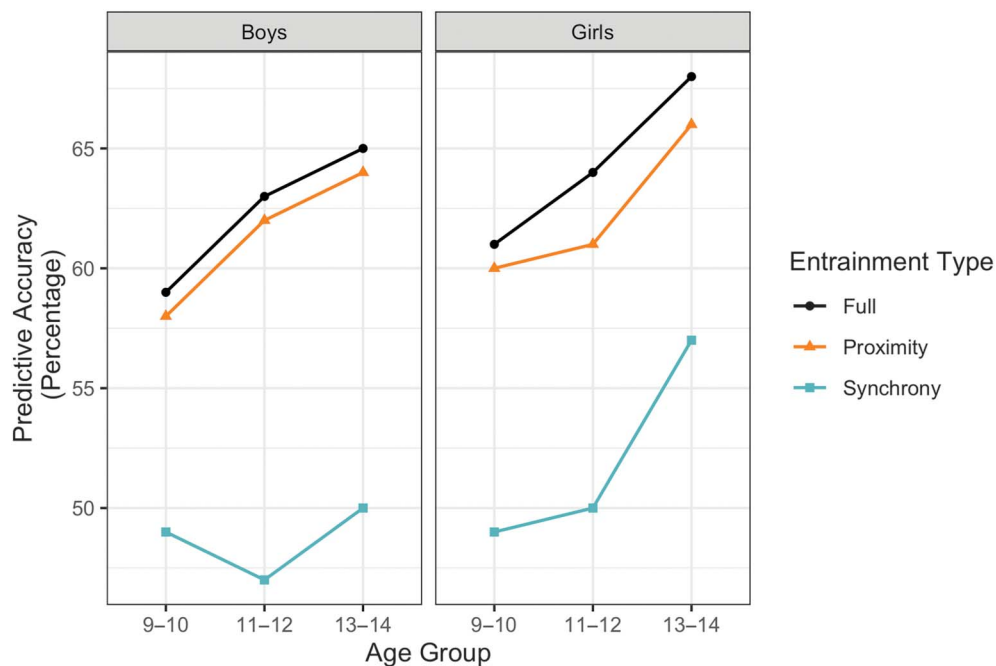
Figure 7. Comparison of actual conversational quality and conversational efficiency scores for each participant and model predicted scores based on entrainment data.


(from a range between about 1.5 and 10 min), whereas the prediction error for conversational quality was less than 1 unit (on a 7-point Likert scale).

Conversational Success by Acoustic Feature Set

Our next analysis focused on the relative importance of entrainment of each acoustic feature set on conversational success. Results are presented in Figure 8 (see also

Table 6 in Supplemental Material S3). For conversational efficiency, the feature set with the highest predictive accuracy was MFCC. This was followed by LTAS, EMS, voice report, and rhythm metrics. Results were similar for measures of conversational quality. Of note, the predictive accuracy was higher when all acoustic feature sets were present than when any feature set was considered in isolation.

Figure 8. Predictive accuracy for conversational success by acoustic feature set. Full represents models containing all acoustic feature sets. EMS = envelope modulation spectrum; LTAS = long-term average spectrum; MFCC = mel-frequency cepstrum coefficient; VR = voice report.
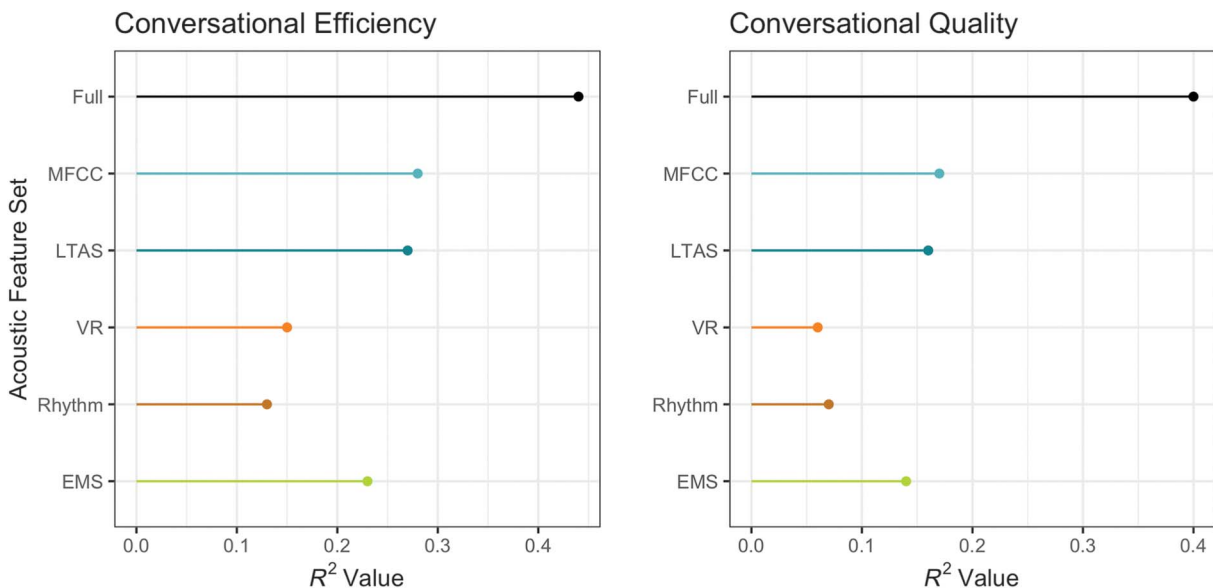
Conversational Success by Entrainment Type

Our final analysis focused on the relative importance of each entrainment type on conversational success. Results are presented in Figure 9 (see also Table 7 in Supplemental Material S3). For both conversational quality and conversational efficiency, proximity led to a higher predictive accuracy than synchrony. For conversational efficiency, predictive accuracy was higher when both synchrony and proximity scores were included in the model than when either type of entrainment was considered separately. However, for conversational quality, the model that only included proximity scores led to greater predictive accuracy than the model where both synchrony and proximity were included.

Discussion

The purpose of this study was to characterize the speech entrainment patterns of early adolescents. More specifically, we investigated the developmental trajectory of speech entrainment across early adolescents and the relationship between entrainment and conversational success in the conversations of this population. To gain a more detailed understanding of entrainment patterns in early adolescence, we also examined how this trajectory and the relationship between entrainment and success varied across different speech dimensions and entrainment types. While a small number of existing studies have explored speech entrainment of children and adolescents, this is the first study, to our knowledge, to study the entrainment patterns of this age group in naturalistic

conversations between peers and to consider multiple speech dimensions and entrainment types.

Our first objective was to characterize the developmental trajectory of speech entrainment in early adolescence. We hypothesized that we would see an increase in entrainment skills across early adolescent groups due to physical, cognitive, and social developmental changes, and for the given sample, this hypothesis was confirmed. Entrainment was present in the conversations of the 9- to 10-year group above the level of chance. However, the degree of entrainment present in these conversations was relatively small (i.e., predictive value was 59%). Entrainment levels increased in the 11- to 12-year group (predictive value was 61%) and subsequently in the 13- to 14-year group (i.e., predictive value was 64%). Importantly, this trajectory was present in both the holistic measure of entrainment and most stratified measures of entrainment across different acoustic feature sets and types of entrainment, illustrating the robustness of these developmental patterns. We also examined the entrainment patterns of the conversations of a comparative adult corpus. The predictive accuracy for adults within our study (i.e., 66%) was comparable to the predictive accuracy from perceptual studies (e.g., Aguilar et al., 2016; Pardo et al., 2018) or other studies that have employed predictive modeling (e.g., Ostrand & Chodroff, 2021; Willi et al., 2018), demonstrating the validity of our methodology for studying entrainment. Importantly, in this sample, predictive accuracy was substantively higher for adults than the 9- to 10-year group and 11- to 12-year group and slightly higher than the 13- to 14-year group. Thus, taken together, our findings

Figure 9. Predictive accuracy for conversational success by entrainment type. Full represents models containing both entrainment types.
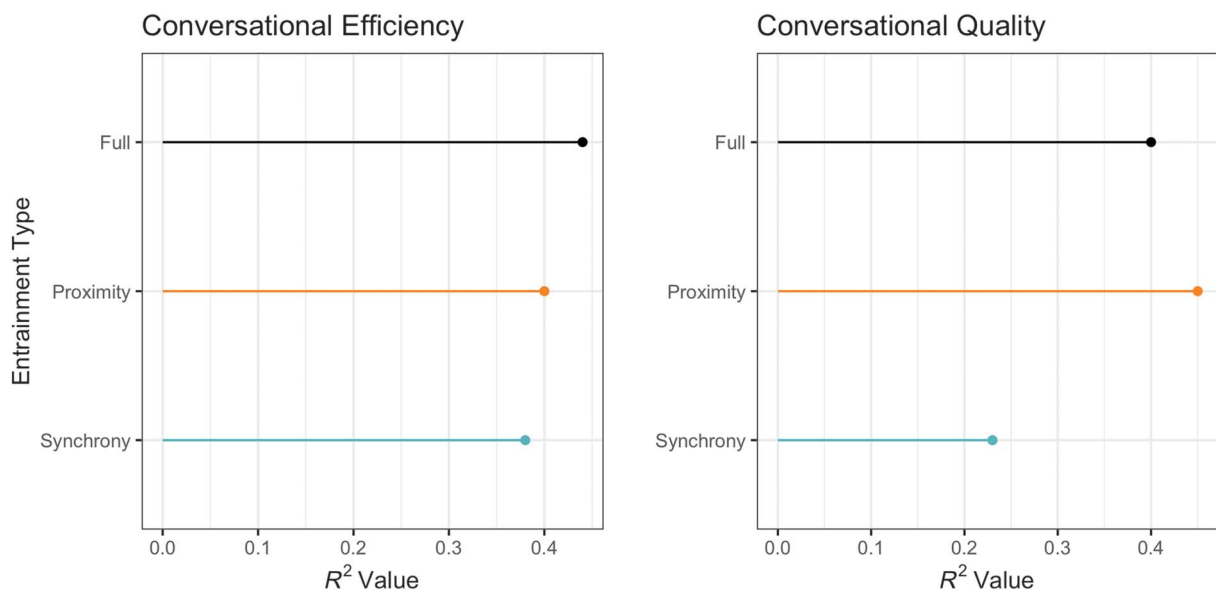
indicate considerable increases in entrainment skills across adolescence (i.e., considerable changes from 9/10 years to adulthood). However, this development is gradual, occurring slowly as adolescents develop the underlying skills necessary for entrainment to occur. That we see a small increase in entrainment between the 13- to 14-year-old group and adults suggests possible continued development into later adolescence. While more research is certainly needed to support such a conclusion, there is good theoretical reason to believe this may be the case. While some of the skills necessary for entrainment may plateau by the end of early adolescence, many continue to develop through later adolescents (e.g., speech motor control: Walsh & Smith, 2002; certain cognitive abilities: Hartshorne & Germine, 2015; Icenogle et al., 2019). Additionally, while peers play an important role in early adolescence, they take an even more prominent role in later adolescence (Buhrmester, 1990; Lam et al., 2014) with mid/late adolescents spending an estimated 50% of their waking hours engaged in peer interaction (Csikszentmihalyi & Larson, 1984). Regardless of the exact timeframe in which adult-like levels are attained, the protracted development of entrainment skills into adolescence implicates the complexity of entrainment and provides indirect empirical evidence in support of Lewandowski and Jilka's (2019) model in which many factors must be in place for entrainment to occur.

Our next finding regards differences in entrainment across gender. We found that, across all age groups in this sample, girls entrained slightly more than boys. These differences were, in some instances, small and, accordingly, should be interpreted with caution. However, we note that this pattern was present in both the holistic measure of entrainment and most stratified measures of entrainment across different acoustic feature sets and types of entrainment. These results are likely due to differences in peer interactions between girls and boys of this age group. For instance, early adolescent girls spend more time with their friends than boys (Larson & Richards, 1991). Additionally, girls are more likely to base friendship on intimacy, emotional support, and self-disclosure (e.g., Brendgen et al., 2001; Camarena et al., 1990; Radmacher & Azmitia, 2006) and, accordingly, spend nearly twice as much time engaged in conversation than boys (Raffaelli & Duckett, 1989). As such, increased opportunities for practice (through increased peer interaction) as well as increased motivation to entrain may account for the entrainment differences observed in this study.

Our second objective was to examine the relationship between entrainment and conversational success. To investigate this relationship, we used both an objective measure of conversational efficiency and a subjective measure of conversational quality. Our findings showed that

speech entrainment was highly predictive of both measures of conversational success, providing further validity to our current methodology. While these findings are in line with research showing similar results in adult conversations (e.g., Borrie et al., 2019; Gregory et al., 1997; Schweitzer et al., 2017), this is the first study, to our knowledge, to explore this relationship in child/adolescent conversations. The fact that entrainment exists within adolescent conversations and is predictive of conversational success carries important implications. The importance of friendship in adolescents' lives has been well documented in the literature. Beyond the myriad emotional benefits (Levitt et al., 1993; Lodder et al., 2017; Pachucki et al., 2015), adolescents with good friends perform better academically (Gallardo et al., 2016; Vaquera & Kao, 2008; Wentzel et al., 2018), develop better social skills over time (Glick & Rose, 2011; Rubin et al., 2004), and engage in less risky behaviors (Brady et al., 2009; Kamper & Ostrov, 2013; Mcelhaney et al., 2006). Given the importance of peer interactions, and the role of conversations in these interactions (e.g., Csikszentmihalyi & Larson, 1984; Larson, 2001; Raffaelli & Duckett, 1989), there is a recognized need to better understand the typical interactive patterns of this age group (Dahl et al., 2018; Turkstra, 2000; Turkstra et al., 2003). Here, we show that entrainment not only is present in the conversations of early adolescents but also plays an important role in conversational outcomes.

To gain a more in-depth view of entrainment, we examined how entrainment patterns (i.e., both the occurrence of entrainment and its relationships with conversational success) varied across speech dimension and entrainment type. We found that entrainment across feature sets that most strongly represent articulatory dimensions of speech (i.e., LTAS and MFCC) occurred at higher levels and was more indicative of conversational success than features sets that represent phonation (i.e., voice report) or rhythm (i.e., rhythm metrics and EMS). These findings are in line with Borrie et al. (2019) who, using a different methodological approach, found similar entrainment patterns across these same speech dimensions in adult conversations. We also found that proximity occurred at higher levels than synchrony. Importantly, models that included both proximity and synchrony across all five acoustic feature sets almost always led to higher predictive accuracy than models that only contained a single entrainment type or acoustic feature set. Several researchers have highlighted the need for holistic approaches that capture the multi-dimensionality of speech entrainment, rather than focusing on individual acoustic features (Borrie et al., 2019; Ostrand & Chodroff, 2021; Pardo et al., 2018). Here, we provide empirical support for these assertions and a valid methodology for capturing speech entrainment across different acoustic features and entrainment types.

Limitations and Future Directions

Given the relative novelty of our study and findings, there are many avenues for continued investigation in this area. First, through our findings, we showed an increase in the entrainment patterns across early adolescents, with the possibility of continued development in later adolescence. However, as our study did not include older adolescents, we were unable to determine the developmental trajectory of their entrainment skills. Accordingly, future work should focus on characterizing the speech entrainment patterns of this age group. Future research could also examine developmental patterns across different types of conversations, languages, and cultures. Additional types of entrainment (i.e., dynamic and/or global entrainment) could be explored as well. While we conjecture that our primary findings (i.e., that entrainment increases during adolescence and is predictive of conversational success) will be largely generalizable, more specific findings will likely vary across contexts. Next, our findings showed a developmental increase in speech entrainment across early adolescents. This is likely due to increases across motor production and cognitive abilities as well as social motivation. However, which specific developmental skills are most important for entrainment is still largely unknown. Therefore, future research investigating individual differences in the entrainment patterns of adolescents and the skills that underlie these differences is necessary. Finally, while our study focused on the speech entrainment patterns of neurotypical populations, our findings offer important clinical implications. It is likely that individuals with neurodevelopmental disorders affecting the underlying abilities/motivation necessary for speech entrainment may exhibit difficulties entraining to others. Currently, there are a couple of studies that have found evidence of entrainment deficits in autism (e.g., Lehnert-LeHouillier et al., 2020; Patel et al., 2022; Wynn et al., 2018), and it is likely that similar patterns are found in other neurodevelopmental disorders such as attention-deficit/hyperactivity disorder, global developmental delay, and various communication disorders. Accordingly, future work should focus on understanding the speech entrainment patterns of neurodivergent adolescents and the impact on conversational outcomes.

Conclusions

This study represents the first study, to our knowledge, to explore speech entrainment in peer conversations during adolescence. We found that speech entrainment increases during early adolescence and is predictive of both conversational quality and conversational efficiency. Furthermore, we demonstrated the benefits of using a multidimensional methodology that incorporates different types of entrainment across multiple speech dimensions. These findings offer a number of implications and future directions for continued investigation of entrainment in neurotypical and neurodivergent adolescent populations.

Author Contributions

Camille J. Wynn: Conceptualization (Lead), Formal analysis (Lead), Funding acquisition (Lead), Methodology (Lead), Project administration (Lead), Software (Supporting), Writing – original draft (Lead), Writing – review and editing (Lead). **Tyson S. Barrett:** Conceptualization (Supporting), Formal analysis (Lead), Funding acquisition (Supporting), Methodology (Supporting), Software (Supporting), Supervision (Supporting), Writing – original draft (Supporting), Writing – review and editing (Supporting). **Visar Berisha:** Formal analysis (Supporting), Funding acquisition (Supporting), Software (Lead), Writing – review and editing (Supporting). **Julie M. Liss:** Conceptualization (Supporting), Formal analysis (Supporting), Funding acquisition (Supporting), Writing – review and editing (Supporting). **Stephanie A. Borrie:** Conceptualization (Lead), Formal analysis (Supporting), Funding acquisition (Lead), Methodology (Lead), Project administration (Supporting), Supervision (Lead), Writing – original draft (Supporting), Writing – review and editing (Supporting).

Data Availability Statement

The data sets analyzed during this study are a part of the kidLUCID and LUCID corpora available in the SpeechBox repository at https://speechbox.linguistics.northwestern.edu/#!/home. Additionally, statistical code can be found at https://osf.io/zmbwd/.

Acknowledgments

References

Aguilar, L. J., Downey, G., Krauss, R. M., Pardo, J. S., Lane, S., & Bolger, N. (2016). A dyadic perspective on speech accommodation and social connection: Both partners'

rejection sensitivity matters. *Journal of Personality, 84*(2), 165–177. https://doi.org/10.1111/jopy.12149

Baker, R., & Hazan, V. (2011). DiapixUK: Task materials for the elicitation of multiple spontaneous speech dialogs. *Behavior Research Methods, 43*(3), 761–770. https://doi.org/10.3758/s13428-011-0075-y

Balaam, M., Fitzpatrick, G., Good, J., & Harris, E. (2011). Enhancing interactional synchrony with an ambient display. In *CHI '11: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 867–876). Association for Computing Machinery. https://doi.org/10.1145/1978942.1979070

Bernieri, F. J., Davis, J. M., Rosenthal, R., & Knee, C. R. (1994). Interactional synchrony and rapport: Measuring synchrony in displays devoid of sound and facial affect. *Personality and Social Psychology Bulletin, 20*(3), 303–311. https://doi.org/10.1177/0146167294203008

Boersma, P. (1993). Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound. In *Proceedings of the Institute of Phonetic Sciences* (Vol. 17, No. 1193, pp. 97–110), Amsterdam, the Netherlands. Institute of Phonetic Sciences, University of Amsterdam.

Boersma, P., & Weenink, D. (2020). *Praat: Doing phonetics by computer* (version 6.1) [computer software]. http://www.praat.org

Borrie, S. A., Barrett, T. S., Liss, J. M., & Berisha, V. (2020). Sync pending: Characterizing conversational entrainment in dysarthria using a multidimensional, clinically informed approach. *Journal of Speech, Language, and Hearing Research, 63*(1), 83–94. https://doi.org/10.1044/2019_JSLHR-19-00194

Borrie, S. A., Barrett, T. S., Willi, M. M., & Berisha, V. (2019). Syncing up for a good conversation: A clinically meaningful methodology for capturing conversational entrainment in the speech domain. *Journal of Speech, Language, and Hearing Research, 62*(2), 283–296. https://doi.org/10.1044/2018_JSLHR-S-18-0210

Borrie, S. A., & Delfino, C. R. (2017). Conversational entrainment of vocal fry in young adult female American English speakers. *Journal of Voice, 31*(4), P513.e25–513.e32. https://doi.org/10.1016/j.jvoice.2016.12.005

Borrie, S. A., Lubold, N., & Pon-Barry, H. (2015). Disordered speech disrupts conversational entrainment: A study of acoustic-prosodic entrainment and communicative success in populations with communication challenges. *Frontiers in Psychology, 6,* 1187. https://doi.org/10.3389/fpsyg.2015.01187

Borrie, S. A., Wynn, C. J., Berisha, V., Lubold, N., Willi, M. M., Coelho, C. A., & Barrett, T. S. (2020). Conversational coordination of articulation responds to context: A clinical test case with traumatic brain injury. *Journal of Speech, Language, and Hearing Research, 63*(8), 2567–2577. https://doi.org/10.1044/2020_JSLHR-20-00104

Bradlow, A. R. (n.d.) *SpeechBox*. Retrieved March 8, 2023, from https://speechbox.linguistics.northwestern.edu

Brady, S. S., Dolcini, M. M., Harper, G. W., & Pollack, L. M. (2009). Supportive friendships moderate the association between stressful life events and sexual risk taking among African American adolescents. *Health Psychology: Official Journal of the Division of Health Psychology, American Psychological Association, 28*(2), 238–248. https://doi.org/10.1037/a0013240

Brendgen, M., Markiewicz, D., Doyle, A. B., & Bukowski, W. M. (2001). The relations between friendship quality, ranked-friendship preference, and adolescents' behavior with their friends. *Merrill-Palmer Quarterly, 47*(3), 395–415. https://doi.org/10.1353/mpq.2001.0013

Buhrmester, D. (1990). Intimacy of friendship, interpersonal competence, and adjustment during preadolescence and

adolescence. *Child Development, 61*(4), 1101–1111. https://doi.org/10.2307/1130878

Buhrmester, D., & Furman, W. (1987). The development of companionship and intimacy. *Child Development, 58*(4), 1101–1113. https://doi.org/10.2307/1130550

Camarena, P. M., Sarigiani, P. A., & Petersen, A. C. (1990). Gender-specific pathways to intimacy in early adolescence. *Journal of Youth and Adolescence, 19*(1), 19–32. https://doi.org/10.1007/BF01539442

Carriedo, N., Corral, A., Montoro, P. R., Herrero, L., & Rucián, M. (2016). Development of the updating executive function: From 7-year-olds to young adults. *Developmental Psychology, 52*(4), 666–678. https://doi.org/10.1037/dev0000091

Csikszentmihalyi, M., & Larson, R. (1984). *Being adolescent: Conflict and growth in the teenage years*. Basic Books.

Dahl, R. E., Allen, N. B., Wilbrecht, L., & Suleiman, A. B. (2018). Importance of investing in adolescence from a developmental science perspective. *Nature, 554*(7693), 441–450. https://doi.org/10.1038/nature25770

Davis, S. B., & Mermelstein, P. (1990). Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences. In A. Waibel & K.-F. Lee (Eds.), *Readings in speech recognition* (pp. 65–74). Morgan Kaufmann. https://doi.org/10.1016/B978-0-08-051584-7.50010-3.

De Looze, C., Scherer, S., Vaughan, B., & Campbell, N. (2014). Investigating automatic measurements of prosodic accommodation and its dynamics in social interaction. *Speech Communication, 58,* 11–34. https://doi.org/10.1016/j.specom.2013.10.002

Demetriou, A., Christou, C., Spanoudis, G., & Platsidou, M. (2002). The development of mental processing: Efficiency, working memory, and thinking. *Monographs of the Society for Research in Child Development, 67*(1), vii–viii. https://doi.org/10.1111/1540-5834.671173

Dreyfuss, M., Caudle, K., Drysdale, A. T., Johnston, N. E., Cohen, A. O., Somerville, L. H., Galván, A., Tottenham, N., Hare, T. A., & Casey, B. J. (2014). Teens impulsively react rather than retreat from threat. *Developmental Neuroscience, 36*(3–4), 220–227. https://doi.org/10.1159/000357755

Dideriksen, C., Christiansen, M. H., Tylén, K., Dingemanse, M., & Fusaroli, R. (2022). Quantifying the interplay of conversational devices in building mutual understanding. *Journal of Experimental Psychology: General*. Advance online publication. https://doi.org/10.1037/xge0001301

Feldman, S. S., & Elliott, G. R. (1990). *At the threshold: The developing adolescent*. Harvard University Press.

Ferguson, H. J., Brunsdon, V. E. A., & Bradford, E. E. F. (2021). The developmental trajectories of executive function from adolescence to old age. *Scientific Reports, 11*(1), 1382. https://doi.org/10.1038/s41598-020-80866-1

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

Gallardo, L. O., Barrasa, A., & Guevara-Viejo, F. (2016). Positive peer relationships and academic achievement across early and midadolescence. *Social Behavior and Personality: An International Journal, 44*(10), 1637–1648. https://doi.org/10.2224/sbp.2016.44.10.1637

Giles, H., Coupland, J., & Coupland, N. (1991). *Contexts of accommodation: Developments in applied sociolinguistics* (1st ed.). Cambridge University Press. https://doi.org/10.1017/CBO9780511663673

Glick, G. C., & Rose, A. J. (2011). Prospective associations between friendship adjustment and social strategies:

Friendship as a context for building social skills. *Developmental Psychology, 47*(4), 1117–1132. https://doi.org/10.1037/a0023277

Gregory, S. W., Dagan, K., & Webster, S. (1997). Evaluating the relation of vocal accommodation in conversation partners' fundamental frequencies to perceptions of communication quality. *Journal of Nonverbal Behavior, 21*(1), 23–43. https://doi.org/10.1023/A:1024995717773

Hartshorne, J. K., & Germine, L. T. (2015). When does cognitive functioning peak? The asynchronous rise and fall of different cognitive abilities across the life span. *Psychological Science, 26*(4), 433–443. https://doi.org/10.1177/0956797614567339

Hazan, V., Tuomainen, O., & Pettinato, M. (2016). Suprasegmental characteristics of spontaneous speech produced in good and challenging communicative conditions by talkers aged 9-14 years. *Journal of Speech, Language, and Hearing Research, 59*(6), S1596–S1607. https://doi.org/10.1044/2016_JSLHR-S-15-0046

Helseth, S., & Misvaer, N. (2010). Adolescents' perceptions of quality of life: What it is and what matters. *Journal of Clinical Nursing, 19*(9–10), 1454–1461. https://doi.org/10.1111/j.1365-2702.2009.03069.x

Hill, J. P. (1980). *Understanding early adolescence: A framework*. Center for Early Adolescence.

Icenogle, G., Steinberg, L., Duell, N., Chein, J., Chang, L., Chaudhary, N., Di Giunta, L., Dodge, K. A., Fanti, K. A., Lansford, J. E., Oburu, P., Pastorelli, C., Skinner, A. T., Sorbring, E., Tapanya, S., Uribe Tirado, L. M., Alampay, L. P., Al-Hassan, S. M., Takash, H. M. S., & Bacchini, D. (2019). Adolescents' cognitive capacity reaches adult levels prior to their psychosocial maturity: Evidence for a "maturity gap" in a multinational, cross-sectional sample. *Law and Human Behavior, 43*(1), 69–85. https://doi.org/10.1037/lhb0000315

Ingham, R. J., Sato, W., Finn, P., & Belknap, H. (2001). The modification of speech naturalness during rhythmic stimulation treatment of stuttering. *Journal of Speech, Language, and Hearing Research, 44*(4), 841–852. https://doi.org/10.1044/1092-4388(2001/066)

Kail, R. V., & Ferrer, E. (2007). Processing speed in childhood and adolescence: Longitudinal models for examining developmental change. *Child Development, 78*(6), 1760–1770. https://doi.org/10.1111/j.1467-8624.2007.01088.x

Kamper, K. E., & Ostrov, J. M. (2013). Relational aggression in middle childhood predicting adolescent social-psychological adjustment: The role of friendship quality. *Journal of Clinical Child & Adolescent Psychology, 42*(6), 855–862. https://doi.org/10.1080/15374416.2013.844595

Karns, C. M., Isbell, E., Giuliano, R. J., & Neville, H. J. (2015). Auditory attention in childhood and adolescence: An event-related potential study of spatial selective attention to one of two simultaneous stories. *Developmental Cognitive Neuroscience, 13*, 53–67. https://doi.org/10.1016/j.dcn.2015.03.001

Kuhn, M. (2017). *Caret: Classification and regression training (R package version 6.0-90)*. https://CRAN.R-project.org/package=caret

La Gaipa, J. J. (1979). A developmental study of the meaning of friendship in adolescence. *Journal of Adolescence, 2*(3), 201–213. https://doi.org/10.1016/S0140-1971(79)80012-3

Lam, C. B., McHale, S. M., & Crouter, A. C. (2014). Time with peers from middle childhood to late adolescence: Developmental course and adjustment correlates. *Child Development, 85*(4), 1677–1693. https://doi.org/10.1111/cdev.12235

Larson, R. W. (1983). Adolescents' daily experience with family and friends: Contrasting opportunity systems. *Journal of Marriage and Family, 45*(4), 739–750. https://doi.org/10.2307/351787

Larson, R. W. (2001). How U.S. children and adolescents spend time: What it does (and doesn't) tell us about their development. *Current Directions in Psychological Science, 10*(5), 160–164. https://doi.org/10.1111/1467-8721.00139

Larson, R. W., & Richards, M. H. (1991). Daily companionship in late childhood and early adolescence: Changing developmental contexts. *Child Development, 62*(2), 284–300. https://doi.org/10.2307/1131003

Lehnert-LeHouillier, H., Terrazas, S., & Sandoval, S. (2020). Prosodic entrainment in conversations of verbal children and teens on the autism spectrum. *Frontiers in Psychology, 11*, 582221. https://doi.org/10.3389/fpsyg.2020.582221

Levitt, M. J., Guacci-Franco, N., & Levitt, J. L. (1993). Convoys of social support in childhood and early adolescence: Structure and function. *Developmental Psychology, 29*(5), 811–818. https://doi.org/10.1037/0012-1649.29.5.811

Lewandowski, N. (2012). *Talent in nonnative phonetic convergence*. Doctoral dissertation, Institut für Maschinelle Sprachverarbeitung, Universität Stuttgart.

Lewandowski, N., & Jilka, M. (2019). Phonetic convergence, language talent, personality and attention. *Frontiers in Communication, 4*, Article 18. https://doi.org/10.3389/fcomm.2019.00018

Liss, J. M., LeGendre, S., & Lotto, A. J. (2010). Discriminating dysarthria type from envelope modulation spectra. *Journal of Speech, Language, and Hearing Research, 53*(5), 1246–1255. https://doi.org/10.1044/1092-4388(2010/09-0121)

Lodder, G. M. A., Scholte, R. H. J., Goossens, L., & Verhagen, M. (2017). Loneliness in early adolescence: Friendship quantity, friendship quality, and dyadic processes. *Journal of Clinical Child & Adolescent Psychology, 46*(5), 709–720. https://doi.org/10.1080/15374416.2015.1070352

Lubold, N., Borrie, S. A., Barrett, T. S., Willi, M., & Berisha, V. (2019). Do conversational partners entrain on articulatory precision? In *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH* (pp. 1931–1935). https://doi.org/10.21437/Interspeech.2019-1786

Lubold, N., & Pon-Barry, H. (2014). Acoustic-prosodic entrainment and rapport in collaborative learning dialogues. In *MLA '14: Proceedings of the 2014 ACM workshop on multimodal learning analytics workshop and grand challenge* (pp. 5–12). Association for Computing Machinery.

Luna, B., Garver, K. E., Urban, T. A., Lazar, N. A., & Sweeney, J. A. (2004). Maturation of cognitive processes from late childhood to adulthood. *Child Development, 75*(5), 1357–1372. https://doi.org/10.1111/j.1467-8624.2004.00745.x

Manson, J. H., Bryant, G. A., Gervais, M. M., & Kline, M. A. (2013). Convergence of speech rate in conversation predicts cooperation. *Evolution and Human Behavior, 34*(6), 419–426. https://doi.org/10.1016/j.evolhumbehav.2013.08.001

Mcelhaney, K. B., Immele, A., Smith, F. D., & Allen, J. P. (2006). Attachment organization as a moderator of the link between friendship quality and adolescent delinquency. *Attachment & Human Development, 8*(1), 33–46. https://doi.org/10.1080/14616730600585250

Memmert, D. (2014). Inattentional blindness to unexpected events in 8–15-year-olds. *Cognitive Development, 32*, 103–109. https://doi.org/10.1016/j.cogdev.2014.09.002

Michalsky, J., & Schoormann, H. (2017). Pitch convergence as an effect of perceived attractiveness and likability. In *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH* (pp. 2254–2256).

Mizuno, K., Tanaka, M., Fukuda, S., Sasabe, T., Imai-Matsumura, K., & Watanabe, Y. (2011). Changes in cognitive functions of students in the transitional period from

elementary school to junior high school. *Brain & Development, 33*(5), 412–420. https://doi.org/10.1016/j.braindev.2010. 07.005

Natale, M. (1975). Convergence of mean vocal intensity in dyadic communication as a function of social desirability. *Journal of Personality and Social Psychology, 32*(5), 790–804. http://doi. org.dist.lib.usu.edu/10.1037/0022-3514.32.5.790

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. https://doi.org/10.1016/j.wocn.2021.101074

Oviatt, S., Darves, C., & Coulston, R. (2004). Toward adaptive conversational interfaces: Modeling speech convergence with animated personas. *ACM Transactions on Computer-Human Interaction (TOCHI), 11*(3), 300–328. https://doi.org/10.1145/ 1017494.1017498

Pachucki, M. C., Ozer, E. J., Barrat, A., & Cattuto, C. (2015). Mental health and social networks in early adolescence: A dynamic study of objectively-measured social interaction behaviors. *Social Science & Medicine, 125,* 40–50. https://doi. org/10.1016/j.socscimed.2014.04.015

Pardo, J. S. (2013). Measuring phonetic convergence in speech production. *Frontiers in Psychology, 4,* 559. https://doi.org/10. 3389/fpsyg.2013.00559

Pardo, J. S., Urmanche, A., Wilman, S., Wiener, J., Mason, N., Francis, K., & Ward, M. (2018). A comparison of phonetic convergence in conversational interaction and speech shadowing. *Journal of Phonetics, 69,* 1–11. https://doi.org/10.1016/j.wocn. 2018.04.001

Patel, S. P., Cole, J., Lau, J. C. Y., Fragnito, G., & Losh, M. (2022). Verbal entrainment in autism spectrum disorder and first-degree relatives. *Scientific reports, 12*(1), Article 11496. https://doi.org/10.1038/s41598-022-12945-4

Petrone, C., D'Alessandro, D., & Falk, S. (2021). Working memory differences in prosodic imitation. *Journal of Phonetics, 89,* 101100. https://doi.org/10.1016/j.wocn.2021.101100

Phillips-Silver, J., Aktipis, C. A., & Bryant, G. A. (2010). The ecology of entrainment: Foundations of coordinated rhythmic movement. *Music Perception: An Interdisciplinary Journal, 28*(1), 3–14. https://doi.org/10.1525/mp.2010.28.1.3

Platt, B., Kadosh, K. C., & Lau, J. Y. F. (2013). The role of peer rejection in adolescent depression. *Depression and Anxiety, 30*(9), 809–821. https://doi.org/10.1002/da.22120

Polyanskaya, L., Samuel, A. G., & Ordin, M. (2019). Speech rhythm convergence as a social coalition signal. *Evolutionary Psychology, 17*(3), Article 1474704919879335. https://doi.org/ 10.1177/1474704919879335

Radmacher, K., & Azmitia, M. (2006). Are there gendered pathways to intimacy in early adolescents' and emerging adults' friendships? *Journal of Adolescent Research, 21*(4), 415–448. https://doi.org/10.1177/0743558406287402

Raffaelli, M., & Duckett, E. (1989). "We were just talking . . .": Conversations in early adolescence. *Journal of Youth and Adolescence, 18*(6), 567–582. https://doi.org/10.1007/BF02139074

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

Reichel, U. D., Beňuš, Š., & Mády, K. (2018). Entrainment profiles: Comparison by gender, role, and feature set. *Speech Communication, 100,* 46–57. https://doi.org/10.1016/j.specom. 2018.04.009

Rubia, K., Smith, A. B., Woolley, J., Nosarti, C., Heyman, I., Taylor, E., & Brammer, M. (2006). Progressive increase of frontostriatal brain activation from childhood to adulthood during event-related tasks of cognitive control. *Human Brain Mapping, 27*(12), 973–993. https://doi.org/10.1002/hbm.20237

Rubin, K. H., Dwyer, K. M., Booth-LaForce, C., Kim, A. H., Burgess, K. B., & Rose-Krasnor, L. (2004). Attachment, friendship, and psychosocial functioning in early adolescence. *The Journal of Early Adolescence, 24*(4), 326–356. https://doi. org/10.1177/0272431604268530

Schertz, J., & Johnson, E. K. (2022). Voice onset time imitation in teens versus adults. *Journal of Speech, Language, and Hearing Research, 65*(5), 1839–1850. https://doi.org/10.1044/2022_ JSLHR-21-00460

Schweitzer, A., & Lewandowski, N. (2013). Convergence of articulation rate in spontaneous speech. In *Proceedings of the 14th Annual Conference of the International Speech Communication Association, INTERSPEECH* (pp. 525–529).

Schweitzer, A., Lewandowski, N., & Duran, D. (2017). Social attractiveness in dialogs. In *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH* (pp. 2243–2247).

Sebastian, C., Viding, E., Williams, K. D., & Blakemore, S.-J. (2010). Social brain development and the affective consequences of ostracism in adolescence. *Brain and Cognition, 72*(1), 134–145. https://doi.org/10.1016/j.bandc.2009.06.008

Smith, A., & Zelaznik, H. N. (2004). Development of functional synergies for speech motor coordination in childhood and adolescence. *Developmental Psychobiology, 45*(1), 22–33. https://doi.org/10.1002/dev.20009

Somerville, L. H. (2013). Special issue on the teenage brain: Sensitivity to social evaluation. *Current Directions in Psychological Science, 22*(2), 121–127. https://doi.org/10.1177/ 0963721413476512

Somerville, L. H., Hare, T., & Casey, B. J. (2011). Frontostriatal maturation predicts cognitive control failure to appetitive cues in adolescents. *Journal of Cognitive Neuroscience, 23*(9), 2123–2134. https://doi.org/10.1162/jocn.2010.21572

Somerville, L. H., Jones, R. M., Ruberry, E. J., Dyke, J. P., Glover, G., & Casey, B. J. (2013). The medial prefrontal cortex and the emergence of self-conscious emotion in adolescence. *Psychological Science, 24*(8), 1554–1562. https://doi.org/ 10.1177/0956797613475633

Steinberg, L. (2020). *Adolescence* (12th ed.). McGraw-Hill.

Stroud, L. R., Foster, E., Papandonatos, G. D., Handwerger, K., Granger, D. A., Kivlighan, K. T., & Niaura, R. (2009). Stress response and the adolescent transition: Performance versus peer rejection stressors. *Development and Psychopathology, 21*(1), 47–68. https://doi.org/10.1017/S0954579409000042

Todd, N. P. M., Lee, C. S., & O'Boyle, D. J. (2002). A sensorimotor theory of temporal tracking and beat induction. *Psychological Research, 66*(1), 26–39. https://doi.org/10.1007/s004260100071

Turkstra, L. S. (2000). Should my shirt be tucked in or left out? The communication context of adolescence. *Aphasiology, 14*(4), 349–364. https://doi.org/10.1080/026870300401405

Turkstra, L. S., Ciccia, A., & Seaton, C. (2003). Interactive behaviors in adolescent conversation dyads. *Language, Speech, and Hearing Services in Schools, 34*(2), 117–127. https://doi.org/ 10.1044/0161-1461(2003/010)

Van Engen, K. J., Baese-Berk, M., Baker, R. E., Choi, A., Kim, M., & Bradlow, A. R. (2010). The Wildcat Corpus of native and foreign-accented English: Communicative efficiency across conversational dyads with varying language alignment profiles. *Language and Speech, 53*(4), 510–540. https://doi.org/10.1177/ 0023830910372495

Vaquera, E., & Kao, G. (2008). Do you like me as much as I like you? Friendship reciprocity and its effects on school outcomes

among adolescents. *Social Science Research, 37*(1), 55–72. https://doi.org/10.1016/j.ssresearch.2006.11.002

Walsh, B., & Smith, A. (2002). Articulatory movements in adolescents: Evidence for protracted development of speech motor control processes. *Journal of Speech, Language, and Hearing Research, 45*(6), 1119–1133. https://doi.org/10.1044/1092-4388(2002/090)

Wentzel, K. R., Jablansky, S., & Scalise, N. R. (2018). Do friendships afford academic benefits? A meta-analytic study. *Educational Psychology Review, 30*(4), 1241–1267. https://doi.org/10.1007/s10648-018-9447-5

Willi, M. M., Borrie, S. A., Barrett, T. S., Tu, M., & Berisha, V. (2018). A discriminative acoustic-prosodic approach for measuring local entrainment. In *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH* (pp. 581–585).

Williams, B. R., Ponesse, J. S., Schachar, R. J., Logan, G. D., & Tannock, R. (1999). Development of inhibitory control across the life span. *Developmental Psychology, 35*(1), 205–213. https://doi.org/10.1037//0012-1649.35.1.205

Wynn, C. J., Barrett, T. S., & Borrie, S. A. (2022). Rhythm perception, speaking rate entrainment, and conversational quality: A mediated model. *Journal of Speech, Language, and Hearing Research, 65*(6), 2187–2203. https://doi.org/10.1044/2022_JSLHR-21-00293

Wynn, C. J., & Borrie, S. A. (2020). Methodology matters: The impact of research design on conversational entrainment outcomes. *Journal of Speech, Language, and Hearing Research, 63*(5), 1352–1360. https://doi.org/10.1044/2020_JSLHR-19-00243

Wynn, C. J., & Borrie, S. A. (2022). Classifying conversational entrainment of speech behavior: An expanded framework and review. *Journal of Phonetics, 94,* 101173. https://doi.org/10.1016/j.wocn.2022.101173

Wynn, C. J., Borrie, S. A., & Pope, K. A. (2019). Going with the flow: An examination of entrainment in typically developing children. *Journal of Speech, Language, and Hearing Research, 62*(10), 3706–3713. https://doi.org/10.1044/2019_JSLHR-S-19-0116

Wynn, C. J., Borrie, S. A., & Sellers, T. P. (2018). Speech rate entrainment in children and adults with and without autism spectrum disorder. *American Journal of Speech-Language Pathology, 27*(3), 965–974. https://doi.org/10.1044/2018_AJSLP-17-0134

Yu, A. C. L., Abrego-Collier, C., & Sonderegger, M. (2013). Phonetic imitation from an individual-difference perspective: Subjective attitude, personality and "autistic" traits. *PLOS ONE, 8*(9), Article e74746. https://doi.org/10.1371/journal.pone.0074746