

Research Article

Going With the Flow: An Examination of Entrainment in Typically Developing Children

Camille J. Wynn,^a Stephanie A. Borrie,^a and Kiersten A. Pope^a

Purpose: Conversational entrainment is the tendency for individuals to modify their behavior to more closely converge with the behavior of their communication partner and is an important aspect of successful interaction. Evidence of entrainment in adults is robust, yet research regarding its development in children is sparse. Here, we investigate the emergence of entrainment skills in typically developing children. **Method:** Data were collected from a total of 50 typically developing children between the ages of 5 and 14 years. Children participated in a quasiconversational paradigm with a virtual interlocutor. Speech rate of the interlocutor was digitally manipulated to produce fast and slow speech rate conditions.

Results: Data from the fast and slow conditions were compared using linear mixed models. Results indicated that children, regardless of age, did not alter their speech to match the rate of the virtual interlocutor.

Conclusions: Findings suggest that entrainment in children may not be as robust as entrainment in adults and therefore not adequately captured with the current experimental paradigm. Modifications to the current paradigm will help identify a methodology sufficiently sensitive to capture the speech alignment phenomenon in children and provide much needed information regarding the typical stages of entrainment development.

From a very early age, children begin to engage in conversation with others. The success of these conversations is predicated on the development of a wide array of speech, language, and pragmatic skills. One area relating to these interactions that has received little attention regards the development of conversational entrainment. Conversational entrainment, also known as *alignment* (Pickering & Garrod, 2004), *accommodation* (Street, Street, & Van Kleek, 1983), and *convergence* (Natale, 1975), is the tendency for individuals to align their behaviors with one another during communicative interactions. That is, over the course of a conversation, individuals will adjust their communicative behaviors in order to become more like the behaviors of their conversation partner. These adaptations are pervasive and occur at an unconscious level, suggesting that this communication phenomenon is a fundamental part of human interaction (Chartrand & Bargh, 1999).

Within typical adult populations, entrainment of speech behaviors is well studied. Adults have been observed to entrain on a variety of acoustic-prosodic features including intensity, pitch, speaking rate, and vocal quality (e.g., Borrie & Delfino, 2017; Borrie & Liss, 2014; Gregory, 1990; Levitan & Hirschberg, 2011). These findings have been observed in several different languages and geographic locations (e.g., Freud, Ezrati-Vinacour, & Amir, 2018; Levitan, Beňuš, Gravano, & Hirschberg, 2015; Xia, Levitan, & Hirschberg, 2014). Additionally, entrainment has been studied in many different speaking contexts, from tightly controlled laboratory settings in which participants respond to prerecorded stimuli (e.g., Goldinger, 1998; Jungers & Hupp, 2009; Pardo, Urmanche, Wilman, & Wiener, 2017) to naturalistic conversations between two people with little environmental modification (e.g., Coupland, 1984; Natale, 1975). Variations in the degree to which individuals entrain to the speech patterns of their conversational partners have been discussed in relation to gender (Levitan et al., 2012; Pardo et al., 2017), participant role (Pardo, 2006), and social preference (Babel, McGuire, Walters, & Nicholls, 2014).

Conversational entrainment has been reported to play a beneficial role in communicative interactions. For example, entrainment correlates with enhanced comprehension of spoken language (Pickering & Garrod, 2004), greater conversational fluidity (Levitan et al., 2012; Wilson & Wilson,

^aDepartment of Communicative Disorders and Deaf Education, Utah State University, Logan

Correspondence to Camille J. Wynn:
camille.wynn@aggiemail.usu.edu

Editor-in-Chief: Bharath Chandrasekaran
Editor: Julie Anderson

Received March 13, 2019

Revision received April 26, 2019

Accepted July 5, 2019

https://doi.org/10.1044/2019_JSLHR-S-19-0116

Disclosure: The authors have declared that no competing interests existed at the time of publication.

2005), and reduced interruptions between dialogue partners (Local, 2007). Furthermore, entrainment predicts increased efficiency on goal-directed dialogue tasks (Borrie, Barrett, Willi, & Berisha, 2019) and greater cooperation between communication partners (Manson, Bryant, Gervais, & Kline, 2013). Individuals who exhibit proficiency in speech entrainment are perceived as being more competent, friendly, and self-confident (Putman & Street, 1984; Schweitzer, Lewandowski, & Duran, 2017). It is therefore not surprising that individuals who entrain well in conversation report stronger relationships and rapport than those who do not (Lee et al., 2010; Pardo, Gibbons, Suppes, & Krauss, 2012).

Although there is much research investigating conversational entrainment in adults, little is known about the phenomenon in pediatric populations. The small body of research that has explored entrainment in children has largely targeted speaking rate. However, despite the literature's focus on this single acoustic feature, general conclusions are scarce. Studies have differed in regard to nontrivial aspects of methodology, and even when methodologies are comparable, findings are equivocal. For example, a key question is the age at which children develop the ability to entrain to the speech rate of others. Using an experimental paradigm in which children responded to spoken stimuli in slow versus fast speech rate conditions, Eaton and Ratner (2013) observed speech rate entrainment in children between 3 and 4 years old. However, using an almost identical paradigm, Hupp and Jungers (2009) found no evidence of speech rate entrainment in children who were 4 years of age. When a more naturalistic methodology was employed in which mothers consciously modulated their speech rate, Guitar and Marchinkoski (2001) showed evidence of entrainment in children 3 years of age, whereas Ratner (1992) found no evidence of entrainment in children of that same age. These inconclusive findings across studies are not limited to young children. Similar discrepancies have been found in studies involving school-aged children and adolescents (Oviatt, Darves, & Coulston, 2004; Wynn, Borrie, & Sellers, 2018).

In addition to differences in findings, other factors make it difficult to generate solid conclusions about the age at which entrainment skills are developed. First, most existing studies have investigated entrainment within a relatively narrow age range. For example, a number of studies have examined entrainment skills in children within an 18-month age range (e.g., Eaton & Ratner, 2013; Ko, Seidl, Cristia, Reimchen, & Soderstrom, 2016; Street et al., 1983). Even those with larger age ranges were generally limited to ranges of less than 3 years (e.g., Jungers & Hupp, 2009; Oviatt et al., 2004; Ratner, 1992). Additionally, most existing studies have employed a small number of participants. Guitar and Marchinkoski (2001), for example, used a sample size of six, and Street et al. (1983) used a sample size of four. Small sample sizes not only make it difficult to reliably generalize findings to larger populations but also make it problematic to compare performance across age groups. This can be highlighted by a previous study carried out by Wynn et al. (2018). In this study, the entrainment patterns of children between ages 6 and 14 years were investigated

using a sample of 30 participants. While collective results did not show overall entrainment patterns within the sample, an interaction effect between age and condition was significant. These findings suggest that entrainment abilities were present in older children, but the sample size was too small to stratify results and further compare differences in findings by age group.

Given the limited research in this area and disparate findings among the few existing studies, additional studies regarding speech rate entrainment in children are warranted. Such research is important for many reasons. As discussed, entrainment plays an important role in successful conversation and is, therefore, likely an important part of a child's social development. Thus, disruptions in this phenomenon may yield significant consequences, affecting a child's ability to engage in meaningful conversation and communicate effectively with their peers. Successful speech entrainment requires individuals to identify the speech behaviors of their communication partner and then modify their own speech to more closely align with those behaviors. Therefore, populations with deficits in the perception and/or the production of speech may experience entrainment impairments (Borrie & Liss, 2014). Indeed, there is a small, but growing body of evidence demonstrating entrainment deficits in individuals with various communication disorders including autism spectrum disorder (Wynn et al., 2018), dysarthria (Borrie, Lubold, & Pon-Barry, 2015), hearing impairments (Freeman & Pisoni, 2017), and fluency disorders (Sawyer, Matteson, Ou, & Nagase, 2017).

While research regarding speech entrainment in children with communication disorders is important, the current lack of knowledge regarding this communication phenomenon in typical pediatric populations makes more extensive research in clinical populations somewhat premature. To fully appreciate the nature of entrainment deficits, conclusions regarding when the skill emerges in typical development is required. Here, we investigate speech rate entrainment in a large sample of 48 typically developing children, who span the ages of 5–14 years. We employ a controlled quasi-conversational paradigm to address the following key research question: At what age do typically developing children entrain their speech rate? Research has indicated that the amount of time children engage in conversation dramatically increases during pre-adolescence (Rafaelli & Duckett, 1989). Additionally, during this time, individuals begin to develop more intimate and empathic relationships (Samter, 2003). Therefore, we hypothesize that speech entrainment will emerge during this stage of development (i.e., between 11 and 13 years). This work will set the stage for further investigations into the developmental nature of speech entrainment in both typical and clinical populations.

Method

Participants

Participants included 48 (27 males, 21 females) typically developing children between the ages of 5 and 14 years

($M = 10.0$, $SD = 2.7$). Data from two additional participants were removed because of language scores outside of normal limits. In order to ensure an even spread, the sample was stratified with 9–10 participants recruited for each of the following age ranges: 5–6 years, 7–8 years, 9–10 years, 11–12 years, and 13–14 years. All participants were native speakers of American English per participant and parent report. Prior to the experimental task, all participants passed a hearing screening administered at 20 dB for 1000, 2000, and 4000 Hz. Participants' language abilities were confirmed to be within typical limits (i.e., scaled score ≥ 7) using the Following Directions ($M = 10.5$, $SD = 1.9$) and Recalling Sentences ($M = 11.2$, $SD = 2.1$) subtest of the Clinical Evaluation of Language Fundamentals–Fifth Edition (Wiig, Semel, & Secord, 2013). Nonverbal cognitive abilities were confirmed within typical limits (i.e., standard scores greater than or equal to 85) by the Matrices subtest ($M = 107.8$, $SD = 18.62$) of the Kaufman Brief Intelligence Test–Second Edition (Kaufman & Kaufman, 2004).

Stimuli

The entrainment task used was a minimally altered version of the entrainment task used by Wynn et al. (2018). Audiovisual stimuli were created in a sound-attenuated booth with an industry standard microphone (Shure SM58) and video camera (Canon EOS 70D).¹ The speaker was positioned against a neutral backdrop with the camera positioned to capture a view of her head and shoulders. Recordings were captured digitally on a memory card at 48 kHz (16-bit sampling rate) and stored as individual recording files.

In each recording, a 22-year-old female speaker of American English holds a picture from a popular children's book near her face. She introduces the picture, requests that participants describe what they see, and provides examples of what participants could talk about for each picture (see the Appendix for a sample transcript). Each recording is approximately 20–25 s in length. Recordings were digitally manipulated to create a slow version (80% of the original rate) and a fast version (120% of the original rate) of each video clip.² In total, 32 total recordings were produced as experimental stimuli, with 16 clips containing fast speech rate and 16 clips containing slow speech rate. The trials were embedded in a web-based application, hosted on a secure university server.

¹The decision to use audiovisual rather than audio-only stimuli was to increase the naturalness of the interaction task and an effort to keep children engaged during the entire duration of the task.

²In order to provide a natural appearance, manipulations to video clips were kept as minimal as possible while still providing sufficient variability in the speech rate of the speaker. Prior to the experiment, all research team members viewed the video clips and concluded they still retained a natural appearance. However, we acknowledge that digital manipulations performed over the entire utterance may not precisely replicate natural changes in speech rate.

Procedure

The experiment was conducted in a quiet room. Participants were told that they would be participating in a study examining speech patterns of children. However, no other explanation about the purpose of the study was given at that time. After obtaining informed consent, participants were seated in front of a computer screen in order to view and respond to the audiovisual stimuli. The researcher explained that the participant would be watching a series of short video clips showing a person talking about some pictures and that immediately following each clip, the picture described in the video would appear on the screen. The participants 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 s 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).

The procedure began with two practice trials, using clips with the women's normal speaking rate. Researchers coached the 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 clip and providing a verbal response. In total, participants completed 32 experimental trials. Trials were divided into two experimental sets containing 16 trials each (i.e., eight trials of both the fast and slow conditions). Between experimental sets, participants completed additional language and cognitive testing. Although all participants viewed the video clips in the same order, the speed of the recordings were presented in a random order. Thus, the presentation of individual fast versus slow stimuli videos was different for each participant. Total time to complete each experimental set was approximately 20 min. All 48 participants completed the task in its entirety. Additionally, all participants followed directions, and no participant required additional instructions or prompting to refocus during any portion of the experiment.

Data Analysis

The total data set for the entrainment task consisted of 1,536 audio response recordings—768 response recordings for the slow condition and 768 response recordings for the fast condition. A trained research assistant used acoustic analysis software, Praat (Boersma & Weenink, 2018), to calculate speech rate (syllables per second) for each response recording. The research assistants orthographically transcribed each response recording and counted the number of syllables for each production. They then measured the entire duration of the response recording, beginning with the moment the child began articulating their response and ending when articulation of the child's response (within

the 15-s time frame) ceased.³ Speech rate for each response recording for each participant was then calculated by dividing the total number of syllables by the duration measure. As per data analysis procedures of Wynn et al. (2018), involuntary or nonspeech sounds (e.g., hiccups, laughing, coughing) were not included as syllables in the speech rate calculation, and the length of the nonspeech sound was removed from the duration measure. As the purpose of this task was to identify speech rate production and not analyze the content of participants' output, all other verbal outputs, including whole word repetitions, part word repetitions, and filler words (e.g., uh, um), were included as syllables in the speech rate calculation. Approximately 25% of the total entrainment task data (13 participants' data sets) was randomly selected by a computer-generated random number list and re-analyzed by a different research assistant to obtain interrater reliability for speech rate calculations. Comparison indicated high agreement between the two judges, with a Pearson correlation *r* score of .97.

Results

Average speech rate in syllables per second in response to slow and fast stimuli conditions was recorded for each participant. Linear mixed models were conducted using the lme4 package in the R statistical environment (lme4 package Version 1.1-19 and R Version 3.5.2; Bates, Machler, Bolker, & Walker, 2015; R Core Team, 2018). This type of analysis was 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. For the models, the random effects structure included a random intercept by participant. The fixed effects included the within-participant factor of condition (i.e., slow stimuli vs. fast stimuli) and the between-participant factor of age. In order to account for potential confounding factors, two additional variables were included within the models. First, experimental set (i.e., first or second set of experimental trials) was included as a fixed effect in order to account for a possible decline in entrainment over time due to fatigue or task disengagement. Second, gender was included as fixed effects to account for potential variability between male and female participants. Thus, the specific formula for the first model (Model 1) was: $\text{lmer}(\text{rate} \sim \text{condition} * \text{age} + \text{condition} * \text{set} + \text{condition} * \text{gender} + (1 | \text{participant}))$. Because there were no significant interactions within this model (see below), interaction terms were removed in order to assess main effects. Thus, the specific formula for the second model (Model 2) was $\text{lmer}(\text{rate} \sim \text{condition} + \text{age} + \text{set} + \text{gender} + (1 | \text{participant}))$. Therefore, results for interaction effects are reported from Model 1 and results for main effects are reported from Model 2.

³Because our analysis focused on speech rate rather than articulation rate, pauses during responses were not removed from the duration measure. However, it is important to note that responses were generally free from long internal pauses, and no obvious variability in utterance-internal pauses across age groups was noted.

Analysis of the results revealed a significant main effect of age ($b = .155, p < .001$), but no significant main effect of condition ($b = .043, p < .068$). Additionally, there was no significant interaction between condition and age (see Table 1). Thus, children's speech rate increased with age; however, regardless of age, children did not modulate their speech rate depending on the speech rate of the virtual interlocutor. There was also a significant main effect for experimental set ($b = .197, p < .001$), indicating that children spoke more quickly in response to the second set of video clips than in response to the first set of video clips. There was, however, no interaction between condition and experimental set. Additionally, there were no significant main effects for gender ($b = .055, p < .673$) and no interaction between gender and condition.

Discussion

This study examined whether children aligned their speech rate to the speech rate of a virtual interlocutor in a quasiconversational experimental paradigm. Our results indicated a main effect for age—older children employed faster speech rates than younger children. This finding is in line with previous research showing an increase in children's speech rate with age (Haselager, Slis, & Rietveld, 1991; Nip & Green, 2013). Results also indicate a main effect for set—children spoke faster in the second set than in the first set of trials. It is likely that the slight increase (i.e., an average increase of .2 syllables per second) from the first to the second set of experimental trials was a result of greater familiarity with the task. Our primary analysis regarding speech rate entrainment revealed nonsignificant results. That is, there were no differences in participant's speech rate in fast and slow conditions, suggesting that children, regardless of age, did not entrain their speech. Given the relatively large sample size, it is unlikely that the insignificant findings can be explained by lack of data. Rather, we deduce two possible interpretations. First, it is possible that by the age of 14 years, children have not yet developed the ability to entrain to the speech rates of others. Second, the current paradigm may not be sufficiently sensitive to capture conversational entrainment in children. We conjecture that these two interpretations are not mutually exclusive; rather, the explanation of the current findings likely hinges on a combination of both.

Table 1. Results of the fixed effects for Model 1.

Term	Estimate	SE	<i>t</i> value	<i>p</i> value
Intercept	1.344	0.249	5.402	< .001
Condition	0.046	0.094	0.485	.628
Age	0.153	0.025	6.088	< .001
Set	0.186	0.034	5.421	< .001
Gender	0.041	0.136	0.301	.765
Condition × Age	0.003	0.009	0.366	.714
Condition × Set	0.020	0.049	0.421	.674
Condition × Gender	0.031	0.050	0.617	.538

Although little is known about the developmental trajectory of conversational entrainment, it is unlikely that the skill is acquired instantaneously. Rather, like other aspects of speech and social development, entrainment development may be viewed as a multistep process, in which skills emerge over time and, through continued practice, become increasingly refined and solidified. When considered from this vantage point, our findings do not necessarily indicate that speech rate entrainment is completely absent in children's communicative interactions. Rather, the findings may suggest that entrainment abilities are not sufficiently robust to be successfully deployed in the current experimental paradigm. Thus, further study of this phenomenon in children requires adjustments to the current paradigm. Systematically, investigating the effects of paradigm modification will not only help identify a methodology sensitive to emerging entrainment skills, but will provide opportunity to learn about the stages of typical entrainment development and the fundamental milestones that precede maturation of this skill. Continuing with a quasiconversational paradigm allows for adequate experimental control while offering a simulated conversation requisite for meaningful findings. Entrainment requires an individual to perceive the rhythmic cues of their communication partner and integrate them into their own speech (Phillips-Silver, Aktipis, & Bryant, 2010; Todd, Lee, & O'Boyle, 2002). Therefore, paradigm modifications that support these perception and production behaviors will likely be advantageous. We offer several ideas below.

First, in this study, stimuli clips were presented in a randomized order so that fast and slow recordings were interspersed with one another. Therefore, entrainment within the paradigm required participants to make frequent adjustments to both their perceptions of the target speech rate and the rate or their own spoken productions. Although this type of presentation has been shown to induce speech entrainment in adults (e.g., Borrie & Liss, 2014), the rapid alteration between speech rate conditions may have been too challenging for children. Because both rhythm perception and production skills are still maturing through late childhood and early adolescence (Persellin, 1992; Uptis, 1987), this group may require longer exposure to the speech of their conversation partner to learn the rhythmic patterns and integrate them into their own speech. Consistent with this speculation, Oviatt et al. (2004) found that children modulated their speech rate to match that of a virtual interlocutor within long-duration dialogues. However, within short-term dialogues, no adjustments in speech rate were noted. Similarly, Hupp and Jungers (2009) found that, while differences between the fast and slow speech rate conditions were not detected in the first block of trials within their study, differences were found in the second block, implicating entrainment as a function of time. Thus, adjusting the current paradigm so that fast and slow conditions are presented in separate blocks may afford the support needed to capture less refined entrainment skills.

Another important consideration regards the type of stimuli presented. In an effort to maintain engagement

within the current paradigm, children watched audiovisual clips of the speaker rather than simply listening to audio recordings. However, the use of visual stimuli may have actually hindered the children's ability to attend to the acoustic information. In a recent study, Schweitzer, Walsh, and Schweitzer (2017) found that adults aligned their pitch accents with one another more in an audio-only than an audiovisual condition. Similarly, Savino, Lapertosa, and Refice (2018) found that, while entrainment varied by prosodic feature and by individual, overall entrainment occurred more frequently in an auditory-only condition relative to an audiovisual condition. One possible reason for these findings is that, during multisensory tasks, attention is divided and individuals are forced to distribute cognitive resources to both modalities (Loose, Kaufmann, Auer, & Lange, 2003; Vohn et al., 2007). Research has demonstrated that children are better able to maintain attention in low-load visual environments (e.g., environments with plain white walls) than in high-load visual environments (e.g., environments with brightly colored pictures on walls; Rodrigues & Pandeirada, 2018; Stern-Ellran, Zilcha-Mano, Sebba, & Levit Binnun, 2016). Consequently, what is moderately difficult but manageable for adults may actually be an extremely difficult task for younger interlocutors. Thus, in the case of verbal entrainment, in which successful entrainment relies heavily on auditory perception, visual stimuli may have served as a critical distractor. It may be beneficial, therefore, to first establish when entrainment develops in children solely with auditory information, and use those findings to examine when the more advanced skills of audiovisual entrainment arise.

The length of the paradigm is also an important element to consider. In the current paradigm, children were asked to engage in the entrainment task for 40 min (i.e., two sets of 16 pictures lasting approximately 20 min each). Additionally, language and cognitive testing was performed between the two sets of pictures, further lengthening the time of the experiment. There is strong evidence that children are able to sustain their attention on a given task for significantly less time than adults (Laurie-Rose, Bennett-Murphy, Curtindale, Granger, & Walker, 2005; Rebok et al., 1997; Zhan et al., 2011). This may be particularly important to consider in the realm of conversation as it is not until late adolescence that individuals shift from activity-centered activities to more conversation-based interactions (McNelles & Connolly, 1999). Therefore, participating in conversational tasks for 40 min may have been too long for children, leading to disengagement, and decreased attention. As such, reducing task length may increase the likelihood that children will attend to the stimuli throughout the duration of the study, consequently, providing more opportunity to perceive and integrate the speech rhythm patterns of their conversational partner. Our analysis did not reveal any difference between entrainment within the first and second set of pictures, suggesting that a loss of attention over time does not exclusively account for the lack of significant results within the current study. However, this does not preclude the idea that attention or the lack thereof,

coupled with other factors, may influence findings to some degree.

Several additional modifications may also increase sensitivity of the entrainment paradigm. In the current paradigm, individual conversational turns lasted 15–25 s. Decreasing the turn length of both the exposure and response productions affords more immediate feedback, which may aid entrainment. This idea is supported by studies, which found that children modified their speech rate when exposure and response productions were limited to a few seconds (Eaton & Ratner, 2013; Oviatt et al., 2004). Additionally, varying individual characteristics of the virtual interlocutor could also be efficacious. For example, as children show many differences in interactions between peers and adults and between same-sex peers and opposite-sex peers (Larson & McKinley, 1998; Turkstra, 2001), changing the age and/or gender of the stimuli may lead to differences in findings. Furthermore, modifying the nature of the conversational task may also be efficacious. For example, questions about participants and their lives (e.g., *What's your favorite thing to do when you are at home?*; Hadley, 1998) may help the child maintain focus throughout the duration of the experiment. This may be particularly important for older children, who may be disengaged in a simple picture description task. Additionally, once the developmental nature of entrainment is established in well-controlled settings, studies involving embodied interactions between in-person participants should ensue.

In sum, the overall aim of this article was to examine conversational entrainment in children between 5 and 14 years using a well-controlled experimental paradigm. Our results indicated no entrainment among the age range employed within in the study. Our findings, taken with previous research, may suggest that entrainment skills develop progressively throughout childhood, and that the current paradigm was not sufficiently sensitive to capture emerging skills. Therefore, we offer ideas regarding methodological shifts for continued research in this area.

Acknowledgments

This research was supported by National Institute on Deafness and Other Communication Disorders Grant R21DC016084, awarded to Stephanie Borrie. Data collection and analysis for this project was led by Kiersten Pope as part of her master's thesis in the Human Interaction Lab (Borrie) at Utah State University. The data included in this report was presented at the American Speech-Language-Hearing Association Convention, Boston, Massachusetts, in November 2018. We gratefully acknowledge Tyson Barrett for statistical input and research assistants in the Human Interaction Lab for assistance with data collection and analysis.

References

Babel, M., McGuire, G., Walters, S., & Nicholls, A. (2014). Novelty and social preference in phonetic accommodation. *Laboratory Phonology, 5*, 123–150.

- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48.
- Boersma, P., & Weenink, D. (2018). Praat: Doing phonetics by computer (Version 6.0) [Computer software]. Retrieved from <http://www.praat.org>
- Borrie, S. A., Barrett, T. S., Willi, M. M., & Berisha, V. (2019). Synching 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*, 283–296.
- Borrie, S. A., & Delfino, C. R. (2017). Conversational entrainment of vocal fry in young adult female American English speakers. *Journal of Voice, 31*(4), 513.e25–513.e32.
- Borrie, S. A., & Liss, J. M. (2014). Rhythm as a coordinating device: Entrainment with disordered speech. *Journal of Speech, Language, and Hearing Research, 57*(3), 815–824.
- 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.
- Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception-behavior link and social interaction. *Journal of Personality and Social Psychology, 76*, 893–910.
- Coupland, N. (1984). Accommodation at work: Some phonological data and their implications. *International Journal of the Sociology of Language, 1984*(46), 49–70.
- Eaton, C. T., & Ratner, N. B. (2013). Rate and phonological variation in preschool children: Effects of modeling and directed influence. *Journal of Speech, Language, and Hearing Research, 56*(6), 1751–1763.
- Freeman, V., & Pisoni, D. B. (2017). Speech rate, rate-matching, and intelligibility in early-implanted cochlear implant users. *The Journal of the Acoustical Society of America, 142*(2), 1043–1054.
- Freud, D., Ezrati-Vinacour, R., & Amir, O. (2018). Speech rate adjustment of adults during conversation. *Journal of Fluency Disorders, 57*, 1–10.
- Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review, 105*, 251–279.
- Gregory, S. W. (1990). Analysis of fundamental frequency reveals covariation in interview partners' speech. *Journal of Nonverbal Behavior, 14*, 237–251.
- Guitar, B., & Marchinkoski, L. (2001). Influence of mothers' slower speech on their children's speech rate. *Journal of Speech, Language, and Hearing Research, 44*(4), 853–861.
- Hadley, P. A. (1998). Language sampling protocols for eliciting text-level discourse. *Language, Speech, and Hearing Services in Schools, 29*(3), 132–147.
- Haselager, G. J. T., Slis, I. H., & Rietveld, A. C. M. (1991). An alternative method of study in the development of speech rate. *Clinical Linguistics & Phonetics, 5*, 53–63.
- Hupp, J. M., & Jungers, M. K. (2009). Speech priming: An examination of rate and syntactic persistence in preschoolers. *British Journal of Developmental Psychology, 27*(2), 495–504.
- Jungers, M. K., & Hupp, J. M. (2009). Speech priming: Evidence for rate persistence in unscripted speech. *Language and Cognitive Processes, 24*(4), 611–624.
- Kaufman, A. S., & Kaufman, N. L. (2004). *Kaufman Brief Intelligence Test—Second Edition (KBIT-2)*. Bloomington, MN: Pearson.
- Ko, E.-S., Seidl, A., Cristia, A., Reimchen, M., & Soderstrom, M. (2016). Entrainment of prosody in the interaction of mothers with their young children. *Journal of Child Language, 43*(2), 284–309.

- Larson, V. L., & McKinley, N. L. (1998). Characteristics of adolescents' conversations: A longitudinal study. *Clinical Linguistics & Phonetics*, 12, 183–203.
- Laurie-Rose, C., Bennett-Murphy, L., Curtindale, L. M., Granger, A. L., & Walker, H. B. (2005). Equating tasks and sustaining attention in children and adults: The methodological and theoretical utility of d' matching. *Perception & Psychophysics*, 67, 254–263.
- Lee, C., Black, M., Katsamanis, A., Lammert, A., Baucom, B., Christensen, A., . . . Narayanan, S. (2010). Quantification of prosodic entrainment in affective spontaneous spoken interactions of married couples. *Proceedings of the 11th Annual Conference of the International Speech Communication Association, 2010*, 793–796.
- Levitan, R., Beňuš, Š., Gravano, A., & Hirschberg, J. (2015). Acoustic-prosodic entrainment in Slovak, Spanish, English and Chinese: A cross-linguistic comparison. *Proceedings of the 16th Annual Meeting of the Special Interest Group on Discourse and Dialogue, 2015*, 325–334.
- Levitan, R., Gravano, A., Willson, L., Benus, S., Hirschberg, J., & Nenkova, A. (2012). Acoustic-prosodic entrainment and social behavior. *Proceedings of the 2012 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, 2012*, 11–19.
- Levitan, R., & Hirschberg, J. (2011). Measuring acoustic-prosodic entrainment with respect to multiple levels and dimensions. In *Proceedings of the 12th Annual Conference of the International Speech Communication Association, 2011*, 3081–3084.
- Local, J. (2007). Phonetic detail and the organisation of talk-in-interaction. *Proceedings of the 16th International Congress of Phonetic Sciences, 2007*, 1–10.
- Loose, R., Kaufmann, C., Auer, D. P., & Lange, K. W. (2003). Human prefrontal and sensory cortical activity during divided attention tasks. *Human Brain Mapping*, 18, 249–259.
- 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.
- McNelles, L. R., & Connolly, J. A. (1999). Intimacy between adolescent friends: Age and gender differences in intimate affect and intimate behaviors. *Journal of Research on Adolescence*, 9(2), 143–159.
- Natale, M. (1975). Convergence of mean vocal intensity in dyadic communication as a function of social desirability. *Journal of Personality and Social Psychology*, 32, 790–804.
- Nip, I. S. B., & Green, J. R. (2013). Increases in cognitive and linguistic processing primarily account for increases in speaking rate with age. *Child Development*, 84(4), 1324–1337.
- Oviatt, S. L., Darves, C., & Coulston, R. (2004). Toward adaptive conversational interfaces: Modeling speech convergence with animated personas. *ACM Transactions on Computer-Human Interaction*, 11(3), 300–328.
- Pardo, J. S. (2006). On phonetic convergence during conversational interaction. *The Journal of the Acoustical Society of America*, 119(4), 2382–2393.
- Pardo, J. S., Gibbons, R., Suppes, A., & Krauss, R. M. (2012). Phonetic convergence in college roommates. *Journal of Phonetics*, 40, 190–197.
- Pardo, J. S., Urmanche, A., Wilman, S., & Wiener, J. (2017). Phonetic convergence across multiple measures and model talkers. *Attention, Perception, & Psychophysics*, 79(2), 637–659.
- Persellin, D. (1992). Responses to rhythm patterns when presented to children through auditory, visual and kinesthetic modalities. *Journal of Research in Music Education*, 40(4), 306–315.
- Phillips-Silver, J., Aktipis, C. A., & Bryant, G. (2010). The ecology of entrainment: Foundations of coordinated rhythmic movement. *Music Perception*, 28(1), 3–14.
- Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, 27, 169–226.
- Putman, W. B., & Street, R. L., Jr. (1984). The conception and perception of noncontent speech performance: Implications for speech accommodation theory. *International Journal of the Sociology of Language*, 46, 97–114.
- R Core Team. (2018). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <https://www.R-project.org/>
- Rafaelli, M., & Duckett, E. (1989). “We were just talking...”: Conversations in early adolescence. *Journal of Youth and Adolescence*, 18(6), 567–582.
- Ratner, N. B. (1992). Measurable outcomes of instructions to modify normal parent-child verbal interactions: Implications for indirect stuttering therapy. *Journal of Speech and Hearing Research*, 35(1), 14–20.
- Rebok, G. W., Smith, C. B., Pascualvaca, D. M., Mirsky, A. F., Anthony, B. J., & Kellam, S. G. (1997). Developmental changes in attentional performance in urban children from eight to thirteen years. *Child Neuropsychology*, 3, 28–46.
- Rodrigues, P. F. S., & Pandeirada, J. N. S. (2018). When visual stimulation of the surrounding environment affects children's cognitive performance. *Journal of Experimental Child Psychology*, 176, 140–149.
- Samter, W. (2003). Friendship interaction skills across the life span. In J. O. Greene & B. R. Burlinson (Eds.), *Handbook of communication and social interaction skills* (pp. 637–684). Mahwah, NJ: Erlbaum.
- Savino, M., Lapertosa, L., & Refice, M. (2018). Seeing or not seeing your conversational partner: The influence of interaction modality on prosodic entrainment. In A. Karpov, O. Jokisch, & R. Potapova (Series Eds.), *Lecture Notes in Computer Science: Speech and computer* (Vol. 11096, pp. 574–584). https://doi.org/10.1007/978-3-319-99579-3_59
- Sawyer, J., Matteson, C., Ou, H., & Nagase, T. (2017). The effects of parent-focused slow relaxed speech intervention on articulation rate, response time latency, and fluency in preschool children who stutter. *Journal of Speech, Language, and Hearing Research*, 60(4), 794–809.
- Schweitzer, A., Lewandowski, N., & Duran, D. (2017). Social attractiveness in dialogs. *Proceedings of the 18th Annual Conference of the International Speech Communication Association, 2017*, 2243–2247.
- Schweitzer, K., Walsh, M., & Schweitzer, A. (2017). To see or not to see: Interlocutor visibility and likeability influence convergence in intonation. *Proceedings of the Annual Conference of the International Speech Communication Association, 2017*, 919–923.
- Stern-Eliran, K., Zilcha-Mano, S., Sebba, R., & Levit Binnun, N. (2016). Disruptive effects of colorful vs. non-colorful play area on structured play—A pilot study with preschoolers. *Frontiers in Psychology*, 7, 1661.
- Street, R. L., Street, N. J., & Van Kleek, A. (1983). Speech convergence among talkative and reticent three year-olds. *Language Sciences*, 5(1), 79–96.
- 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.
- Turkstra, L. S. (2001). Partner effects in adolescent conversations. *Journal of Communication Disorders*, 34(1–2), 151–162.
- Uptis, R. (1987). Children's understanding of rhythm: The relationship between development and music training. *Psychomusicology: A Journal of Research in Music Cognition*, 7(1), 41–60.

-
- Vohn, R., Fimm, B., Weber, J., Schnitker, R., Thron, A., Spijkers, W., . . . Sturm, W. (2007). Management of attentional resources in within-modal and cross-modal divided attention tasks: An fMRI study. *Human Brain Mapping, 28*, 1267–1275.
- Wiig, E. H., Semel, E., & Secord, W. A. (2013). *Clinical Evaluation of Language Fundamentals—Fifth Edition (CELF-5)*. Bloomington, MN: NCS Pearson.
- Wilson, M., & Wilson, T. P. (2005). An oscillator model of the timing of turn-taking. *Psychonomic Bulletin & Review, 12*, 957–968.
- 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.
- Xia, Z., Levitan, R., & Hirschberg, J. (2014). Prosodic entrainment in Mandarin and English: A cross-linguistic comparison. *Proceedings of the International Conference on Speech Prosody, 2014*, 65–69.
- Zhan, J. Y., Wilding, J., Cornish, K., Shao, J., Xie, C. H., Wang, Y. X., . . . Zhao, Z.-Y. (2011). Charting the developmental trajectories of attention and executive function in Chinese school-aged children. *Child Neuropsychology, 17*(1), 82–95.

Appendix

Example Transcript of Stimuli Recording

This is a picture from a book called *The Berenstain Bears Go Green*. I want you to describe this picture for me. You can tell me about what the houses look like. You can tell me about what the weather is like outside, or you can tell me about what the bears are doing or what they are wearing. Remember to keep talking until the timer runs out.
