ASHA American Speech-Language-Hearing Association

JSLHR

Research Article

Conversational Speech Behaviors Are Context Dependent

Camille J. Wynn,[a] Tyson S. Barrett,[b] and Stephanie A. Borrie[b]

[a]Department of Communication Sciences and Disorders, University of Houston, TX [b]Department of Communicative Disorders and Deaf Education, Utah State University, Logan

ARTICLE INFO

Article History:
Received October 5, 2023
Revision received January 1, 2024
Accepted February 12, 2024

Editor-in-Chief: Cara E. Stepp
Editor: Shaheen N. Awan

https://doi.org/10.1044/2024_JSLHR-23-00622

ABSTRACT

Purpose: According to the interpersonal synergy model of spoken dialogue, interlocutors modify their communicative behaviors to meet the contextual demands of a given conversation. Although a growing body of research supports this postulation for linguistic behaviors (e.g., semantics, syntax), little is understood about how this model applies to speech behaviors (e.g., speech rate, pitch). The purpose of this study is to test the hypothesis that interlocutors adjust their speech behaviors across different conversational tasks with different conversational goals.
Method: In this study, 28 participants each engaged in two different types of conversations (i.e., relational and informational) with two partners (i.e., Partner 1 and Partner 2), yielding a total of 112 conversations. We compared six acoustic measures of participant speech behavior across conversational task and partner.
Results: Linear mixed-effects models demonstrated significant differences between speech feature measures in informational and relational conversations. Furthermore, these findings were generally robust across conversations with different partners.
Conclusions: Results suggest that contextual demands influence speech behaviors. These findings provide empirical support for the interpersonal synergy model and highlight important considerations for assessing speech behaviors in individuals with communication disorders.

Human behavior is context dependent. The way in which we walk (e.g., Hu et al., 2018; Kuntapun et al., 2020), eat (Ishida et al., 2002; Steele et al., 2015), perceive objects (Chun, 2000; Todorović, 2010), evaluate emotions (Barrett et al., 2011), make decisions (Mellers et al., 1998), and perform countless other daily activities depends on the context in which we perform them. This is especially evident in the realm of communication—speech and language behaviors differ across communicative tasks. For example, a large body of research has documented differences in the speech characteristics (e.g., speech rate, pitch) of read speech versus spontaneous spoken monologues (e.g., de Ruiter, 2015; Furui et al., 2005; Howell & Kadi-Hanifi, 1991; Silverman et al., 1992). Furthermore, different types of monologic speech tasks such as informal narratives and conference presentations exhibit different acoustic properties (Nakamura et al., 2008; Tasko & McClean, 2004). Beyond speech characteristics, variations in linguistic characteristics, such as lexical choices and syntactic structures, can also be found between communicative tasks including formal speeches, radio broadcasts, classroom lectures, and religious sermons (Bentum et al., 2019; Wiggers & Rothkrantz, 2007).

One area in which context may play an especially prominent role in speech and language behavior is within conversation. According to the interpersonal synergy model of spoken dialogue, interlocutors modify their communicative behaviors to meet the contextual demands of the conversation (Fusaroli et al., 2014).[1] These demands may be internal; modifications may reflect an interlocutor's conversational goals and the degree to which certain

Correspondence to Camille J. Wynn: cjwynn@uh.edu. *Disclosure: The authors have declared that no competing financial or nonfinancial interests existed at the time of publication.*

[1]Other theoretical models also support the idea of communicative adaptations based on conversational context. For example, the hypo–hyper theory (Lindblom, 1990) asserts that speakers vary the precision of their speech based on listener demands. Although such theories are useful and offer similar perspectives, we focus on the interpersonal synergy model because of its broad application to many different acoustic features and types of contextual demands.

 Copyright © 2024 American Speech-Language-Hearing Association

Downloaded from: https://pubs.asha.org Stephanie Borrie on 05/13/2024, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

behaviors help accomplish these goals. Demands may also be external. For example, interlocutors may modify their speech based on linguistic or cognitive demands brought about by the given conversational task. Notably, internal and external demands are inextricably linked. For instance, an interlocutor's goal may dictate the type of conversational task that they choose to engage in. A small but growing body of research supports the interpersonal synergy model. Dideriksen et al. (2023) compared the linguistic properties of informational conversations (i.e., conversations in which participants exchanged information to complete a collaborative task) to relational conversations (i.e., conversations in which participants asked and answered questions about themselves to establish rapport with one another). The authors found that interlocutors were more likely to use conversational behaviors such as targeted forms of conversational repair (i.e., requests for clarification) in informational conversations, which required higher levels of understanding to successfully complete the given task. Contrastingly, conversational behaviors, such as back channels (e.g., *yeah*, *uh-huh*), were used more frequently in relational conversations, presumably to convey attention and engagement. Thus, the behaviors observed in the conversations were shaped by contextual differences brought about by different conversational tasks and goals. In a related line of study, Reitter et al. (2006) found that dyads exhibited higher levels of syntactic entrainment (i.e., modifying their syntactic structures to become more similar to their conversational partner) in informational than relational conversations. The authors suggest that, in informational conversations, interlocutors may rely more on this strategy to reduce cognitive load and foster the mutual understanding necessary to accomplish their goals. Similar patterns of linguistic adaptation have frequently been found in the interactions of parents who alter several aspects of their language (i.e., utterance length, lexical diversity, syntactic complexity, linguistic style) to meet the demands of different types of interactions with their children (Crain-Thoreson et al., 2001; Haden & Fivush, 1996; Hoff-Ginsberg, 1991).

Although there is a growing body of research supporting the interpersonal synergy model in regards to linguistic behaviors (i.e., what was said), there is little research examining the role of contextual demands on speech behaviors (i.e., how the message was produced). However, there is ample reason to believe that speech behaviors (e.g., speech rate, pitch) will also be shaped by contextual demands. Research has shown that speech behaviors are often used to enhance communication. For example, individuals use speech signals to convey emotion (Carl et al., 2022), highlight information (Koiso et al., 1998), and signal interest and engagement (Gustafson & Neiberg, 2010). Thus, one can imagine that, in conversations with different goals (i.e., internal demands) that require communication of different types of information, speech patterns will vary. Furthermore, speech behaviors are often impacted by external demands, such as the conversational task and the accompanying motoric, linguistic, social, and cognitive demands placed upon a speaker (Kleinow & Smith, 2006; Lively et al., 1993; Tasko & McClean, 2004). Accordingly, when these demands differ, speech behaviors may differ as well. Thus, there is reason to believe that the interpersonal synergy model of spoken dialogue can be applied beyond linguistic behavior to speech behaviors. However, this hypothesis has never, to our knowledge, been systematically investigated.

Acoustic data from speech signal features (e.g., speech rate, pitch) may vary by context in two distinct ways (Diamond & Jefferies, 2001). First, the central tendency or average acoustic feature value tells us the typical level of a behavior in a given context and is the property most commonly studied in quantitative studies of speech research. Second, although often overlooked in speech research, the variability of acoustic feature values also carries information. Two conversations with the same typical feature values can differ meaningfully in their variability. This variability indicates the degree to which individuals modulate their behaviors from their typical level. Thus, a holistic examination of how speech varies by conversational context should account for both the central tendency and variability within the speech signal.

The purpose of this study is to explore the degree to which the interpersonal synergy model applies to speech behaviors. More specifically, we test the hypothesis that interlocutors adjust their speech behaviors as they attempt to meet the contextual demands of different types of conversational tasks with different conversational goals. To test this hypothesis, we compare the properties (i.e., central tendency and variability) of three acoustic features (i.e., pitch, speech rate, and articulatory precision) across two different types of conversations. Acoustic features were selected to represent a diverse set of speech signal properties across phonatory (i.e., pitch), rhythmic (speech rate), and articulatory (i.e., articulatory precision) dimensions. Conversation tasks were selected based on the work of Yeomans et al. (2022), who suggest that conversations can be broadly classified across two dimensions. First, the informational dimension focuses on the extent to which conversations are used to give or receive accurate information. We elicit conversations with high-informational intent (termed herein informational conversations) using the Diapix task, a cooperative game in which participants must give and receive information to successfully complete the task. Second, the relational dimension focuses on the extent to which conversations focus on establishing rapport or building a relationship. We elicit conversations

Downloaded from: https://pubs.asha.org Stephanie Borrie on 05/13/2024, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

with high-relational intent (termed herein relational conversations) by instructing participants to get to know each other through asking and answering questions about themselves. To appropriately account for potential confounding factors and adjust for personal differences, we employ a repeated-measures paradigm in which participants engage in each type of conversation with two different partners. While not a primary question of the study, this affords the opportunity to explore the degree to which conversing with different partners impacts speech behaviors as well.

Method

This study makes use of a corpus of 112 audio-recorded conversations and was carried out with ethical approval from the institutional review board at Arizona State University. In this corpus, 28 participants engaged in two different types of conversations (i.e., relational and informational) with two partners. Thus, each participant had four conversations, providing intraindividual information on typical behavior and variability of behaviors both within and across conversations. The order in which participants engaged in conversations with partners was counterbalanced across the data collection process. More details are provided below.

Participants

Participants consisted of 28 neurotypical women between the ages of 18 and 29 years. All participants were proficient speakers of American English as a first-learned language. Additionally, participants had no self-reported speech, language, or hearing deficits. All participants were blinded to the specific purpose of the study.

Conversation Partners

The two partners were also neurotypical women. The first partner (i.e., Partner 1) was 23 years old. The second partner (i.e., Partner 2) was 27 years old. Both partners were native speakers of American English with no self-reported speech, language, or hearing deficits. Additionally, both were blinded to the specific purpose of the study.

Procedure

For each conversation, one participant and one partner were seated facing one another. A microphone (Shure SM58) was positioned in front of each participant, and a second microphone was positioned in front of their partner at a mouth-to-microphone distance of 30 cm. Both microphones were connected to a digital recorder (Tascam DR-40). Separate audio channels for each interlocutor and standard settings (48 kHz; 16-bit sampling rate) were employed for recording. Conversational dyads then engaged in two different conversations: one relational conversation and one informational conversation.

Relational Conversations

Similar to previous research (i.e., Borrie et al., 2020; Dideriksen, 2023; Trujillo et al., 2023), relational conversations were elicited using open-ended conversation prompts. Participants were told that the goal of this conversation was to get to know one another. To do this, dyads were given a list of conversation topics such as food, hobbies, travel, and entertainment and were asked to select one of the topics and engage in a conversational exchange about the given topic. Conversations were terminated after 7 min.

Informational Conversations

Informational conversations were elicited using the Diapix task (Van Engen et al., 2010). Participants were told that the goal of this conversation was to successfully complete the Diapix task, a collaborative spot-the-difference game commonly used in speech research (e.g., Borrie et al., 2019; Kim et al., 2011; Wynn et al., 2023). In the Diapix task, each interlocutor is given one of a set of pictures. Pictures are virtually identical; however, in each set, there are 10 small details that differed between pictures. Dyads are told that their task is to talk to one another to find all 10 differences as quickly as possible without looking at each other's pictures. When introducing the Diapix task, no other instructions were given, and dyads were free to verbally interact as they wished to achieve the task. Conversations were terminated when dyads had found all 10 differences. Thus, while conversation length was variable, all conversations lasted at least 3 min.

Data Analysis

Given the variable length of the informational conversations and in order to control for differences between lengths of relational and informational conversations, only the first 3 min of each conversation were analyzed as has been done in previous research (Wynn, Barrett, & Borrie, 2022).[2] Trained research assistants manually coded each audio file, annotating and transcribing individual conversational turns using the Praat textgrid function (Boersma & Weenink, 2020). Conversational turns were defined as units of speech by a single interlocutor (i.e., participant or partner) that were free from pauses greater than 500 ms. To improve the accuracy of acoustic measures, conversational turns that were one word long were removed from analysis.

[2]Overall findings were comparable when full conversations were analyzed.

Downloaded from: https://pubs.asha.org Stephanie Borrie on 05/13/2024, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

For each conversational turn, we extracted data for three acoustic features: pitch, articulatory precision, and speech rate. These features were selected to represent different dimensions of the speech signal (i.e., phonatory, articulatory, and rhythmic dimensions).

Pitch

Pitch values were extracted using an automated Praat script. Using this script, fundamental frequency is extracted using a 5-ms time step and default parameters for pitch floor (75 Hz), pitch ceiling (600 Hz), silence threshold (0.03), and voicing threshold (0.45). Values are then averaged for each individual conversational turn.

Speech Rate

Speech rate was calculated using a syllabification script from the Penn Phonetics Toolkit (Tauberer, 2008). Previous research has shown a high reliability between the automatic script and a hand-coded speech rate calculation (i.e., Pearson's correlation score of .99; Wynn, Barrett, & Borrie, 2022). This script relies on a prespecified list of vowels that are used to identify syllable nuclei. Onsets and codas for each syllable are then determined using a series of linguistic rules. Speaking rate is calculated for each conversational turn by dividing the number of syllables by the turn length (in seconds).

Articulatory Precision

Articulatory precision values were extracted using an automated measure introduced by Tu et al. (2018) and subsequently validated and used by other researchers (Borrie et al., 2020, 2022; Stegmann et al., 2020; Wynn, Josephson, & Borrie, 2022). In this method, spectral acoustics of each observed phone are compared to the spectral acoustics of the target phone of a normative speech sample. Target phones are generated from the LibriCoprus, a publicly available corpus of 1,000 hr of read speech. Acoustic feature values represent the likelihood score between the observed and target phone. Perfect alignment between observed and target phones results in a score of 0. The smaller the score (i.e., the farther from 0), the greater the difference between observed and target phones, indicating more imprecise articulation. Scores for each phone are then averaged for each individual conversational turn.

Statistical Analyses

After acoustic feature values were extracted for each conversational turn, we calculated two summary statistics for each conversation. First, we used the mean to represent the average acoustic feature value of each conversation (i.e., the typical level of the acoustic feature). Second, we used the interquartile range (IQR) to represent the variability of acoustic feature values within each conversation. IQR was selected as our measure of variability, because (a) it makes no assumptions about the shape of the distribution and (b) it is interpretable (i.e., it represents the middle 50% of the observed feature values).

A series of linear mixed-effects models were used to investigate the effects of conversation task (relational vs. informational) and partner (Partner 1 vs. Partner 2) on acoustic feature mean and IQR scores while controlling for the lack of independence in data due to repeated conversations for each participant. All analyses were performed in the R statistical environment (R Version 4.2.1; R Development Core Team, 2022). Data cleaning and visualization relied on the tidyverse package (Wickham et al., 2019). Statistical analyses relied on the lme4, lmerTest, ggplot2, and emmeans packages (Bates et al., 2015; Kuznetsova et al., 2017; Lenth, 2023; Wickham, 2016). All p values reported in conjunction with estimates of effect are based on Satterwaithe approximation to degrees of freedom. All statistical output can be viewed on https://osf.io/cpx2y/.

Results

Descriptive Statistics

Analysis consisted of six acoustic feature measures: pitch mean (M = 218.08, SD = 19.83), pitch IQR (M = 23.85, SD = 11.22), speech rate mean (M = 4.23, SD = 0.48), speech rate IQR (M = 1.54, SD = 0.46), articulatory precision mean (M = −1.21, SD = 0.32), and articulatory precision IQR (M = 0.95, SD = 0.33). Acoustic feature measures were extracted from a total of 6,755 conversational turns across 112 conversations (28 participants with four conversations each). Descriptive data including means and standard deviations for the number of analyzed turns per conversation and the average duration of turns across conversation task and partner are shown in Table 1.

Primary Analysis

Linear mixed-effects models were used to analyze the six dependent variables (i.e., each acoustic feature measure). For each dependent variable, one model examined the main effects of conversation task and partner, and a second model examined the interaction between conversation task and partner. Additionally, each model included a random intercept by participant ID to account for the repeated measures. Figure 1 depicts each speech feature measure across conversational task (i.e., informational vs. relational) and partner (i.e., Partner 1 and Partner 2). Specific findings, organized by speech feature, are detailed below (see Figure 1).

Downloaded from: https://pubs.asha.org Stephanie Borrie on 05/13/2024, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

Table 1. Descriptive statistics for conversational turns.

Variable	*M* (*SD*)			
	Function		Partner	
	Informational	Relational	Confederate 1	Confederate 2
Conversational turns	31.11 (4.69)	26.18 (4.91)	28.39 (5.70)	28.89 (5.09)
Turn duration (in seconds)	2.16 (1.47)	3.37 (3.21)	2.82 (2.68)	2.60 (2.30)

Pitch

Mean

Linear mixed-effects models revealed a significant main effect of task on mean pitch ($b = -9.12$; $p < .001$). Participants spoke with higher pitch in relational conversations than informational conversations. There was also a significant main effect of partner on mean pitch ($b = -3.67$; $p = .047$). Participants spoke with higher pitch in conversations with Partner 1 than with Partner 2. There was no significant interaction between task and partner on mean pitch ($b = -2.13$; $p = .562$).

IQR

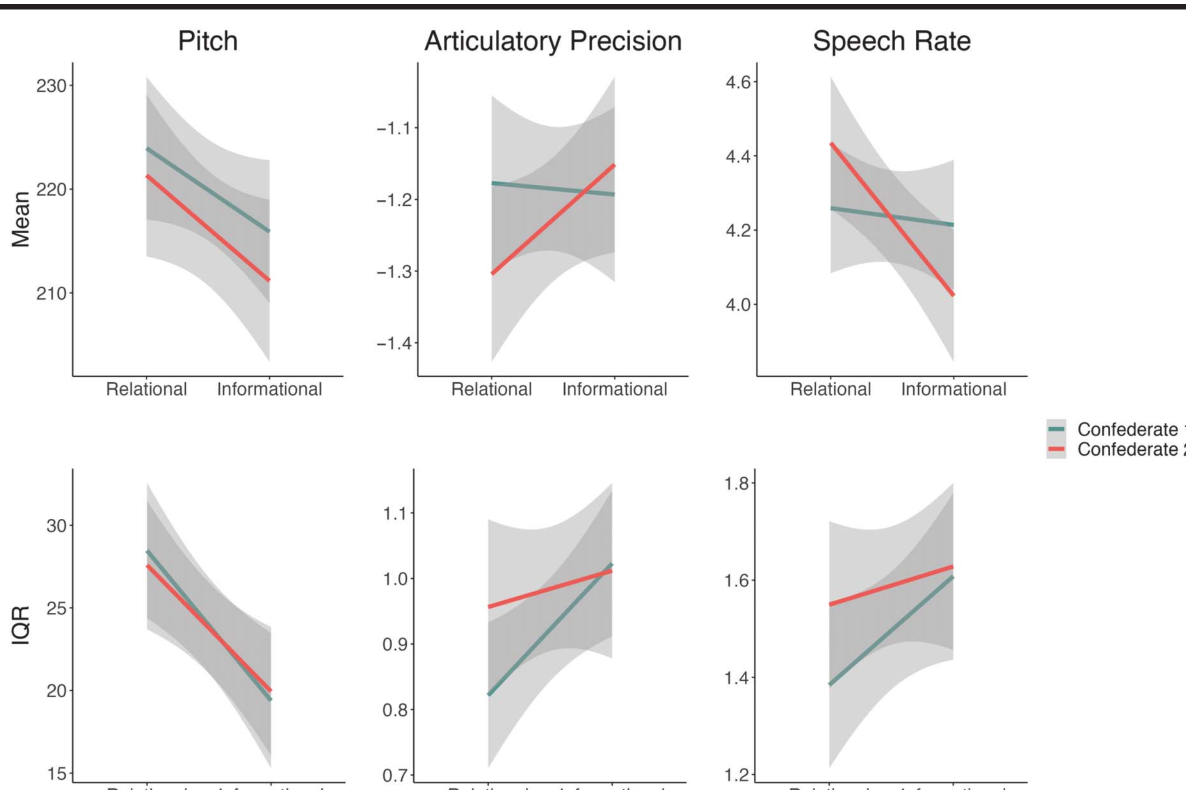
There was a significant main effect of task on pitch IQR ($b = -8.37$; $p < .001$). Participants spoke with more variability in pitch in relational than informational conversations. There was no significant main effect of partner on pitch IQR ($b = -0.16$; $p = .922$). Additionally, there was no significant interaction between task and partner on pitch IQR ($b = 1.4$; $p = .656$).

Speech Rate

Mean

There was a significant main effect of task on mean speech rate ($b = -0.22$; $p < .001$) and no significant main effect of partner on mean speech rate ($b = -0.01$; $p = .890$). However, there was a significant interaction between task and partner on speech rate ($b = -0.37$; $p < .001$). Specifically, while participants spoke faster in relational than informational conversations ($b = 0.41$; $p \leq .001$) with Partner 2, this difference was not present when participants spoke with Partner 1 ($b = 0.05$; $p = .53$).

Figure 1. Speech feature measures by conversational task and partner. IQR = interquartile range.



Downloaded from: https://pubs.asha.org Stephanie Borrie on 05/13/2024, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

IQR

There was a significant main effect of task on speech rate IQR ($b = 0.15$; $p = .030$). Participants spoke with less speech rate variability in relational conversations than informational conversations. There was no significant main effect of partner on speech rate IQR ($b = 0.09$; $p = .18$). Additionally, there was no significant interaction between task and partner on speech rate IQR ($b = -0.15$; $p = .29$).

Articulatory Precision

Mean

There was no significant main effect of task on mean articulatory precision ($b = 0.07$; $p = .139$). There was also no significant main effect of partner on mean articulatory precision ($b = -0.04$; $p = .355$). Furthermore, there was no significant interaction between task and partner on mean articulatory precision ($b = 0.17$; $p = .065$).

IQR

There was a significant main effect of task on articulatory precision IQR ($b = 0.13$; $p = .018$). Participants spoke with less variability in articulatory precision in relational conversations than informational conversations. There was no significant main effect of partner on articulatory precision IQR ($b = 0.06$; $p = .247$). Additionally, there was no significant interaction between task and partner on articulatory precision IQR ($b = -0.15$; $p = .171$).

Discussion

The purpose of this study was to test the hypothesis that interlocutors adjust their speech behaviors across different conversational tasks with different goals. To do this, we examined the central tendency (i.e., mean) and variability (i.e., IQR) of pitch, speech rate, and articulatory precision across informational and relational conversations. As a secondary objective, we also examined if these speech behaviors differed when conversing with different conversational partners and if there was an interaction between conversational task and partner. Specific findings are discussed below.

First, our findings showed significant differences between informational and relational conversations for almost every speech feature measure (i.e., all feature measures except mean articulatory precision). Furthermore, in only one case (i.e., mean speech rate) was there an interaction between conversation task and partner, indicating that these findings were generally robust across conversations with both partners. Thus, our findings provide empirical support for the interpersonal synergy model of spoken dialogue, which postulates that communicative behaviors are used as interlocutors attempt to facilitate successful interactions and achieve communicative goals (Fusaroli et al., 2014). While this study did not explore the mechanisms driving context-dependent adaptations, we offer some postulations regarding why specific changes may have occurred. In some instances, differences may have been a direct result of specific strategies used (either consciously or subconsciously) as interlocutors attempted to meet their internal demands (i.e., accomplish specific goals). For example, research has shown that higher pitch and greater pitch variations are used in conversations to express higher arousal emotional states (Bänziger & Scherer, 2005). Thus, interlocutors may have exhibited higher pitch with more variation in relational conversations where high-arousal emotions could help convey engagement and enthusiasm about what their partner was saying—something less important in informational conversations. Additional research has found that speakers vary their speech rate depending on whether they are presenting new information or continuing to discuss old piece of information (Koiso et al., 1998). Thus, speech rate variation may have been higher in informational conversations because of frequent shifts between presenting new information (i.e., describing new parts of the picture) and discussing old information (i.e., providing additional details or following up about what was previously described) and reliance on this strategy to achieve mutual understanding. Contrastingly, in relational conversations, the introduction of new information may occur less often and signaling a transition may be less important than in informational conversations. Therefore, speech rate variability may have played a less central role. In other instances, speech adaptations may have been influenced by external demands of the conversation, such as the cognitive, linguistic, social, and motoric demands associated with tasks that vary in many different ways (e.g., difficulty, source and level of motivation, time pressures; see Beechey et al., 2019, for a discussion of the role of task effects on conversation). For instance, research has shown that higher cognitive load leads to slower speech rate (Gorovoy et al., 2010; Mitchel et al., 1996). Although the task used to elicit the informational conversations is relatively simple, the collaborative nature of the task coupled with the requirement to solve the problem (i.e., identify all 10 differences) as quickly as possible likely led to a higher cognitive load and subsequent slower speech rate than that observed in relational conversations. Furthermore, a shift between speaking slowly when visually inspecting the Diapix scene and speeding up when not as readily focused on cognitive processing may be partially responsible for the greater levels of speech rate variation seen in the informational task.

Downloaded from: https://pubs.asha.org Stephanie Borrie on 05/13/2024, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

Although not the primary aim of this study, we also examined the role of conversational partner on speech behaviors. We found that only one speech behavior (i.e., mean pitch) varied by partner. Thus, these partner-dependent differences occurred less consistently and to a smaller degree than task-dependent differences. Such findings make good sense. In this study, the contextual tasks (and associated goals) differed substantially. Therefore, the speech behaviors employed also differed substantially, presumably to meet differing contextual demands. Contrastingly, the two conversational partners were relatively similar—both partners were young neurotypical women who were proficient speakers of American English as a first-learned language. As such, context-dependent speech adaptations in relation to partner in the current study were likely to be minimal. However, if partners differ in ways that may be consequential to conversation outcomes, we would expect speech behavior to differ more substantially. Indeed, there is existing empirical evidence that supports this postulation. Studies have shown that interlocutors adapt their speech depending on the age (Kemper, 1994; Shaw & Gordon, 2021), native language (Lee & Baese-Berk, 2020), and neurotype (Borrie et al., 2020; Lubold et al., 2021) of their conversational partner. This is often done in an effort to compensate for limitations (either real or imagined) in their partner's communication abilities. For example, interlocutors often hyperarticulate speech when conversing with someone with a traumatic brain injury relative to a neurotypical control (Borrie et al., 2020). Similarly, interlocutors with normal hearing may modify their vocal level and speech spectrum depending on the degree of hearing impairment exhibited by their conversational partner (Beechey et al., 2020). Thus, our current findings, in conjunction with previous literature, provide robust empirical support for the interpersonal synergies model of spoken dialogue in which context shapes behavior in meaningful ways (Fusaroli et al., 2014).

Clinical Implications

Our findings afford important clinical implications. Assessment of conversational speech is one of the most important components of evaluations for a variety of speech disorders such as motor speech disorders (Duffy, 2019), speech sound disorders (Morrison & Shriberg, 1992), and stuttering (Guitar, 2019). However, although monologic speech is often evaluated in multiple tasks (e.g., picture description, reading passages, narratives) during evaluations, if conversational speech is included in the evaluation, it is often only done within a single task. Although the current study involved individuals without communication disorders, the results suggest that gaining an accurate and holistic picture of an individual's disorder may require assessing speech in multiple types of conversations. Such information offers many benefits to assessment. First, utilizing different conversational tasks may offer insights into specific irregularities or challenges. For example, excessive pitch variability sometimes exhibited in the speech of autistic individuals (e.g., Filipe et al., 2014) may not be apparent in relational conversations where higher degrees of pitch variation are expected, but may be more noticeable in informational conversations where such variation is less frequent. Assessing speech across multiple conversational tasks can also be used to provide information about the degree to which an individual is able to use specific speech strategies to meet their conversational goals. For example, interlocutors in our study employed more speech rate variation in informational conversations (possibly due to the need to signal when new information was being introduced). However, people with apraxia of speech, who frequently present with limitations in their ability to vary their speech rate (Duffy, 2019), may employ similar speech rate patterns across conversational tasks. Thus, different strategies may need to be adopted to ensure that conversational goals are not left unmet. Finally, it is likely that individuals with different disorders will respond differently across tasks, and such information may aid differential diagnosis. For example, an individual with aphasia may show substantial changes between their speech in a simple relational conversation with low-linguistic demands versus a complex informational conversation where the linguistic demands are high. Contrastingly, we may see very little difference in speech behaviors in these two conversation tasks for someone with dysarthria whose motoric deficits can restrict the potential for speech variation (Duffy, 2019).

Limitations and Future Directions

In the current study, we conducted an initial examination to investigate whether contextual demands modulate the speech behaviors of interlocutors in dyadic conversations. Findings offer several potential avenues for continued examination. First, research should continue to explore the ways in which different contextual factors influence speech behaviors. For instance, although we looked at differences across two relatively disparate conversational tasks with differing conversational goals, examining patterns in other types of conversations (i.e., deeper conversations, conversations during competitive tasks) would provide further information about how contextual demands drive changes in speech behavior. Future research could also examine the degree to which speech behaviors change when speaking to a more diverse set of conversational partners. Additional factors such as a conversation location and timing may be important to examine as well. Next, participants from this study came from a relatively homogenous sample. Although we conjecture

Downloaded from: https://pubs.asha.org Stephanie Borrie on 05/13/2024, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

that conversational speech behaviors will always be impacted by context, specific patterns will likely vary across gender, age, culture, language, and neurotype. Thus, future research should explore these patterns across more heterogenous samples to understand how individual factors moderate context-dependent behavior change. Additionally, although our research focused on the ways in which speech is modified within the context of different tasks and goals, we did not explore if and how speech adaptations actually lead to the realizations of these goals. It is possible that some modifications do promote goal attainment, while others, while well intentioned, are maladaptive. Therefore, examining which speech features are most amenable to context and the degree to which these changes alter conversational outcomes would be beneficial in making assessment and treatment decisions.

Conclusions

In summary, this study investigated the role of contextual demands on the acoustic properties of conversational speech. We found significant differences in pitch, speech rate, and articulatory precision of interlocutors engaged in informational relative to relational conversations. Thus, we provide evidence that conversational speech behavior is context dependent and highlights the need to consider the role of context in assessments of individuals with communication disorders.

Data Availability Statement

Data analysis code and model outputs associated with this work are available at the study repository hosted at https://osf.io/cpx2y/.

Acknowledgments

This research was supported by National Institute on Deafness and Other Communication Disorders Grant R01DC020713 (awarded to Stephanie A. Borrie). The authors gratefully acknowledge research assistants in the Human Interaction Lab at Utah State University for assistance with data analysis.

References

Bänziger, T., & Scherer, K. R. (2005). The role of intonation in emotional expressions. *Speech Communication, 46*(3–4), 252–267. https://doi.org/10.1016/j.specom.2005.02.016

Barrett, L. F., Mesquita, B., & Gendron, M. (2011). Context in emotion perception. *Current Directions in Psychological Science, 20*(5), 286–290. https://doi.org/10.1177/0963721411422522

Bates, D., Mächler, 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

Beechey, T., Buchholz, J. M., & Keidser, G. (2019). Eliciting naturalistic conversations: A method for assessing communication ability, subjective experience, and the impacts of noise and hearing impairment. *Journal of Speech, Language, and Hearing Research, 62*(2), 470–484. https://doi.org/10.1044/2018_JSLHR-H-18-0107

Beechey, T., Buchholz, J. M., & Keidser, G. (2020). Hearing impairment increases communication effort during conversations in noise. *Journal of Speech, Language, and Hearing Research, 63*(1), 305–320. https://doi.org/10.1044/2019_JSLHR-19-00201

Bentum, M., ten Bosch, L., van den Bosch, A., & Ernestus, M. (2019). Do speech registers differ in the predictability of words? *Linguistics, 24*(1), 98–130. https://doi.org/10.1075/ijcl.17062.ben

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

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., Wynn, C. J., Berisha, V., & Barrett, T. S. (2022). From speech acoustics to communicative participation in dysarthria: Towards a causal framework. *Journal of Speech, Language, and Hearing Research, 65,* 405–418. https://doi.org/10.1044/2021_JSLHR-21-00306

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

Carl, M., Icht, M., & Ben-David, B. M. (2022). A cross-linguistic validation of the test for rating emotions in speech: Acoustic analyses of emotional sentences in English, German, and Hebrew. *Journal of Speech, Language, and Hearing Research, 65*(3), 991–1000. https://doi.org/10.1044/2021_JSLHR-21-00205

Chun, M. M. (2000). Contextual cueing of visual attention. *Trends in Cognitive Sciences, 4*(5), 170–178. https://doi.org/10.1016/s1364-6613(00)01476-5

Crain-Thoreson, C., Dahlin, M. P., & Powell, T. A. (2001). Parent-child interaction in three conversational contexts: Variations in style and strategy. *New Directions for Child and Adolescent Development, 2001*(92), 23–38. https://doi.org/10.1002/cd.13

de Ruiter, L. E. (2015). Information status marking in spontaneous vs. read speech in story-telling tasks – Evidence from intonation analysis using GToBI. *Journal of Phonetics, 48,* 29–44. https://doi.org/10.1016/j.wocn.2014.10.008

Diamond, I., & Jefferies, J. (2001). *Beginning statistics: An introduction for social scientists.* SAGE. https://doi.org/10.4135/9781446249437

Dideriksen, C., Christiansen, M. H., Tylén, K., Dingemanse, M., & Fusaroli, R. (2023). Quantifying the interplay of conversational devices in building mutual understanding. *Journal of Experimental Psychology: General, 152*(3), 864–889. https://doi.org/10.1037/xge0001301

Duffy, J. R. (2019). Motor speech disorders: Substrates, differential diagnosis, and management. Elsevier.

Downloaded from: https://pubs.asha.org Stephanie Borrie on 05/13/2024, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

Filipe, M. G., Frota, S., Castro, S. L., & Vicente, S. G. (2014). Atypical prosody in Asperger syndrome: Perceptual and acoustic measurements. *Journal of Autism and Developmental Disorders, 44*(8), 1972–1981.

Furui, S., Nakamura, M., Ichiba, T., & Iwano, K. (2005). Analysis and recognition of spontaneous speech using Corpus of Spontaneous Japanese. *Speech Communication, 47*(1), 208–219. https://doi.org/10.1016/j.specom.2005.02.010

Fusaroli, R., Rączaszek-Leonardi, J., & Tylén, K. (2014). Dialog as interpersonal synergy. *New Ideas in Psychology, 32,* 147–157. https://doi.org/10.1016/j.newideapsych.2013.03.005

Gorovoy, K., Tung, J., & Poupart, P. (2010). Automatic speech feature extraction for cognitive load classification. *CMBES Proceedings, 33*(1). https://proceedings.cmbes.ca/index.php/proceedings/article/view/538

Guitar, B. (2019). *Stuttering: An integrated approach to its nature and treatment* (5th ed.). LWW.

Gustafson, J., & Neiberg, D. (2010). *Prosodic cues to engagement in non-lexical response tokens in Swedish*. DiSS-LPSS Joint Workshop 2010, University of Tokyo, Japan, September 25–26, 2010. https://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-52148

Haden, C. A., & Fivush, R. (1996). Contextual variation in maternal conservational styles. *Merrill-Palmer Quarterly, 42*(2), 200–227.

Hoff-Ginsberg, E. (1991). Mother-child conversation in different social classes and communicative settings. *Child Development, 62*(4), 782–796. https://doi.org/10.1111/j.1467-8624.1991.tb01569.x

Howell, P., & Kadi-Hanifi, K. (1991). Comparison of prosodic properties between read and spontaneous speech material. *Speech Communication, 10*(2), 163–169. https://doi.org/10.1016/0167-6393(91)90039-V

Hu, B., Dixon, P. C., Jacobs, J. V., Dennerlein, J. T., & Schiffman, J. M. (2018). Machine learning algorithms based on signals from a single wearable inertial sensor can detect surface- and age-related differences in walking. *Journal of Biomechanics, 71,* 37–42. https://doi.org/10.1016/j.jbiomech.2018.01.005

Ishida, R., Palmer, J. B., & Hiiemae, K. M. (2002). Hyoid motion during swallowing: Factors affecting forward and upward displacement. *Dysphagia, 17*(4), 262–272. https://doi.org/10.1007/s00455-002-0064-5

Kemper, S. (1994). Elderspeak: Speech accommodations to older adults. *Aging, Neuropsychology, and Cognition, 1*(1), 17–28. https://doi.org/10.1080/09289919408251447

Kim, M., Horton, W. S., & Bradlow, A. R. (2011). Phonetic convergence in spontaneous conversations as a function of interlocutor language distance. *Laboratory Phonology, 2*(1). https://doi.org/10.1515/labphon.2011.004

Kleinow, J., & Smith, A. (2006). Potential interactions among linguistic, autonomic, and motor factors in speech. *Developmental Psychobiology, 48*(4), 275–287. https://doi.org/10.1002/dev.20141

Koiso, H., Shimojima, A., & Katagiri, Y. (1998). Collaborative signaling of informational structures by dynamic speech rate. *Language and Speech, 41*(3–4), 323–350. https://doi.org/10.1177/002383099804100405

Kuntapun, J., Silsupadol, P., Kamnardsiri, T., & Lugade, V. (2020). Smartphone monitoring of gait and balance during irregular surface walking and obstacle crossing. *Frontiers in Sports and Active Living, 2.* https://doi.org/10.3389/fspor.2020.560577

Kuznetsova, A., Brockhoff, P. B., Christensen, R. H. B. (2017). lmer Test package: Tests in linear mixed effects models. *Journal of Statistical Software, 82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Lee, D.-Y., & Baese-Berk, M. M. (2020). The maintenance of clear speech in naturalistic conversations. *The Journal of the Acoustical Society of America, 147*(5), 3702–3711. https://doi.org/10.1121/10.0001315

Lenth, R. (2023). *emmeans: Estimated marginal means, aka least-squares means*. R package version 1.8.5. https://CRAN.R-project.org/package=emmeans

Lindblom, B. (1990). Explaining phonetic variation: A sketch of the H&H theory. In W. J. Hardcastle & A. Marchal (Eds.), *Speech production and speech modelling* (pp. 403–439). Springer. https://doi.org/10.1007/978-94-009-2037-8_16

Lively, S. E., Pisoni, D. B., Summers, W. V., & Bernacki, R. H. (1993). Effects of cognitive workload on speech production: Acoustic analyses and perceptual consequences. *Journal of the Acoustical Society of America, 93*(5), 2962–2973. https://doi.org/10.1121/1.405815

Lubold, N., Willi, M., Borrie, S., Barrett, T., & Berisha, V. (2021). Healthy communication partners modify their speech when conversing with individuals with Parkinson's disease. *Journal of Speech, Language, and Hearing Research, 64*(5), 1539–1549. https://doi.org/10.1044/2021_JSLHR-20-00233

Mellers, B., Schwartz, A., & Cooke, A. (1998). Judgment and decision-making. *Annual Review of Psychology, 49*(1), 447–477. https://doi.org/10.1146/annurev.psych.49.1.447

Mitchel, H. L., Hoit, J. D., & Watson, P. J. (1996). Cognitive-linguistic demands and speech breathing. *Journal of Speech and Hearing Research, 39*(1), 93–104. https://doi.org/10.1044/jshr.3901.93

Morrison, J. A., & Shriberg, L. D. (1992). Articulation testing versus conversational speech sampling. *Journal of Speech and Hearing Research, 35*(2), 259–273. https://doi.org/10.1044/jshr.3502.259

Nakamura, M., Iwano, K., & Furui, S. (2008). Differences between acoustic characteristics of spontaneous and read speech and their effects on speech recognition performance. *Computer Speech & Language, 22*(2), 171–184. https://doi.org/10.1016/j.csl.2007.07.003

R Development Core Team. (2022). R: A language and environment for statistical computing. https://www.R-project.org/

Reitter, D., Moore, J. D., & Keller, F. (2006). Priming of syntactic rules in task-oriented dialogue and spontaneous conversation. *Proceedings of the 28th Annual Conference of the Cognitive Science Society*, Article 6.

Shaw, C. A., & Gordon, J. K. (2021). Understanding Elderspeak: An evolutionary concept analysis. *Innovation in Aging, 5*(3), Article igab023. https://doi.org/10.1093/geroni/igab023

Silverman, K., Blaauw, E., Spitz, J., & Pitrelli, J. F. (1992). A prosodic comparison of spontaneous speech and read speech. *Proceedings 2nd International Conference on Spoken Language Processing (ICSLP 1992),* 1299–1302. https://doi.org/10.21437/ICSLP.1992-349

Steele, C. M., Alsanei, W. A., Ayanikalath, S., Barbon, C. E. A., Chen, J., Cichero, J. A. Y., Coutts, K., Dantas, R. O., Duivestein, J., Giosa, L., Hanson, B., Lam, P., Lecko, C., Leigh, C., Nagy, A., Namasivayam, A. M., Nascimento, W. V., Odendaal, I., Smith, C. H., & Wang, H. (2015). The influence of food texture and liquid consistency modification on swallowing physiology and function: A systematic review. *Dysphagia, 30*(1), 2–26. https://doi.org/10.1007/s00455-014-9578-x

Stegmann, G. M., Hahn, S., Liss, J., Shefner, J., Rutkove, S., Shelton, K., Duncan, C. J., & Berisha, V. (2020). Early detection and tracking of bulbar changes in ALS via frequent and remote speech analysis. *NPJ Digital Medicine, 3,* Article 132. https://doi.org/10.1038/s41746-020-00335-x

Tasko, S. M., & McClean, M. D. (2004). Variations in articulatory movement with changes in speech task. *Journal of*

Downloaded from: https://pubs.asha.org Stephanie Borrie on 05/13/2024, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

Speech, Language, and Hearing Research, 47(1), 85–100. https://doi.org/10.1044/1092-4388(2004/008)

Tauberer, J. (2008). *P2TK automated syllabifier*. https://sourceforge.net/p/p2tk/code/HEAD/tree/python/syllabify/

Todorović, D. (2010). Context effects in visual perception and their explanations. *Review of Psychology, 17*(1), 17–32.

Trujillo, J. P., Dideriksen, C., Tylén, K., Christiansen, M. H., & Fusaroli, R. (2023). The dynamic interplay of kinetic and linguistic coordination in Danish and Norwegian conversation. *Cognitive Science, 47*(6), Article e13298. https://doi.org/10.1111/cogs.13298

Tu, M., Grabek, A., Liss, J., & Berisha, V. (2018). Investigating the role of L1 in automatic pronunciation evaluation of L2 speech. In *Proceedings of INTERSPEECH Conference 2018* (pp. 1636–1640). Hyderabad, India. https://arxiv.org/pdf/1807.01738.pdf [PDF]

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

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag. https://doi.org/10.1007/978-3-319-24277-4

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., . . . Yutani, H. (2019). Welcome to the Tidyverse. *Journal of Open Source Software, 4*(43), Article 1686. https://doi.org/10.21105/joss.01686

Wiggers, P., & Rothkrantz, L. J. M. (2007). Exploratory analysis of word use and sentence length in the spoken Dutch corpus. In V. Matoušek & P. Mautner (Eds.), *Text, speech and dialogue* (pp. 366–373). Springer. https://doi.org/10.1007/978-3-540-74628-7_48

Wynn, C. J., Barrett, T. S., Berisha, V., Liss, J. M., & Borrie, S. A. (2023). Speech entrainment in adolescent conversations: A developmental perspective. *Journal of Speech, Language, and Hearing Research, 66*(8S), 3132–3150. https://doi.org/10.1044/2023_JSLHR-22-00263

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

Wynn, C. J., Josephson, E. R., & Borrie, S. A. (2022). An examination of articulatory precision in autistic children and adults. *Journal of Speech, Language, and Hearing Research, 65*(4), 1416–1425. https://doi.org/10.1044/2021_JSLHR-21-00490

Yeomans, M., Schweitzer, M. E., & Brooks, A. W. (2022). The Conversational Circumplex: Identifying, prioritizing, and pursuing informational and relational motives in conversation. *Current Opinion in Psychology, 44,* 293–302. https://doi.org/10.1016/j.copsyc.2021.10.001

Downloaded from: https://pubs.asha.org Stephanie Borrie on 05/13/2024, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions