An Examination of Entrainment in Typically Developing Children

Kiersten A. Pope
Utah State University

Follow this and additional works at: https://digitalcommons.usu.edu/gradreports

Part of the Speech Pathology and Audiology Commons

Recommended Citation
Pope, Kiersten A., "An Examination of Entrainment in Typically Developing Children" (2018). All Graduate Plan B and other Reports. 1339.
https://digitalcommons.usu.edu/gradreports/1339

This Report is brought to you for free and open access by the Graduate Studies at DigitalCommons@USU. It has been accepted for inclusion in All Graduate Plan B and other Reports by an authorized administrator of DigitalCommons@USU. For more information, please contact digitalcommons@usu.edu.
AN EXAMINATION OF ENTRAINMENT IN
TYPICALLY DEVELOPING
CHILDREN

by

Kiersten A. Pope

A thesis submitted in partial fulfillment
of the requirements for the degree

of

MASTER OF SCIENCE

in

Speech-Language Pathology

Approved:

Stephanie A. Borrie, Ph.D.
Major Professor

Ronald B. Gillam, Ph.D.
Committee Member

Tyra Sellers, Ph.D.
Committee Member

Laurens H. Smith, Ph.D.
Interim Vice President for Research and
Dean of the School of Graduate Studies

UTAH STATE UNIVERSITY
Logan, Utah

2018
ACKNOWLEDGMENTS

Thank you to Stephanie Borrie for the mentorship, support, and encouragement needed throughout this process. Thank you to Camille Wynn for experiencing this with me in our simultaneous excitement and frustrations. Thank you to the Human Interaction Lab members for their support, hard work, and dedication to this project. And especially thank you to my family, my friends, and my graduate school cohort.

I could not have done this without them.

Kiersten Pope
ABSTRACT

Purpose: Conversational entrainment describes the tendency for individuals to align their behavior with their communication partners and is essential for successful interaction. Evidence of entrainment in adults is robust, yet research regarding its development is sparse. Here, we investigate the effectiveness of a quasi-conversational paradigm for the purpose of identifying the speech rate entrainment abilities of children.

Method: Data were collected from a total of 50 typically developing children from 5-14 years old. Participants completed an entrainment task to identify the presence of speech rate modification depending on the presence of “fast” or “slow” stimuli. The entrainment task utilized a quasi-conversational design requiring participants to listen and respond to video clips with a random presentation of speech rate conditions.

Results: Results indicated no significant difference between conditions for speech rate production of any age. Taken with previous literature, analysis suggests a lack of sensitivity in the current paradigm.

Conclusions: Understanding how entrainment develops in children will increase knowledge regarding communication in disordered populations, and provide insight into the mechanisms of typical conversational entrainment. By building upon the current paradigm, future research moves toward the establishment of a methodology that simulates entrainment effects in conversation while reasonably eliminating confounding factors of human-to-human interaction.
INTRODUCTION

During communication, individuals often describe an effortless, comfortable, or smooth conversation as one with good “flow”. This flow is considered to be the result of the tendency for people to align their verbal and non-verbal behaviors during communicative interaction (Beňuš, 2014). That is, over the course of a conversation, speakers will adjust their communicative behaviors in order to become more like one another. For example, one conversational partner may speak at a habitually fast rate. When conversing with a partner who speaks at a habitually slow rate, each partner will adjust their speech rate so that they become more similar (Freud, Ezrati-Vinacour, & Amir, 2018). These adaptations, made during regular interactions, are not performed at a conscious level, suggesting that this interchange is woven into the fabric of human interaction (Chartrand & Bargh, 1999). Researchers have coined various names for this phenomenon, including interactional alignment (Pickering & Garrod, 2004), communication accommodation (Street, Street, & Van Kleek, 1983), and conversational convergence (Welkowitz et. al, 1974). To maintain consistency among the most recent work, the term conversational entrainment will be used. Regardless of the label, the basic principle of synchrony and mutual adaptation remains intact across findings.

There is mounting evidence for entrainment in a wide variety of communicative behaviors. Yawning (e.g. Norscia & Palagi, 2011), facial expressions (e.g. Dimberg, Thunberg, & Elmehed, 2000), breathing (e.g. McFarland, 2001) and laughter (Provine, 2004; Scott, Lavan, Chen, & McGettigan, 2014) have all been shown to entrain among communication partners. As such, entrainment is a multisensory process relying not only
on auditory information from the opposing speaker but also on visual feedback (Athreya, Riley, & Davis, 2014; Phillips-Silver, Aktipis, & Bryant, 2010). Additionally, a strong neurological basis for entrainment exists. When an individual is listening to a speech signal, their neural responses in the auditory cortex have the ability to “phase-lock” onto the external stimuli, a concept that is usually designated as “neural entrainment” (Lakatos, Karmos, Mehta, Ulbert, & Schroeder, 2008, p. 110). This is a dynamic system, as a wide variety of acoustic and prosodic information such as speech rate, amplitude, and stress, is processed in real time. Using these rhythmic speech patterns, the brain is able to predict forthcoming signals, ultimately increasing not only the synchronicity of partners but also their ability to be understood (Peelle & Davis, 2012).

More specifically, conversational entrainment itself serves many vital roles in human interaction. Among these are a myriad of cognitive, social, and emotional functions. From a cognitive perspective, when conversational partners entrain with one another, the efficiency of communication increases. Gill (2012, p. 111) underlined the importance of this concept when she said that “information transfer is managed with prosodic and rhythmically synchronized movements of body and voice”. Additionally, Pickering and Garrod (2004) suggest that individuals draw on the lexical, semantic, syntactic, and phonologic representations of their partners to aid in the construction of their own message, thus reducing the necessary cognitive resources for the exchange.

Emotionally, entrainment can serve as a point of initiation and interest. Ireland and colleagues (2011, p. 41) found indications that “similarity in language style uniquely predicts relationship initiation beyond self-reported similarity” when looking at participants in a speed dating event. Manson, Bryant, Gervais and Kline (2013)
discovered a correlation between speech rate convergence and cooperation among strangers during a partner task. Dunbar and Mejia (2013) reported similar findings with married couples with couples who showed higher levels of entrainment reporting a more equal power balance within their relationship. In a review of entrainment literature, Beňuš (2014, p. 807) observed that the “degree of entrainment correlates not only with social distance, belief about communicative competence, or likeable and engaging personalities, but it might also correlate with empathy”. Thus, conversational entrainment serves as a facilitator for emotional connection.

Lastly, research indicates that entrainment encompasses a wide variety of social functions. Gill (2012) describes entrainment as a “social survival” skill, as it contributes substantially to personal relationships. Indeed, as entrainment serves both cognitive and emotional functions, these domains play key roles in the initiation, development, and maintenance of social rapport. The interaction of speech and non-speech features lend themselves to turn taking actions, and individuals are able to identify when to transition between speakers (Local, 2007). Conversely, conversations containing longer response latencies, a sign of poorly maintained interaction, are “negatively associated with entrainment” (Levitan et al., 2012, p.18). Consequently, conversational entrainment is an important, overarching mechanism of social interaction, and lends itself to the quality of interpersonal communication.

The evidence for speech entrainment in typical adults is relatively robust across many acoustic-prosodic features. Research suggests that during conversation adults entrain on a variety of elements including intensity, pitch, speaking rate and voice quality (e.g. Borrie & Delfino, 2017; Borrie & Liss, 2014; Gregory, 1990; Levitan & Hirschberg,
Pardo and colleagues have found extensive evidence supporting the convergence of phonetic productions over the course of a conversation, which is a finding that has been confirmed by subsequent research teams (Abney, Kello, & Warlaumont, 2015; Dias & Rosenblum, 2016; Pardo, 2006, 2013; Pardo, Urmanche, Wilman, & Wiener, 2017). The features of entrainment have also been documented across many languages, suggesting that processes underlying entrainment to speech rhythms would not be unique to language (Beňuš, 2014; Kawamori, Kawabata, & Shimazu, 1998; Levitan, Beňuš, Gravano, & Hirschberg, 2015; Lidji, Palmer, Peretz, & Morningstar, 2011). The wide variety of speech entrainment research provides a rich evidence base for the prevalence of entrainment among adults.

One specific prosodic feature often examined within entrainment research is that of speech rate. Speech rate is ideal to focus on in exploratory studies as it is relatively easy to measure and manipulate, and contributes to the rhythmic properties of speech. Current literature contains paradigms of speech rate manipulation including both digital reduction (Jungers & Hupp, 2009; Oviatt, Darves, & Coulston, 2004) and human speaker reduction of rate (Guitar & Marchinkoski, 2001; Street et al., 1983). Additionally, speech rate is a clinically relevant feature, as a slow speech rate is often used by clinicians to provide a model for a wide variety of disordered populations, including fluency disorders, dysarthria, apraxia, and voice disorders (Freud et al., 2018). For the purposes of the current study, and to offer comparable methodologies and findings, entrainment of speech rate will be the focus of further discussion.

While evidence of speech rate entrainment in adults is strong, the same cannot be said relating to children. The literature about speech rate entrainment in children is not
only scarce, but findings are equivocal. Some studies have identified the presence of speech rate entrainment in children of various ages. Welkowitz and colleagues (1976) found a convergence of pause duration in conversations between school aged children. They discovered that the degree of convergence was greater in older children (6;9 years compared to 5;8 years). Thus, they hypothesized that entrainment ability is a function of age. With data comprising multiple conversational dyads across several days, Ko and colleagues (2016, p. 2) found “small, but significant, effects of mothers and their [12-30 month old] children influencing each other’s speech”. Guitar and Marchinkoski (2001) used a single subject design with three year old participants and their mothers. This team concluded that in five out of six mother-child dyads, the child participants significantly altered their speech rate to match their mothers in the slow speech rate condition. Though these studies represent emerging evidence relating to entrainment development, they are not yet representative of a larger corpus of scientific findings, as seen in adult populations.

But while some findings support the presence of speech rate entrainment abilities in a wide variety of ages, other research teams have yet to identify a significant level of conversational entrainment, whether in natural or simulated conversation. Bernstein-Ratner (1992) used a natural play environment to elicit conversational samples from children ages three to five and their mothers. Mothers were instructed to speak at a normal rate in one condition and slow down their rate in a second condition. No significant difference in the children’s speech rate was reported between conditions. Most recently are the findings of Wynn, Borrie, and Sellers (2018), who used a quasi-conversational paradigm with pre-recorded video stimuli. Again, no significant difference
between slow and fast speech rate conditions was found for the child participants aged six to fourteen. However, using the same paradigm, significance between conditions was identified among typical adult participants. A further exploration into the methodologies used in the current literature of entrainment in children reveals a diversity of paradigms. Table 1 summarizes the basic methodologies and findings of studies with a specific focus on speech rate, highlighting the conflicting findings between teams with similar frameworks for collecting data. Contradictions in findings across the literature highlight the need for a methodology that is representative of conversational speech and yet replicable with reasonable experimental control.

Table 1. Current methodology relating to speech rate entrainment of children.

<table>
<thead>
<tr>
<th>Reference</th>
<th>Participants</th>
<th>Methodology</th>
<th>Identified Entrainment?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bernstein-Ratner (1992)</td>
<td>Ages: 3-5 n = 20</td>
<td>Natural play environment with mothers altering speech rate</td>
<td>No</td>
</tr>
<tr>
<td>Guitar &amp; Marchinkoski (2001)</td>
<td>Age: 3 n = 12</td>
<td>Natural play environment with mothers altering speech rate</td>
<td>Yes</td>
</tr>
<tr>
<td>Eaton &amp; Bernstein Ratner (2013)</td>
<td>Ages: 3-4 n = 40</td>
<td>Computer paradigm with priming sentence followed by novel production</td>
<td>No</td>
</tr>
<tr>
<td>Hupp &amp; Jungers (2009)</td>
<td>Ages: 4-5 n = 55</td>
<td>Computer paradigm with priming sentence followed by novel production</td>
<td>Yes</td>
</tr>
<tr>
<td>Wynn, Borrie, &amp; Sellers (2018)</td>
<td>Ages: 6-14 n = 15</td>
<td>Turn taking picture description task (quasi-conversational) with pre-recorded video stimuli</td>
<td>No</td>
</tr>
</tbody>
</table>
As discussed previously, entrainment is vital to successful conversation and social interaction. Consequently, a faulty or underdeveloped entrainment system can result in difficulties with building personal relationships. Borrie and Liss (2014) noted that “any impairment in the capacity to perceive, produce, and/or integrate rhythmic information has the potential to disrupt entrainment and negatively impact communication success.” Because rhythm deficits are present in many communication disorders, the potential for subsequent entrainment deficits in these populations is also likely. Indeed, there is a growing body of evidence suggesting impairments in entrainment abilities of both children and adults in several communication disorders including dysarthria (Borrie & Liss, 2014), hearing impairments (Freeman & Pisoni, 2017), fluency disorders (Sawyer, Matteson, Ou, & Nagase, 2017), and autism spectrum disorder (Wynn, Borrie, & Sellers, 2018).

Given the presence of entrainment impairments in such populations, research regarding entrainment is important from both a theoretical and clinical perspective. However, in order to understand the implications of entrainment deficits in children, we need a clear understanding of when and how these abilities are cultivated in typically developing children. A prerequisite to forming a theoretical framework regarding the development of entrainment is identifying a paradigm capable of measuring it in a reliable, controlled, and ecologically valid way. Although there are several paradigms currently employed to study entrainment in adults, these are often unsuitable for children. For example, the Montclair Map task, and Diapix task both require verbal problem solving between participants (Pardo et al., 2018; Van Engen et al., 2010), which are too difficult for younger children to adequately complete. Another methodology used with
adult populations requires participants to engage in scripted conversations by reading provided stimuli (Schultz et al., 2016). This approach is also not feasible for use with children, as those in early school age years are unable to read. Lastly, other current paradigms often involve a general conversational task between two participants (Sawyer et al., 2017). While this methodology is useful in gathering large amounts of data regarding the complex nature of conversation, these approaches prove difficult to isolate the abilities of a singular conversational partner. A controlled paradigm that allows the behaviors of a single participant to be analyzed in isolation serves to eliminate the confounding effects of other participants. The aims of the current study are to test a paradigm for capturing speech rate entrainment that can identify significant, replicable findings regarding the development of this integral ability.

METHOD

Overview

After obtaining informed consent, participants completed two key tasks: (i) a rhythm perception task, and (ii) an entrainment task. Additionally, participants completed brief standardized assessments for language and non-verbal intelligence.

Participants

Data were collected from a total of 50 typically developing children, comprising the following five age groups: 5-6 years, 7-8 years, 9-10 years, 11-12 years, and 13-14 years. Each group contained 10 participants. 28 participants were male, while 22 were
female. All participants were native speakers of American English per participant and parent report. Participants passed a hearing screening administered at 20 dB for 1000, 2000, and 4000 Hz. Participants’ language abilities were confirmed to be within normal limits (i.e. scaled score greater than or equal to 7) using the Following Directions and Recalling Sentences subtest of the Clinical Evaluation of Language Fundamentals Fifth Edition (CELF-5; Wiig, Semel, & Secord, 2013). Cognitive abilities were confirmed within normal limits (i.e. standard scores greater than or equal to 75) by using the Matrices subtest of the Kaufman Brief Intelligence Test 2nd Edition (KBIT-2; A. Kaufman, N. Kaufman, 2004).

Stimuli

Entrainment task. The entrainment task used was a minimally altered version of the entrainment task used by Wynn and colleagues (Wynn, Borrie, & Sellers, 2018). Stimuli consisted of 16 pictures taken from a children’s picture book. In pre-recorded video clips, a 22-year-old female speaker of American English introduces the picture, requests that participants describe what they see, and provides examples of what participants could talk about for each picture. Each recording is approximately 20-25s in length, with 18 recordings in total. Video clips were then digitally manipulated to create a version with a slower speech rate (80% of the original rate) and a version with a faster speech rate (120% of the original rate). Two clips were not altered so they could be used as training trials to ensure participants’ understanding of the task. Thus, 32 total recordings were produced as experimental stimuli, with 16 clips per each speech rate condition. The trials were embedded in a web-based application (hosted on a secure
university server). Participants completed all 32 trials in two sets of 16, one administered at the beginning of the experimental session, and the second completed at the end. Each set contained 8 trials of the fast and slow conditions respectively. The recordings were administered in a randomized order.

*Rhythm Perception.* The rhythm perception task contained stimuli previously used by Gordon and colleagues with children ranging in age from 5 yrs to 11 yrs (Gordon et al., 2015; Wieland, McAuley, Dilley, & Chang, 2015). The task was designed to measure rhythm perception abilities by having participants distinguish between similar and different rhythmic patterns. Stimuli were comprised of eight simple and eight complex rhythmic sequences, for a total of 16 rhythms. Participants completed 32 trials for this task, including same and different variants of all rhythm sequences. Rhythms range from 5-7 intervals long, and consisted of a pure tone “beep”. A rhythm was categorized as “simple” or “complex” based on the presence or absence of a strong beat.

**Procedure**

*Entrainment task.* During the entrainment task, participants were seated front of a computer screen in order to view and respond to the videos. The researcher explained to the participant that they would be watching a series of short video clips showing a person talking about some pictures. They were instructed to watch each video and then describe the picture shown in the clip. Participants were informed that they should continue talking for 15 seconds during the response period. A visual timer was displayed on screen to indicate the end of each trial, and to encourage participants to speak for the entire
duration of the response period. Each participant’s speech samples were audio-recorded using a headset with an attached microphone (Astro A50 Wireless System, San Francisco, CA).

The procedure began with two practice trials, using clips that were not digitally altered. Researchers coached or modeled to participants as necessary during the trial tasks until they demonstrated sufficient understanding of the task. Participants then continued with the experimental trials by viewing each stimuli video and providing a response. Each participant completed two sets of 16 trials, with randomized presentation of slow and fast conditions. Total time to complete this task was approximately 20 minutes.

*Rhythm perception task.* This procedure was integrated into a computer game format functioning on E-Prime 3 software (Psychology Software Tools, Inc.). Participants were seated at a laptop computer and wore sound-attenuating headphones throughout the task (Sennheiser HD 650 PRO, Old Lyme, CT). During the task, participants viewed a character named “Randy Drummer” who played each original rhythm two times. Then, either the same rhythm sequence or a different rhythm was played. Participants were asked to determine if the subsequent rhythm was the same or different as the original by indicating if the new rhythm was played by the “Sandy Same” character or the “Doggy Different” character. Choices were indicated by pressing the corresponding button on the computer. See Gordon and colleagues (2015) for a visual summary of each trial.

Participants began by completing four practice trials comprised of both a same and a different variant of one simple and one complex rhythm. The researcher coached each participant as needed during the practice trials to ensure comprehension of the task.
During the trials, participants received feedback from the computer program indicating if their selections were correct or incorrect. After the practice trials, participants completed 32 experimental trials containing the same and different variants of all 16 rhythms. The procedure was counterbalanced by randomizing the order of appearance for each rhythm. This task required approximately 20 minutes to complete.

**Data Analysis**

The total data set for the entrainment task consisted of 1600 audio response recordings—800 response recordings for the slow condition and 800 response recordings for the fast condition. Research assistants used acoustic analysis software, Praat (Boersma & Weenink, 2015), to calculate speech rate (syllables per second) for each response recording. Research assistants transcribed each response recording, and counted the number of syllables for each production. Involuntary or non-speech sounds (e.g., hiccups, laughing, coughing) were not included in the transcription or speech rate calculation. All other verbal output, including whole word repetitions, part word repetitions, and filler words (e.g., uh, um) were included in the speech rate calculation as the purpose of this task was to identify speech rate production, and not analyze the content of participants’ output. Approximately 25% of the total entrainment task data (416 response productions) were coded by a trained research assistant to obtain interrater reliability for speech rate calculations. Data sets were chosen for this purpose by using a random number generator. Comparison indicated high agreement between the two judges, with a Pearson correlation r score of .97. Data for the rhythm perception task were recorded using the E-Prime software, resulting in a raw score for each participant.
Maximum performance on the rhythm perception task constituted a raw score of 32 (one point for each experimental trial).

**RESULTS**

Linear mixed models were used to investigate the effects of condition and age on average speech rate while controlling for the lack of independence in the data due to the repeated measures. Average speech rate in syllables per second in response to slow and fast stimuli conditions was recorded for each participant. Figure 1 illustrates these findings. Within-participant factor was Condition (slow stimuli vs fast stimuli), and the between participant factor was Age. Analysis of the results revealed a significant main effect of age (see Figure 1). Thus, as age increased, speech rate also increased. There was no significant main effect of condition and no significant interaction between condition and age (see Table 2), indicating that regardless of age, children did not modulate their speech rate to match that of the virtual interlocutor.

Rhythm scores were also computed for each participant (M = 25.58; SD = 4.81). Because entrainment was not detected in any age group, an analysis between entrainment and rhythm perception was not conducted. Analysis between rhythm perception score and age revealed a moderate positive correlation, \( r(48) = .48, p < .001 \), meaning as age increased, rhythm perception score also increased. This supports the theory that as age increases, rhythmic entrainment ability will improve, as rhythm perception is hypothesized to be a prerequisite of entrainment capability (Phillips-Silver et al., 2010).
Figure 1. *Relationship between speech rate during slow and fast conditions by age group.*

Table 2. *Linear mixed models results relating to condition and age*

<table>
<thead>
<tr>
<th>Effects</th>
<th>t-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Condition</td>
<td>-0.669</td>
<td>.504</td>
</tr>
<tr>
<td>Age</td>
<td>6.644</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Interaction (Condition x Age)</td>
<td>0.196</td>
<td>.845</td>
</tr>
</tbody>
</table>
DISCUSSION

The present study examined the use of a quasi-conversational paradigm for identifying speech rate entrainment in typical child populations. The results revealed that children did not modify their speech to align to the speech rate of the virtual interlocuter within the given paradigm. This finding, or lack thereof, in children populations is consistent with the findings of research teams listed in Table 1 (Bernstein-Ratner, 1992; Torrington Eaton & Bernstein Ratner, 2013; Wynn et al., 2018). Comparisons of these findings with other literature containing similar paradigms but statistically significant measures of speech rate entrainment (see Table 1) suggests that the lack of a significant finding reflects the design features of the current experimental paradigm, rather than a complete absence of speech rate entrainment within this population.

The current paradigm offered many advantages from a methodological standpoint. First, the use of a computer program ensured for standardization across all participants. Digital manipulation of the video clips allowed for all participants to experience the same stimuli, and thus the same experimental rates for the slow and fast conditions, respectively. Unlike paradigms which required human experimenters to adjust their rate in real time (e.g. Guitar & Marchinkoski, 2001), results were not influenced by variability in performance of confederate speakers. Additionally, a paradigm employing quasi-conversation with a virtual interlocutor eliminates confounding variables that may arise between two in vivo interlocutors engaged in conversation. Furthermore, this high level of control is balanced by the turn-taking between participant and stimuli, while the use of spontaneous productions maintains some degree of naturalness and ecological validity.
However, the lack of speech rate entrainment observed in the current study suggests that this paradigm may not be sufficiently sensitive to capture entrainment in children populations. While preserving the benefits of the current design, there are multiple potential modifications that could be implemented to potentially eliminate the sensitivity deficits of this paradigm. As theoretical rationale exists for employing these modifications, it is unclear which individual change or combination of changes will provide the most useful scenario for identifying entrainment in children. Below, ideas for modification of the current paradigm, as well as the theoretical underpinnings, are discussed.

First, an adjustment to the nature of the task may elicit more salient findings. Our current task involved a picture description following prerecorded questions and comments. An adaptation to this feature may include providing questions and comments about neutral topics (e.g. For my last birthday I went to the park. What did you do for your last birthday?). By removing the picture description component, participants would perhaps be inclined to actively listen to the stimuli and provide an answer based on the given topic, rather than arbitrarily describing visual scenes. While this topic of conversation might be as equally mundane as a picture description task, by answering personally relevant questions participants may be more likely to produce authentic answers, and consequently enhance the overall authenticity of the interaction itself. This technique is similar to those used to gather language samples, such as the Hadley protocol (Hadley, 1998). This method of introducing a topic has been shown to be effective in eliciting high level, unplanned discourse.
The next modifiable component involves the length of the task and individual turns. As the length of the entrainment task was 30 minutes (two sets of approximately 15 minutes each), participants may have experienced fatigue or become disengaged over the course of all 32 entrainment trials. Additionally, the turn length of 15 seconds could also be shortened to allow the conversational simulation to rapidly alternate between interlocutors. One example is Hupp and Jungers (2009), who also used prerecorded stimuli in a computer format. However, participant responses were limited to a single sentence repetition following the presentation of that same sentence in either a slow or fast condition. This research team identified statistically significant differences between conditions. Additionally, Oviatt and colleagues (2004) also reported significant findings by using a paradigm which provided stimuli turns ranging from 1.27 to 2.01 seconds in length. While intended to encourage response production, the timer used in the current paradigm may have added pressure for participants to provide as much commentary as possible, thus altering the anticipated naturalness of a response. Current literature regarding conversational turn taking reports that individuals formulate their responses during the turn of their partner, lending to typical response latencies in adults lasting an average of 236 milliseconds (Levinson, 2016; Stivers et al., 2009). By shortening the turn length of stimuli and responses, participants will have the opportunity to engage in this rapid exchange of turns, creating the need for partners to perceptually engage with the task.

The presentation of stimuli trials also provides an opportunity for modification. The current paradigm utilized a random presentation, which required participants to listen to a sequence of trials that shifted unexpectedly between slow and fast speech rates. An
alternative method of presentation would employ a format in which participants viewed
and responded to blocks of consecutive fast clips followed by consecutive slow clips.
This type of paradigm is supported by theories of the mechanism of entrainment.
Pickering and Garrod (2004) posit that entrainment of speech features in dialogue is
dependent on a speaker’s ability to perceive, produce, and adjust rhythmic signals based
on input received from their conversational partner. The current random presentation of
stimuli diminishes this ability by rapidly and abruptly moving from slow to fast rates,
potentially disrupting the perceptual learning process. This is further supported by the
work of Oviatt and colleagues (2004), in which participants were sometimes required to
quickly alternate between animated conversational partners. Findings indicated that
participants were able to entrain their amplitude and pause duration regardless of rapid
transitions, yet this was not similarly indicated in the analysis of speech rate. However,
participants did demonstrate entrainment of speech rate when conversing with a single
animated character. This suggests that the feature of speech rate requires adequate time
for children to perceive and integrate into their own speech patterns. Thus, by segmenting
the presentation of stimuli by speed, participants would have the opportunity to
perceptually adapt to the given speech rate, similar to the experience of interacting with
different conversational partners with varied habitual rates.

Remaining modifiable components include the acoustic properties of the stimuli
itself. In this study, the video samples captured the speech of a 24-year-old female
speaking American English. Alternative versions of stimuli could include male speakers,
or younger speakers. There is evidence to suggest that gender not only influences the
acoustic features produced by the speaker, but also the perceptions of the listener
Specifically, the gender of interlocutors influences the listener’s perception of the speaker’s speech rate. While the gender of participants was relatively balanced in the current study (28 males, 22 females), it is possible that greater entrainment effects could be identified by including some stimuli clips of a male speaker. Further exploration is needed to identify the extent of the effects of a particular conversational partner’s attributes on the presence of conversational entrainment in children.

In sum, a number of ideas for modifying the current paradigm exist, with the goal of establishing a methodology that can identify typical and disordered conversational entrainment in a controlled setting. Future studies in this area will also make use of adult populations to validate the experimental methodology. There is ample evidence of speech rate entrainment in healthy adults, thus an absence of entrainment within adult populations would suggest a lack of sensitivity in a given paradigm. Further, once a validated experimental paradigm for capturing speech rate entrainment is established, examining individual variables associated with speech entrainment can be explored, including a developmental timeline for this ability. Other than age, potential factors to a typical developmental process include language ability, social aptitude, musical experience (relevant to perception and production of rhythmic signals), and other general traits such as temperament. Identifying the precipitating factors for typical development of entrainment will increase our understanding of the underlying mechanisms of this ability. In this way, researchers will also gather insight into the detriments and potential remedies for difficulties in entrainment caused by communicative disorders. This crucial
research requires a sound methodological foundation in order to provide theoretically and clinically meaningful information.
REFERENCES


https://doi.org/10.1177/0956797610392928


