Research Article

Methodology Matters: The Impact of Research Design on Conversational Entrainment Outcomes

Camille J. Wynn[a] and Stephanie A. Borrie[a]

Purpose: Conversational entrainment describes the tendency for individuals to alter their communicative behaviors to more closely align with those of their conversation partner. This communication phenomenon has been widely studied, and thus, the methodologies used to examine it are diverse. Here, we summarize key differences in research design and present a test case to examine the effect of methodology on entrainment outcomes.

Method: Sixty neurotypical adults were randomly assigned to experimental groups formed by a 2 × 2 factorial combination of two independent variables: stimuli organization (blocked vs. random presentation) and stimuli modality (auditory-only vs. audiovisual stimuli). Individuals participated in a quasiconversational design in which the speech of a virtual interlocutor was manipulated to produce fast and slow speech rate conditions.

Results: There was a significant effect of stimuli organization on entrainment outcomes. Individuals in the blocked, but not the random, groups altered their speech rate to align with the speech rate of the virtual interlocutor. There were no effect of stimuli modality and no interaction between modality and organization on entrainment outcomes.

Conclusion: Findings highlight the importance of methodological decisions on entrainment outcomes. This underscores the need for more comprehensive research regarding entrainment methodology.

C ommunication is anything but simple. This process, in which ideas are exchanged through a dynamic interplay of words, facial expressions, body postures, gestures, and prosodic features, is intricate and multifaceted. It is no surprise, therefore, that research in this area is riddled with obstacles and difficulties. The study of human communication requires adherence to subtle details and slight nuances while still maintaining a larger and more holistic perspective. Because of these complexities, careful attention to methodological choices employed in the study of communication is essential. Seemingly insignificant variations in the ways in which studies are designed may yield vastly different outcomes. Research questions in this area must focus not only on what answers can be found but also on how the overall research design may impact those findings. One area in which this idea is particularly important is in research regarding conversational entrainment.

Known by several different terms including alignment, accommodation, and convergence, conversational entrainment can be defined as the synchronization of behavior throughout communicative interactions. Put simply, during conversation, individuals will unconsciously alter their verbal and nonverbal behaviors to more closely align with the behaviors of their conversation partner. Within the realm of speech, conversational entrainment has been documented across many acoustic–prosodic features including speech rate (e.g., Freud et al., 2018; Manson et al., 2013), vocal intensity (e.g., Local, 2007; Natale, 1975), pitch properties (Borrie & Liss, 2014; C. Lee et al., 2010), and vocal quality (Borrie & Delfino, 2017; Levitan & Hirschberg, 2011). Furthermore, research has not only documented the pervasive occurrence of conversational entrainment but also has revealed a myriad of associated benefits including increased conversational fluidity (Local, 2007; Wilson & Wilson, 2005), better cooperation (Manson et al., 2013), and greater empathy, rapport, and closeness (C. Lee et al., 2010; Pardo et al., 2012).

Conversational entrainment has been studied for decades across a wide variety of disciplines including psychology, linguistics, neuroscience, sociology, and computer science, and the types of methodologies used to study this phenomenon are as numerous and diverse as these

[a]Department of Communicative Disorders and Deaf Education, Utah State University, Logan

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

disciplines. Regarding entrainment of speech production behaviors, we conceptualize methodologies along a spectrum, ranging from highly naturalistic conversational settings to tightly controlled experimental conditions. On one end of the spectrum, naturalistic studies examine entrainment in embodied conversations, which seek to replicate real-world contexts. Dyads are often free to discuss topics of their choice (e.g., Natale, 1975; A. Schweitzer & Lewandowski, 2013), though researchers may add some level of control by assigning specific conversation topics (e.g., Frued et al., 2018; Webb, 1969) or presenting dyads with a collaborative, goal-directed task (i.e., Map Task, Pardo et al., 2019; Diapix Task, Van Engen et al., 2010). Beyond topic of conversation, research designs may form dyads from two participants (e.g., Borrie et al., 2015; Pardo, 2006) or make use of a confederate to converse with each participant (e.g., Borrie & Delfino, 2017; Gregory, 1990). Furthermore, while some confederates are given no explicit instructions about how to use their speech, others are told to consciously alter their speech patterns (e.g., speak slowly) to meet specific requirements of different experimental conditions (e.g., Freud et al., 2018; Schultz et al., 2016). In some studies, conversational participants engage in face-to-face conversation (e.g., Borrie et al., 2020; Giles, 1973). In others, participants are separated by a curtain or divider (e.g., Levitan et al., 2012; Louwerse et al., 2012). Furthermore, conversation partners are often strangers (e.g., Pardo, 2006; Pardo et al., 2010), but some studies target participants who are familiar with one another such as roommates (Pardo et al., 2012) or married couples (C. Lee et al., 2010).

On the opposite end of the spectrum are studies that employ highly structured settings to study entrainment, which, while inducing some loss of ecological validity, afford a level of experimental control simply not achieved in more naturalistic contexts. The typical nature of these research designs is to make use of prerecorded speech stimuli; participants listen to the stimuli recordings and are asked to provide some type of verbal response. As with the more natural, conversation-based methodologies, there are many elements that differ across studies. Stimuli, for example, may consist of speech produced by a single speaker (e.g., Babel et al., 2013; Nguyen et al., 2012) or by multiple speakers (e.g., Babel et al., 2014; Namy et al., 2002). Recordings may be unaltered (e.g., Goldinger, 1998; Porter & Castellanos, 1980) or digitally manipulated to create different experimental conditions such as slow or fast speaking rates (e.g., Fowler et al., 2003; Staum Casasanto et al., 2010). Individual trials may consist of recordings that are several sentences (e.g., Kosslyn & Matt, 1977; Wynn et al., 2018), a single sentence (e.g., Borrie & Liss, 2014; Jungers et al., 2002), or even just a single word (Goldinger, 1998), syllable (Sanchez et al., 2010), or phoneme (Sato et al., 2013). Some stimuli sets are presented alongside a visual component such as a video or a still picture (e.g., Babel, 2012; Jungers et al., 2016), whereas others are exclusively auditory (Miller et al., 2013; Pardo et al., 2017). The types of verbal responses following stimuli presentation also vary. Participants may

be asked to read aloud scripted sentences (e.g., Borrie & Liss, 2014; Kosslyn & Matt, 1977), repeat the utterance of the stimuli (e.g., Goldinger, 1998; Pardo et al., 2013), or form their own spontaneous productions (e.g., Jungers & Hupp, 2009; Wynn et al., 2018).

Given the amount of research investigating entrainment in the speech domain, the diversity of methodologies is not surprising, nor is it problematic. In fact, using such a wide array of research designs can be extremely beneficial. All research designs carry inherent limitations, but through multimethodological research, the strengths of certain methods can offset the weaknesses of others. What is problematic, however, is that few studies have evaluated how these differences in approaches may alter research findings. While one hopes that variations in methodology do not implicate big-picture outcomes, this notion is likely too idealistic to hold true in all situations. Indeed, there are probably contexts in which entrainment is more easily achieved than others. Consequently, some methodological choices may alter study outcomes, and if this is the case, the validity of results may be largely dependent upon the degree to which research designs are well understood and well justified. Therefore, a line of study focused on methodologies and their effects on entrainment is an important component of research in this area.

As a first step, research should examine individual methodological factors and the role they play on research findings. Within the realm of speech, there are a small handful of studies that have explored this issue, primarily in respect to stimuli modality (i.e., audio vs. audiovisual aspects of stimuli; see Dias & Rosenblum, 2011; A. Schweitzer & Lewandowski, 2013; K. Schweitzer et al., 2017). However, research regarding other aspects of methodology is virtually nonexistent. Furthermore, focusing on individual factors, independent of one another, may not be sufficient. It is plausible that different aspects of the research design interact with each other to alter outcomes in ways that go undetected when such factors are studied in isolation. Accordingly, research investigating multiple factors and their relationships with each other may provide a more comprehensive picture regarding research design. In this study, we explore the influence of methodology on entrainment outcomes. Our hypothesis is twofold: (a) There are certain methodological factors that influence entrainment outcomes, and (b) outcomes may be the result of an interaction between two or more factors. To test these hypotheses, we present a test case in which we examine the effect of two methodological factors on entrainment outcomes.

First, we focus on stimuli organization—whether stimuli are presented in blocked sets or in a random fashion. In blocked presentation schedules, participants listen and respond to a stimulus set containing several recordings from the same speaker or condition before being presented with a different stimulus set. For example, in a study on speech rate entrainment, Jungers and Hupp (2009) employed a blocked paradigm in which participants listened and responded to recordings from one speech rate condition (e.g., slow rate condition) before listening and responding to

recordings from another condition (e.g., fast rate condition). Contrasting blocked schedules are randomized schedules in which trials from different speakers or conditions are interspersed at random with each other. For example, in a study similar to that of Jungers and Hupp (2009), participants were also asked to listen and respond to recordings from different speech rate conditions. However, rather than presenting stimulus sets in blocks, recordings from each condition were randomly interspersed with one another. It is probable that conversational entrainment is a dynamic process, in which individuals make gradual adjustments to their own speech patterns as they become more comfortable and familiar with the patterns of their conversation partner. Therefore, we would assume that blocked paradigms, which allow more time for familiarity and adaptation, would be more sensitive to capture entrainment than randomized paradigms. However, to date, this idea has not been investigated.

Another aspect of the methodology that may influence entrainment outcomes is stimuli modality—whether participants are presented with auditory stimuli only or whether auditory stimuli are coupled with visual information. As mentioned previously, there is a small body of research investigating the effect of stimuli modality on entrainment outcomes. However, research in this area is disparate and inconclusive. Some studies have shown that entrainment outcomes are more pronounced when communication partners are able to see each other than when they are only able to hear each other, indicating that visual cues may augment auditory information to enhance entrainment (Dias & Rosenblum, 2011). Others studies have shown greater entrainment in auditory-only contexts, suggesting that, in audiovisual contexts, individuals may exert greater focus toward nonverbal cues being presented, thus reducing the cognitive resources available to attend to verbal information (e.g., K. Schweitzer et al., 2017). Still, others show no significant difference between the two modalities (e.g., A. Schweitzer & Lewandowski, 2013). Given the disparate findings, it is possible that stimuli modality influences entrainment outcomes, but does so in tandem with other methodological factors that, together, interact to shape the given result. For example, individuals in random stimuli organizations may be more successful in an audiovisual condition where they can rely on visual information to compensate for the rapidly changing acoustic patterns. Contrastingly, individuals in the blocked organizations may be less reliant and therefore more distracted by extraneous visual information. Therefore, they may be more successful in an auditory-only condition. Thus, it is possible that stimuli modality and organization interact with each other, in unique ways, to affect entrainment outcomes.

In this study, we employ a quasiconversational paradigm and focus on entrainment of speech rate to target two key research questions: (a) Is speech rate entrainment modulated by differences in the presentation of stimuli organization and/or the modality in which the speech stimuli are presented? (b) Is speech rate entrainment modulated by an interaction between the presentation of stimuli organization and stimuli modality? If significant, these results would highlight the importance of methodological decisions on entrainment outcomes and underscore the need for further research regarding entrainment methodology.

Method

Overview

Participants, randomly assigned to one of four experimental groups (see Figure 1), engaged in a quasiconversation with a virtual interlocutor. Groups ($n = 15$) were formed by a 2 × 2 factorial combination of two independent variables: stimuli organization (blocked vs. random) and stimuli modality (audiovisual vs. audio-only). Linear mixed models were used to compare differences in speech rate entrainment outcomes between groups.

Participants

Participants included 60 (51 women, nine men) neurotypical adults between the ages of 18 and 33 years ($M = 22.6$, $SD = 2.6$). All participants were native speakers of American English with no identified intellectual, language, or speech impairments. Prior to the experimental task, all participants passed a pure-tone hearing screening administered at 20 dB for 1000, 2000, and 4000 Hz in both ears.

Entrainment Task and Stimuli

The entrainment task used was an 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 a video camera (Canon EOS 70D). 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 is positioned against a neutral backdrop with the camera positioned to capture a view of her head and shoulders. The woman 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

Figure 1. Participants were randomly assigned to one of four experimental groups formed by a 2 × 2 factorial combination of stimuli organization and stimuli modality contexts.

		Stimuli Modality	
		Audio-only	Audio-visual
Stimuli Organization	Blocked	$n = 15$	$n = 15$
	Random	$n = 15$	$n = 15$

examples of what participants could talk about for each picture (see the Appendix for a sample transcript). A total of 16 audiovisual recordings, each approximately 20–25 s in length, were produced. Audio stimuli, for the audio-only groups, were created by extracting .mp3 files from the audiovisual recordings. Both audio and audiovisual recordings were then digitally manipulated to form two different conditions. Half of the recordings were assigned to the slow condition (digitally manipulated to 80% of the original rate), and half were assigned to the fast condition (digitally manipulated to 120% of the original rate). This level of rate modification was chosen in order to provide a natural appearance, while still providing sufficient variability in the speech rate of the speaker. Trials were embedded in a web-based application, hosted on a secure university server.

Procedure

The experiment was conducted in a quiet room. Participants were blind to the purpose of the study. After obtaining informed consent, participants were seated in front of a computer screen to undergo the experimental task. The researcher explained to the participants that they would be watching/listening to a series of short video clips of a woman talking about some pictures and that, immediately following each clip, the picture described by the woman would appear on the screen. The participants were instructed to watch/listen to each recording and then describe the picture. They were informed that they should continue talking for 15 s, until a visual cue (stop sign symbol on the screen) and an auditory cue (short beep) let them know that it was time to stop. The experimental procedure began with one practice trial, using an unaltered clip of a novel picture. Participants then viewed and responded to each of the 16 stimuli clips. The average total time to complete the procedure was 15 min.

As noted in the overview, prior to the experimental procedure, participants were randomly assigned to one of the four experimental groups (again, see Figure 1). First, in order to determine the effect of stimuli organization, stimuli were presented to participants either in blocks or randomly. Participants in the two blocked groups were presented with stimuli of the same speech rate condition in blocks (i.e., eight fast and then eight slow, or vice versa), whereas participants in the two randomized groups were presented with stimuli of both conditions randomly interspersed with each other. In addition to stimuli organization, the effect of stimuli modality was examined. Participants in the two audiovisual groups were presented with audiovisual clips, whereas participants in the two audio-only groups were presented with audio clips while looking at a neutrally colored backdrop on the computer screen. Within both the audio-only and audiovisual blocked groups, presentation of the stimuli was counterbalanced so that equal numbers of individuals were presented with the fast condition first and with the slow condition first.

Data Analysis

The total data set for the entrainment task consisted of 960 audio response recordings—480 response recordings for the slow condition and 480 response recordings for the fast condition. Two trained research assistants used an acoustic analysis software, Praat (Boersma & Weenink, 2018), to manually calculate speech rate in syllables per second for each individual response recording. Research assistants orthographically transcribed each response recording and counted the number of syllables for each production. All 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. They then measured the entire duration of the response recording, beginning with the moment the participant began articulating their response and ending when articulation of the participant's response (within the 15-s time frame) ceased. Speech rate for each response recording was then calculated by dividing the total number of syllables by the duration measure. During analysis, assistants were blinded to the participant group and the response condition of each recording. In order to obtain interrater reliability for speech rate calculations, 25% of the total data set (15 participant data sets) was randomly selected by a computer-generated random number list and analyzed by both research assistants. Comparison indicated high agreement between the two judges, with a Pearson correlation r score of .99.

Linear mixed models were used to analyze the data in the R statistical environment (R Version 3.5.2; R Development Core Team, 2018) using the lme4 package (lme4 package Version 1.1–19; Bates et al., 2015). This analysis was used to investigate the effects of speech rate condition, stimuli organization, and stimuli modality on average speech rate while controlling for intraindividual variability across the repeated measures. Effects were tested using a series of nested mixed-effects models to determine the best-fitting model. For each model, the random effects structure included a random intercept by participant. Fixed effects within each model included the within-participant factor of condition (i.e., slow stimuli vs. fast stimuli) and the between-participant factors of stimuli organization (i.e., blocked vs. random) and stimuli modality (i.e., audio vs. audiovisual). Additionally, sex was included in each model as a fixed effect to account for potential variability between male and female participants. The first model included main effects for condition, organization, and modality (Model 1). The second model included a two-way interaction between condition and organization while controlling for modality (Model 2). The third model included a two-way interaction between condition and organization and a two-way interaction between condition and modality (Model 3). Finally, the last model included a three-way interaction between condition, organization, and modality (Model 4).

Results

Based on likelihood ratio tests, the most parsimonious best-fitting model was Model 2, which included the

two-way interaction between condition and organization, but not between condition and modality (see Table 1). Thus, speech rate entrainment was influenced by the organization of stimuli but not by the modality of stimuli. Analysis of Model 2 revealed a significant interaction between condition and organization (see Table 2). This interaction, as illustrated in Figure 2, revealed that participants entrained their speech rate with the virtual interlocutor in the blocked organization condition, but not in the random organization condition.

Our findings, demonstrating a significant effect of stimuli organization on speech rate entrainment, warranted further analysis to investigate whether the difference between groups could be accounted for by changes in entrainment over time. We hypothesized that the degree of entrainment would increase within the blocked condition over time as participants became more familiar with the speech rate patterns of the interlocutor and subsequently adjusted their own patterns, while the degree of entrainment (or lack thereof) in the random condition would remain stable over time. Therefore, a secondary analysis using linear mixed models was conducted to examine the interaction between block, condition, and time (as measured by trial number). As with the previous analysis, the random effects structure included a random intercept by participant, and sex was included as a fixed effect. Results revealed no significant interaction between block, condition, and time, indicating that the difference in the degree of entrainment between groups did not change over time.

Discussion

The aim of this study was to examine the influence of methodology on entrainment outcomes in healthy adult populations. As a test case, we focused on the effects of stimuli organization and stimuli modality on speech rate entrainment with a virtual interlocutor. Our hypothesis consisted of two separate components. First, certain methodological factors will influence entrainment outcomes. Second, outcomes may be the result of an interaction between two or more factors. We comment on the current findings and how they relate to these hypotheses below.

Our findings showed that the methodological factor of stimuli organization played a significant role in entrainment outcomes. Individuals in the blocked groups spoke more slowly in the slow condition than in the fast condition, revealing significant levels of entrainment. In contrast, individuals in the random groups did not modulate their speech rate depending on the condition. We postulated that blocked stimuli presentation may be more amendable to entrainment because it allows individuals greater opportunity to learn, and subsequently adapt, to the interlocutor's speech rate. The interaction between condition, group, and clip number, however, was nonsignificant, implicating that differences between the blocked and random groups cannot be accounted for as a function of time. Thus, the result of significant entrainment in the blocked but not random groups warrants exploration of other explanations. One prominent model of learning, the model of contextual interference, makes predictions about learning in blocked versus random contexts. One of the model predictions is that, during initial learning of a new task, individuals will perform better in blocked presentation formats than randomized formats (Shea & Morgan, 1979). According to the action-plan reconstruction hypothesis, a more specific conceptualization of this model, individuals form an action plan before engaging in a given task. In blocked contexts, these plans remain in the working memory, and individuals are able to reactivate them on successive trials. However, in randomized contexts, differing actions plans cannot be stored in the working memory at the same time. Therefore, individuals are required to reconstruct their action plans on every trial, leading to more effortful cognitive processing and increased performance error (T. D. Lee & Magill, 1983, 1985). The effect of stimuli organization on novel task acquisition has been verified in complex rhythmic activities such as instrumental music (e.g., Abushanab & Bishara, 2013) and dance (Bertollo et al., 2010). Conversational entrainment is, of course, a complex rhythmic activity demanding perception and adjustment of rhythmic speech patterns. As such, the current findings, in which speech rate entrainment is evident in the blocked but not random groups, are supported by this framework.

While there was a significant effect of stimuli organization, there were no significant main effects of stimuli

Table 1. Linear mixed-model fit indices for models of interest.

Model	AIC	BIC	Log likelihood	χ^2	χ^2 difference	df	p
Model 1	1,469.6	1,503.7	−727.80	1,455.6			
Model 2	**1,463.5**	**1,502.4**	**−723.73**	**1,447.5**	**8.1410**	**1**	**.004**
Model 3	1,464.0	1,507.8	−723.01	1,446.0	1.4440	1	.229
Model 4	1,467.4	1,520.9	−722.68	1,445.4	0.6593	2	.719

Note. Model 1 included main effects for condition, stimuli organization, and stimuli modality. Model 2 included a two-way interaction term for condition and organization. Model 3 included a two-way interaction term for condition and organization and a two-way interaction term for condition and modality as well as a three-way interaction term for condition, organization, and modality. Bold text represents the best-fitting model based on likelihood ratio tests. AIC = Akaike information criterion; BIC = Bayesian information criterion.

Table 2. Results of Model 2 with a two-way interaction for condition and stimuli organization while controlling for stimuli modality and sex.

Term	B	SE	t value	p value
Intercept	3.969	0.121	32.778	< .001**
Condition	0.129	0.043	3.003	.003*
Organization	0.053	0.137	0.387	.700
Modality	0.043	0.132	0.329	.744
Gender	0.373	0.187	1.998	.051
Condition × Organization	0.174	0.061	2.857	.004*

*p < .01. **p < .001.

modality on entrainment outcomes, nor was there a significant interaction between stimuli organization and modality. This finding is in line with A. Schweitzer and Lewandowski (2013), showing that stimuli modality did not play a significant role in entrainment outcomes. There are several possible explanations for this finding. From a theoretical perspective, some researchers argue that visual information aids entrainment (i.e., provides additional compensatory information; Dias & Rosenblum, 2011), whereas others have argued that visual information is disadvantageous (i.e., presents unnecessary distraction; K. Schweitzer et al., 2017). However, these arguments are not necessarily mutually exclusive. One could imagine that visual information carries both the abovementioned advantages and disadvantages, and in certain contexts such as this study, these counterbalance each other, resulting in similar results for both the audiovisual and audio-only groups. Additionally, from a more pragmatic viewpoint, most adults communicate regularly via both audiovisual (e.g., face-to-face conversations) and audio-only (e.g., telephone conversations) channels. Therefore, adults may be adept at adjusting their speech rate patterns regardless of the communication modality. Our findings contrast those who did find that stimuli modality affected entrainment outcomes (Dias & Rosenblum, 2011; K. Schweitzer et al.,

2017). Therefore, although there was no interaction between stimuli organization and modality in this study, it is also plausible that stimuli modality is moderated by other methodological factors not investigated in this study. For example, while our study and that of A. Schweitzer and Lewandowski (2013) examined entrainment of speech rate, the studies that did find an influence measured pitch accents and articulatory behavior. Thus, it is possible that the effect of stimuli modality may be moderated by the type of acoustic feature being measured.

Beyond simply providing information regarding the effects of stimuli organization and modality on entrainment outcomes, this study affords important evidence that methodology matters. As evidenced in this study, small changes in methodology can yield large transformations of study outcomes. Here, we demonstrate that a simple shift —from a blocked presentation to a random presentation— meant the difference between detecting and not detecting significant levels of entrainment. Such effects are likely not exclusive to stimuli organization. Significant entrainment outcomes may be undercut by any number of other methodological factors. Our findings, therefore, highlight the need for well-designed research studies in this area of communication research. This does not imply that those designs with promises of significant findings should always be used simply for the sake of getting more notable results. It does imply, however, that in research regarding entrainment, conclusions cannot simply be taken at face value, and research design decisions must be carefully considered.

While important for any study, considerations regarding entrainment methodology are especially significant in studies with populations whose entrainment skills may be impaired or not fully developed. For example, in a recent large-scale study by Wynn et al. (2019), speech rate entrainment was not detected in the production behaviors of a large cohort of 48 children between the ages of 5 and 14 years. Rather than concluding that these children simply do not entrain their speech rate, the authors advanced

Figure 2. Entrainment outcomes are affected by stimuli organization but not stimuli modality. Error bars delineate ± 1 *SEM*.
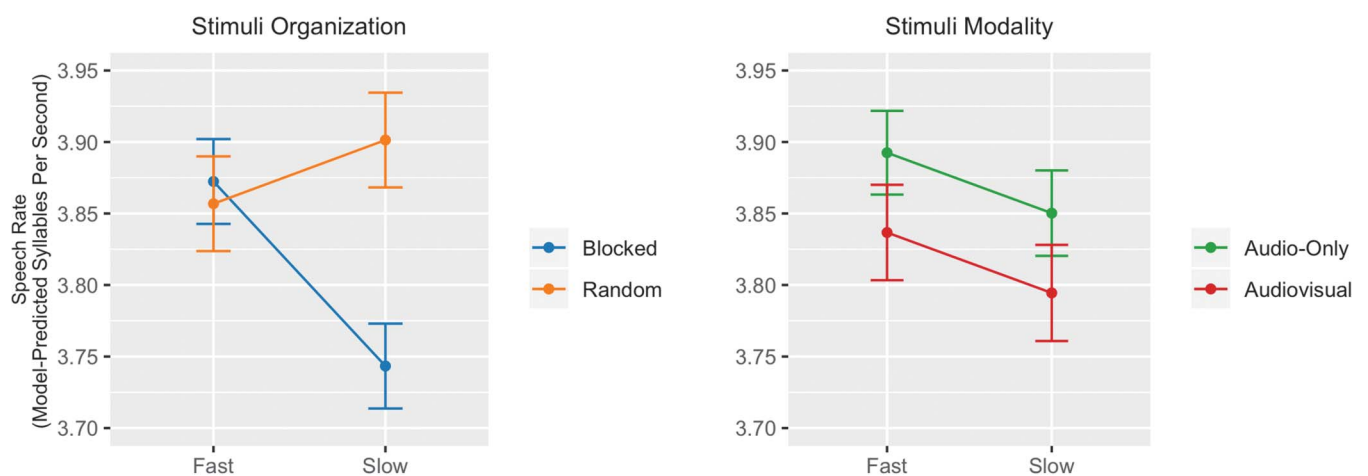
that entrainment is a multistep process in which skills "…emerge over time and, through continued practice, become increasingly refined and solidified" (Wynn et al., 2019, p. 3710). Thus, the design used in that study may have been too challenging to capture emerging entrainment skills. Other studies, however, may employ research designs that are too simple, making it difficult to detect differences in entrainment across development. Thus, identifying factors that increase or decrease the sensitivity of the research designs to entrainment outcomes would assist in making well-informed and suitably justified methodological decisions. Another group in which the influence of methodology decisions may be particularly pertinent is individuals with communication disorders. While still a relatively new area of exploration in speech-language pathology, there is a growing body of research highlighting entrainment deficits in individuals with autism spectrum disorder (Wynn et al., 2018), dysarthria (Borrie et al., 2015, 2020), traumatic brain injury (Gordon et al., 2015), fluency disorders (Sawyer et al., 2017), and hearing impairments (Freeman & Pisoni, 2017). As with children, further research in this area requires balanced designs—ones that are sensitive enough to detect entrainment impairments while still capturing the weaker patterns of entrainment that may exist in these populations. Thus, a greater understanding of the contextual factors that influence entrainment would be beneficial for clinical populations with entrainment deficits.

The benefit gained from exploring entrainment methodologies is not limited to more informed research designs. The study of entrainment methodologies may also offer many theoretical and clinical implications. Indeed, the concept of gaining foundational knowledge through the study of research methodologies has been successful in many areas of social science. For example, in their meta-analysis on educational performance outcomes, de Boer et al. (2014) examined studies that focused on different types of educational intervention and academic outcome measures. However, rather than searching for the interventions that were most effective, the authors leveraged the methodological differences between studies (e.g., computer-implemented vs. teacher-implemented instruction, cooperative vs. individual learning) to identify aspects of teaching, aside from actual content, that were most influential in student outcomes. Boer and colleagues found that methodological differences accounted for over 60% of the variance in the outcome effect sizes between studies and, thus, were able to highlight key aspects of student learning. In another study, researchers found that methodological differences accounted for between-study variability regarding the typical fundamental frequency range in healthy adults (Reich et al., 1990). Although focused on understanding the effects of research design, their findings led to subsequent studies and further conversations about the best way to elicit phonation frequency range in clinical voice evaluations (e.g., Zraick et al., 2000).

We advance that similarly valuable insights may be gained from research focused on entrainment methodologies. For example, while still speculative at this point, our findings that the blocked organization led to greater entrainment than random organization may carry theoretical implications for differences between entrainment in conversations between dyads and multiple speakers. That is, in conversations with more than two conversation partners, where speakers must constantly adjust their speech patterns to align with different speakers, entrainment may be more difficult to achieve. Additionally, study of entrainment methodologies may yield important findings for clinical populations with entrainment deficits. For example, a better understanding of the contextual factors that influence entrainment could inform therapeutic techniques in which communication partners are taught strategies to scaffold conversational environments so that individuals with entrainment impairments are more easily able to synchronize their speech.

Conclusion

In summary, this study investigated the impact of methodological decisions on entrainment outcomes using a test case that examined the main effects and interaction effects of stimuli organization and stimuli modality. We showed that variations in one of these factors, stimuli organization, led to significant differences in entrainment outcomes. Thus, this study affords empirical evidence to support the statement that "methodology matters." It also highlights the need for further, more comprehensive research in this area, including systematic reviews that target the effects of methodological factors in existing studies.

Acknowledgments

References

Abushanab, B., & Bishara, A. J. (2013). Memory and metacognition for piano melodies: Illusory advantages of fixed- over random-order practice. *Memory & Cognition, 41*(6), 928–937. https://doi.org/10.3758/s13421-013-0311-z

Babel, M. (2012). Evidence for phonetic and social selectivity in spontaneous phonetic imitation. *Journal of Phonetics, 40*(1), 177–189. https://doi.org/10.1016/j.wocn.2011.09.001

Babel, M., McAuliffe, M., & Haber, G. (2013). Can mergers-in-progress be unmerged in speech accommodation? *Frontiers in Psychology, 4,* 653. https://doi.org/10.3389/fpsyg.2013.00653

Babel, M., McGuire, G., Walters, S., & Nicholls, A. (2014). Novelty and social preference in phonetic accommodation. *Laboratory Phonology, 5*(1), 123–150. https://doi.org/10.3389/fpsyg.2013.00653

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bertollo, M., Berchicci, M., Carraro, A., Comani, S., & Robazza, C. (2010). Blocked and random practice organization in the learning of rhythmic dance step sequences. *Perceptual and

Motor Skills, 110(1), 77–84. https://doi.org/10.2466/pms.110.1.77-84

Boersma, P., & Weenink, D. (2018). *Praat: Doing phonetics by computer (Version 6.0)* [Computer software]. http://www.praat.org

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. https://doi.org/10.1016/j.jvoice.2016.12.005

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. https://doi.org/10.1044/2014_JSLHR-S-13-0149

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/10.3389/fpsyg.2015.01187

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., 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

de Boer, H., Donker, A. S., & van der Werf, M. P. C. (2014). Effects of the attributes of educational interventions on students' academic performance: A meta-analysis. *Review of Educational Research, 84*(4), 509–545. https://doi.org/10.3102/0034654314540006

Dias, J. W., & Rosenblum, L. D. (2011). Visual influences on interactive speech alignment. *Perception, 40*(12), 1457–1466. https://doi.org/10.1068/p7071

Fowler, C. A., Brown, J. M., Sabadini, L., & Weihing, J. (2003). Rapid access to speech gestures in perception: Evidence from choice and simple response time tasks. *Journal of Memory and Language, 49*(3), 396–413. https://doi.org/10.1016/S0749-596X(03)00072-X

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. https://doi.org/10.1121/1.4998590

Freud, D., Ezrati-Vinacour, R., & Amir, O. (2018). Speech rate adjustment of adults during conversation. *Journal of Fluency Disorders, 57,* 1–10. https://doi.org/10.1016/j.jfludis.2018.06.002

Giles, H. (1973). Accent mobility: A model and some data. *Anthropological Linguistics, 15*(2), 87–105. https://doi.org/10.4236/ojs.2013.35038

Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review, 105*(2), 251–279. https://doi.org/10.1037/0033-295x.105.2.251

Gordon, R. G., Rigon, A., & Duff, M. C. (2015). Conversational synchrony in the communicative interactions of individuals with traumatic brain injury. *Brain Injury, 29*(11), 1300–1308. https://doi.org/10.3109/02699052.2015.1042408

Gregory, S. W. (1990). Analysis of fundamental frequency reveals covariation in interview partners' speech. *Journal of Nonverbal Behavior, 14*(4), 237–251. https://doi.org/10.1007/BF00989318

Jungers, M. K., & Hupp, J. M. (2009). Speech priming: Evidence for rate persistence in unscripted speech. *Language and Cognitive Processes, 24*(4), 611–624. https://doi.org/10.1080/01690960802602241

Jungers, M. K., Hupp, J. M., & Dickerson, S. D. (2016). Language priming by music and speech: Evidence of a shared processing mechanism. *Music Perception: An Interdisciplinary Journal, 34*(1), 33–39. https://doi.org/10.1080/01690960802602241

Jungers, M., Palmer, C., & Speer, S. (2002). Time after time: The coordinating influence of tempo in music and speech. *Cognitive Processing, 1*(1), 21–35.

Kosslyn, S. M., & Matt, A. M. C. (1977). If you speak slowly, do people read your prose slowly? Person-particular speech recoding during reading. *Bulletin of the Psychonomic Society, 9*(4), 250–252. https://doi.org/10.3758/BF03336990

Lee, C., Black, M., Katsamanis, A., Lammert, A., Baucom, B., Christensen, A., Georgiou, P., & Narayanan, S. (2010). Quantification of prosodic entrainment in affective spontaneous spoken interactions of married couples. In *Proceedings of the 11th Annual Conference of the International Speech Communication Association* (pp. 793–796). INTERSPEECH.

Lee, T. D., & Magill, R. A. (1983). The locus of contextual interference in motor-skill acquisition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9*(4), 730–746. https://doi.org/10.1037/0278-7393.9.4.730

Lee, T. D., & Magill, R. A. (1985). Can forgetting facilitate skill acquisition? In D. Goodman, R. B. Wilberg, & I. M. Franks (Eds.), *Differing perspectives in motor learning, memory, and control* (pp. 3–22). Elsevier.

Levitan, R., Gravano, A., Willson, L., Beňuš, Š, Hirschberg, J., & Nenkova, A. (2012). Acoustic-prosodic entrainment and social behavior. In E. Fosler-Lussier, E. Riloff, & S. Bangalore (Eds.), *Proceedings of the 2012 conference of the North American chapter of the association for computational linguistics: Human language technologies* (pp. 11–19). North American Chapter of the Association for Computational Linguistics.

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* (pp. 3081–3084). INTERSPEECH.

Local, J. (2007). Phonetic detail and the organization of talk-in-interaction. In J. Trouvain & W. J. Barry (Eds.), *Proceedings of the 16th ICPhS* (ID 1785). ICPhS.

Louwerse, M. M., Dale, R., Bard, E. G., & Jeuniaux, P. (2012). Behavior matching in multimodal communication is synchronized. *Cognitive Science, 36*(8), 1404–1426. https://doi.org/10.1111/j.1551-6709.2012.01269.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

Miller, R. M., Sanchez, K., & Rosenblum, L. D. (2013). Is speech alignment to talkers or tasks? *Attention, Perception, & Psychophysics, 75*(8), 1817–1826. https://doi.org/10.3758/s13414-013-0517-y

Namy, L. L., Nygaard, L. C., & Sauerteig, D. (2002). Gender differences in vocal accommodation: The role of perception. *Journal of Language and Social Psychology, 21*(4), 422–432. https://doi.org/10.1177/026192702237958

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. https://doi.org/10.1037/0022-3514.32.5.790

Nguyen, N., Dufour, S., & Brunellière, A. (2012). Does imitation facilitate word recognition in a non-native regional accent? *Frontiers in Psychology, 3,* 480. https://doi.org/10.3389/fpsyg.2012.00480

Pardo, J. S. (2006). On phonetic convergence during conversational interaction. *The Journal of the Acoustical Society of America, 119*(4), 2382–2393. https://doi.org/10.1121/1.2178720

Pardo, J. S., Gibbons, R., Suppes, A., & Krauss, R. M. (2012). Phonetic convergence in college roommates. *Journal of Phonetics, 40*(1), 190–197. https://doi.org/10.1016/j.wocn.2011.10.001

Pardo, J. S., Jay, I. C., & Krauss, R. M. (2010). Conversational role influences speech imitation. *Attention, Perception, & Psychophysics, 72*(8), 2254–2264. https://doi.org/10.3758/BF03196699

Pardo, J. S., Jordan, K., Mallari, R., Scanlon, C., & Lewandowski, E. (2013). Phonetic convergence in shadowed speech: The relation between acoustic and perceptual measures. *Journal of Memory and Language, 69*(3), 183–195. https://doi.org/10.1016/j.jml.2013.06.002

Pardo, J. S., Urmanche, A., Gash, H., Wiener, J., Mason, N., Wilman, S., Francis, K., & Decker, A. (2019). The Montclair map task: Balance, efficacy, and efficiency in conversational interaction. *Language and Speech, 62*(2), 378–398. https://doi.org/10.1177/0023830918775435

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. https://doi.org/10.3758/s13414-016-1226-0

Porter, R. J., & Castellanos, F. X. (1980). Speech-production measures of speech perception: Rapid shadowing of VCV syllables. *The Journal of the Acoustical Society of America, 67*(4), 1349–1356. https://doi.org/10.1121/1.384187

R Development Core Team. (2018). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing. https://www.R-project.org/

Reich, A. R., Frederickson, R. R., Mason, J. A., & Schlauch, R. S. (1990). Methodological variables affecting phonational frequency range in adults. *Journal of Speech and Hearing Disorders, 55*(1), 124–131. https://doi.org/10.1044/jshd.5501.124

Sanchez, K., Miller, R. M., & Rosenblum, L. D. (2010). Visual influences on alignment to voice onset time. *Journal of Speech, Language, and Hearing Research, 53*(2), 262–272. https://doi.org/10.1044/1092-4388(2009/08-0247)

Sato, M., Grabski, K., Garnier, M., Granjon, L., Schwartz, J.-L., & Nguyen, N. (2013). Converging toward a common speech code: Imitative and perceptuo-motor recalibration processes in speech production. *Frontiers in Psychology, 4,* 422. https://doi.org/10.3389/fpsyg.2013.00422

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. https://doi.org/10.1044/2016_JSLHR-S-16-0002

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

Schweitzer, K., Walsh, M., & Schweitzer, A. (2017). To see or not to see: Interlocutor visibility and likeability influence convergence in intonation. In *Proceedings of the 18th Annual Conference of the International Speech Communication Association* (pp. 919–923). INTERSPEECH.

Schultz, B. G., O'Brien, I., Phillips, N., McFarland, D. H., Titone, D., & Palmer, C. (2016). Speech rates converge in scripted turn-taking conversations. *Applied Psycholinguistics, 37*(05), 1201–1220. https://doi.org/10.1017/S0142716415000545

Shea, J. B., & Morgan, R. L. (1979). Contextual interference effects on the acquisition, retention, and transfer of a motor skill. *Journal of Experimental Psychology: Human Learning and Memory, 5*(2), 179–187. https://doi.org/10.1037/0278-7393.5.2.179

Staum Casasanto, L., Jasmin, K., & Casasanto, D. (2010). Virtually accommodating: Speech rate accommodation to a virtual interlocutor. In S. Ohlsson & R. Catrambone (Eds.), *Proceedings of the 32nd Annual Conference of the Cognitive Science Society* (pp. 127–132). Cognitive Science Society.

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

Webb, J. T. (1969). Subject speech rates as a function of interviewer behaviour. *Language and Speech, 12*(1), 54–67. https://doi.org/10.1177/002383096901200105

Wilson, M., & Wilson, T. P. (2005). An oscillator model of the timing of turn-taking. *Psychonomic Bulletin & Review, 12*(6), 957–968. https://doi.org/10.3758/BF03206432

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

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

Zraick, R. I., Nelson, J. L., Montague, J. C., & Monoson, P. K. (2000). The effect of task on determination of maximum phonational frequency range. *Journal of Voice, 14*(2), 154. https://doi.org/10.1016/s0892-1997(00)80022-3

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.

Formula for the model constructed for statistical analysis:
    lmer(rate ~ condition*age+ condition*set + condition* gender + (1 | participant), data = children)