ASHA American Speech-Language-Hearing Association

JSLHR

Research Note

# Acoustic Correlates of Timing Typicality in Speakers With Parkinson's Disease

**Saul A. Frankford**[a,b,c] and **Cara E. Stepp**[a,d]

[a]Department of Speech, Language, and Hearing Sciences, Boston University, MA [b]Department of Speech, Language, and Hearing, The University of Texas at Dallas, Richardson [c]Callier Center for Communication Disorders, The University of Texas at Dallas, Richardson [d]Department of Biomedical Engineering, Boston University, MA

ARTICLE INFO

Article History:
Received October 9, 2024
Revision received March 21, 2025
Accepted May 1, 2025

Editor-in-Chief: Rita R. Patel
Editor: Tara McAllister

https://doi.org/10.1044/2025_JSLHR-24-00712

ABSTRACT

**Purpose:** The present study aimed to determine acoustic metrics that can approximate listener perception of the typicality of speech timing in individuals with Parkinson's disease (PD). It was hypothesized that the perception of timing typicality would correlate with measures based on the deviation from speech produced by individuals with typical speech (in speaking duration, in pause time and other disfluencies, and at the word level).
**Method:** Twenty speakers with PD and 40 typical speakers matched in age and sex were recorded reading a standard passage. Acoustic timing measures were calculated for the speakers with PD, both absolute and relative to recordings from the typical speakers. Linear regression models were used to estimate the relationship strength between each acoustic measure and the perception of timing typicality. Models containing all variables and subsets of variables were compared to test study hypotheses.
**Results:** A model consisting of mean word duration and mean interword duration, both in their absolute values and relative to control speakers, explained substantial variance in perceptual judgments of timing typicality in speakers with PD ($R^2$ = .93).
**Conclusions:** Timing measures based on the deviation from normative values and accounting for pausing and disfluencies may provide an acoustic basis for estimating timing typicality in people with PD. Future work should examine these acoustic metrics in a larger sample to determine their utility in research and clinical settings.

Parkinson's disease (PD) is the second most prevalent neurodegenerative disorder (Tysnes & Storstein, 2017) and results in speech impairment (typically hypokinetic dysarthria) in up to 90% of individuals (Ho et al., 1999; Logemann et al., 1978). Among the various speech features associated with hypokinetic dysarthria are disruptions in the rate and rhythm of speech. Auditory–perceptual studies indicate that the speech of people with PD (PwPD) includes inappropriate silences and short rushes and is either faster, slower, or more variable than in typical speakers (Darley et al., 1969; Plowman-Prine et al., 2009), with notable variation across individuals (Ho et al., 1999). Furthermore, timing features have been shown to be correlated with ratings of intelligibility and naturalness (Darley et al., 1969; Frankford et al., 2024), important metrics in both research and clinical settings. Although speech timing disturbances in PwPD are most common in those with greater speech impairment (Ho et al., 1999), reports have identified temporal changes in the speech of individuals with mild dysarthria severity (Lowit et al., 2018; Tjaden & Wilding, 2011), indicating that disrupted speech timing is a characteristic feature of speech in PD.

Several acoustic speech timing features have been proposed to differentiate PwPD and speakers with no neurological disorder. Speaking rate and articulation rate (typically defined as speaking rate when pauses are removed) are the most widely reported measures; however,

Correspondence to Saul A. Frankford: saul.frankford@utdallas.edu.
***Disclosure:*** *The authors have declared that no competing financial or nonfinancial interests existed at the time of publication.*

results across studies are highly variable. Some studies show an increased rate in PwPD (Knowles et al., 2021; McRae et al., 2002; Skodda & Schlegel, 2008), whereas others show a decreased rate (Lowit et al., 2018; Martínez-Sánchez et al., 2016; Nishio & Niimi, 2001) or no differences from typical speakers (Chiu et al., 2021; Martens et al., 2015; Rusz et al., 2021; Walsh & Smith, 2012). Relatedly, it has been demonstrated that PwPD tend to increase speaking rate over the course of a reading passage (Skodda & Schlegel, 2008), similar to the festinating behaviors commonly observed in syllable repetition (Flasskamp et al., 2012; Skodda et al., 2010) and gait (Giladi et al., 2001; Moreau et al., 2007). Studies have also examined features that contribute to speaking rate such as pause time, where PwPD may show significantly increased pausing (Hammen & Yorkston, 1996; Logemann et al., 1978; Nishio & Niimi, 2001; Tjaden & Wilding, 2011), decreased pausing (Skodda & Schlegel, 2008), and/or atypical pausing (Darling-White & Huber, 2020; Hammen & Yorkston, 1996) compared to controls. In parallel with changes in speech rate and pause characteristics, several studies examining speech rhythm measures in PwPD have found that the percentage of speaking time taken up by vowels, the time between successive vowel onsets, and variability in consonant durations may also differentiate speakers with and without PD (Lowit et al., 2018; Maffia et al., 2021; Pettorino et al., 2016). Although these findings demonstrate several speech timing components that are potentially disrupted in PD, the biggest takeaway appears to be that speech timing changes in PD are complex, multifaceted, and variable across individuals.

Given this complexity, assessing speech timing differences across individuals or across several time points may be cumbersome, time-consuming, and/or incomplete. Distilling the factors comprising speech timing disturbances in PD into a single measure would allow for efficient speech timing evaluation for several purposes. Clinically, this would allow for tracking changes due to disease progression or in response to speech therapy. For research, it would be useful to be able to efficiently examine changes in speech timing across different speaking tasks, in speaking modes (e.g., different rates, "clear" vs. loud speech; Tjaden et al., 2013), or in response to different types of speech therapy. An efficient metric of speech timing would also facilitate the examination of the relationship between disrupted timing and other outcome measures such as intelligibility or naturalness. Our previous study (Frankford et al., 2024) introduced an auditory–perceptual construct, termed "timing typicality," intended to capture a listener's impression of how typical the timing of an individual's speech is, separate from other prosodic speech features such as loudness and pitch. This construct was meant to be an efficient way to distill the 10 atypical timing characteristics found in the speech of PwPD in previous work (Dimensions 22–31 in the Appendix of Darley et al., 1969) into one measure, particularly to determine the relationship between speech timing and intelligibility and naturalness. Listeners without formal training were able to provide reliable ratings of this measure with minimal instruction. Therefore, it provides an ideal auditory–perceptual metric for benchmarking potential acoustic correlates of timing disruption in the speech of PwPD.

The present study aimed to evaluate what acoustic metrics are related to listener perception of the typicality of speech timing in PwPD. In particular, we evaluated how well two kinds of timing measures estimated timing typicality: mean word duration (MWD; a measure related to articulation rate) and mean interword interval (MIWI; which characterizes pauses and disfluencies and, when combined with MWD, captures variation in speaking rate) across a standard reading passage. Both measures were evaluated on their own and relative to the average word and interword interval (IWI) durations of a cohort of control speakers, resulting in passage-relative MWD and MIWI metrics. We then evaluated whether extracting temporal deviation from control speakers at the word level (rather than the passage level) would significantly improve the estimation of timing typicality, constructing a word-relative MWD and MIWI. Because a listener's perception of timing typicality necessarily evaluates how a speaker's production deviates from a typical production, it was hypothesized that the deviation in measures relative to control speakers would improve estimation of timing typicality compared to the measures on their own. That is, speakers with both shorter MWDs and MIWIs (faster rate) or longer MWDs and MIWIs (slower rate) would be associated with lower timing typicality scores. It was also hypothesized that the IWI measures (MIWI, passage-relative MIWI, and word-relative MIWI) would significantly contribute to timing typicality above and beyond the duration measures, since they account for pause time and other disfluencies that affect the rhythm of speech. Finally, we hypothesized that measures of duration deviation at the word level (word-relative MWDs and word-relative MIWI) would better estimate timing typicality relative to the absolute and passage-level measures, given that average values over a passage do not capture the range of linguistically driven diversity in speech rhythm characteristics in speech with typical timing.

## Method

All participants (speakers and listeners) provided informed consent through the Boston University Institutional Review Board (No. 2625) or the University of

Washington Institutional Review Board (No. 36181), and all participants received compensation for their time.

## Speakers

Sixty speakers contributed speech recordings for this study. Twenty PwPD (10 males and 10 females[1]) ranged in age from 46 to 81 years ($M$ = 69.7, $SD$ = 7.7) and with times since diagnosis from 1 to 20 years ($M$ = 9.0, $SD$ = 5.8). These participants were previously included in Frankford et al. (2024) and were selected from a larger database of PwPD to typify a range of naturalness and intelligibility values that were not correlated. Briefly, this study examined the relationships between ratings of speech timing and articulatory precision and ratings of naturalness and intelligibility in PwPD. To ensure that intelligibility and naturalness were separable constructs, a certified speech-language pathologist with 8 years of experience listening to speech in PwPD made preliminary ratings of naturalness and intelligibility on speakers from the larger database. Samples were then chosen to approximately equally represent combinations of lower and higher intelligibility and naturalness. Ratings of intelligibility and naturalness from the final sample of inexperienced listeners demonstrated no significant correlation ($r$ = .28, $p$ = .24; see Frankford et al., 2024, for a complete description of the selection process). All speakers were recorded in the "ON" phase of their medication, as determined by self-report. The overall severity of the dysarthria of their speech was rated by the same speech-language pathologist on a visual analog scale (VAS) on which 0 was anchored as *no signs of dysarthria* and 100 was anchored as *very severe dysarthria*. Ratings of their speech ranged between 10 and 65. To provide a reference for the typical production of the reading passage, two neurotypical speakers who reported typical speech and language were matched by sex and age (within 2 years) to each speaker with PD. This procedure yielded a group of 40 speakers (20 males and 20 females) ranging in age from 46 to 81 years ($M$ = 69.5, $SD$ = 7.2).

All speakers read aloud the first six sentences of the "Rainbow Passage" (Fairbanks, 1960). Disfluencies were intentionally left in recordings to reflect the natural components of individuals' speech timing behaviors, and stimuli were amplitude normalized such that the root-mean-square of the amplitude signal was the same across recordings.

### Auditory–Perceptual Ratings

A set of 20 young inexperienced listeners (seven cisgender men, 10 cisgender women, two nonbinary people, and one agender person; nine males and 11 females) aged between 19 and 28 years ($M$ = 21.8, $SD$ = 2.4) rated the speech samples of the PwPD on *timing typicality* using a VAS from 0 to 100 as part of a larger auditory–perceptual study (see Frankford et al., 2024, for complete listener details). Timing typicality was operationally defined as "the extent to which a speaker's speech cadence is as expected in a typical speaker. Atypical timing can be halting, rushed, or too evenly paced for the context. This does not take into account factors like pitch or loudness." All listeners spoke North American English and reported no history of speech, language, hearing, or neurological disorders. In addition, listeners passed a hearing screening using a 25-dB HL cutoff at all tested frequencies (125, 250, 500, 1000, 2000, 4000, and 8000 Hz; American Speech-Language-Hearing Association, 2005). Listeners who formally studied speech, language, or hearing science or had any prior experience with auditory–perceptual ratings of motor speech disorders were excluded. VAS ratings from all raters were averaged together to yield a single timing typicality score for each PwPD. Intrarater reliability (repeated ratings within a session) was found to be only moderate, ICC(A, 1) = .73, potentially given the inexperience of the listeners. However, the interrater reliability for the average of 20 listeners (the outcome measure in the present study) was found to be excellent, ICC(C, $k$) = .95 (Koo & Li, 2016; McGraw & Wong, 1996).[2]

### Acoustic Measures

All recordings were transcribed for their exact linguistic content, and word boundaries were identified using the Penn Phonetics Lab Forced Aligner, an open-source forced alignment toolkit (Yuan & Liberman, 2008). Word boundaries for each participant were imported as a textgrid file into Praat (Boersma & Weenink, 2022), along with the audio, and were checked and corrected for gross errors in word alignment by the first author. The onset and offset time boundaries of each word in the passage were extracted using a custom MATLAB script. In the case of sound or syllable repetitions at the beginning of a word, the boundaries were determined for the first fluent production of the word, and passage words that were mispronounced or omitted from the recording received values

[1]Gender information was not available for this sample.

[2]Intrarater reliability was determined using the formula ICC(A, 1) because the repeated ratings did not systematically differ, and single-rater values were of most interest. Interrater reliability was determined using the formula ICC(C, $k$) because the repeated ratings systematically differed by listener and the reliability of the group average was of most importance for the present study (Koo & Li, 2016; McGraw & Wong, 1996). This measure of interrater reliability is consistent with prior auditory–perceptual literature examining intelligibility (e.g., Knowles et al., 2021; Tjaden et al., 2014; van Brenk et al., 2022).

of "not a number" (NaN). Word duration was computed as the difference between the offset and onset times for each word. This measure is inversely proportional to the articulation rate as measured in words per second. The difference between the offset time of one word and the onset time of the following word was calculated as the IWI, which captures the duration of any intervening pauses, words, or syllables. When a word was coded as NaN because it was mispronounced or omitted, IWI following that word also received a value of NaN.

MWD was calculated for each speaker by averaging word duration values (excluding NaN). This was then averaged across the 40 speakers without PD to obtain a reference based on the passage-level MWD. The absolute deviation from this reference was then computed for each speaker with PD to obtain their passage-relative MWD. The absolute deviation was used to reflect the fact that low ratings of timing typicality may result from either very long or very short MWDs. Similarly, MIWI was calculated by averaging the durations of intervals between words (excluding NaN) for each speaker. The MIWI of the 40 speakers without PD was averaged to obtain a reference based on the passage-level MIWI, and the absolute deviation from this reference was computed for each speaker with PD to obtain their passage-relative MIWI.

To determine similar types of deviations at the level of individual words within the passage, word durations and interword durations corresponding to each word (or the interval following each word) for the 40 speakers without PD were averaged across speakers to yield a "typical" pattern of word durations and IWI durations across the passage. The absolute deviation from each word duration and IWI was then computed for each speaker with PD and averaged across interonset intervals to obtain their word-relative MWD and MIWI.

### Data Analysis

All data analyses were carried out in RStudio (Version 2023.06.1). Our three hypotheses were tested by comparing the amount of variance explained by three sets of models using a general *F* test (*anova* function in package stats). We tested our first hypothesis that relative measures would improve the estimation of timing typicality by comparing a full model with all six acoustic variables with a model that only included absolute measures (MWD and MIWI). We then tested our second hypothesis that IWI measures would improve the estimation of timing typicality by comparing the full model with a model that only included duration measures. Next, we tested our third hypothesis that measures of duration and IWI deviation at the word level (word-relative MWDs and word-relative MIWI) would better estimate timing typicality by comparing the full model with a model only including absolute and passage-relative measures. Finally, to understand which variables individually contributed significantly to explaining variation in timing typicality, we evaluated the individual effects of a single complete model. Because there may be significant collinearity among the variables included, we ran a backward model selection procedure using the *stepAIC* function (package MASS) to find the simplest model that best fit the data. Assumptions for this model were assessed using the regression diagnostics plots using the R base function *plot*.

## Results

Summary statistics for each of the acoustic variables are shown in Table 1. A correlation matrix showing Pearson correlation coefficients among all of the acoustic measures can be found in Table 2. The mean timing typicality rating for the PwPD was 52 (*SD* = 19, range: 15–75). Plots comparing each of the acoustic variables against timing typicality can be found in Figure 1.

Comparing the full model (all six acoustic variables) with the model that only included absolute measures, there was significantly more variance explained by the full model than the reduced model, $R^2$ = .93 versus .11, $F(4, 13)$ = 37.02, $p < .001$. Comparing the full model with the model that only included duration-based measures, there was significantly more variance explained by the full model than the reduced model, $R^2$ = .93 versus .77, $F(3, 13) = 9.19$, $p = .002$. Comparing the full model with the model without word-relative acoustic variables, there was significantly more variance explained by the full model than the reduced model, $R^2$ = .93 versus .86, $F(2, 13) = 5.98$, $p = .01$.

To better understand which acoustic features significantly explained variability in timing typicality, the

**Table 1.** Summary statistics for acoustic variables.

| Acoustic variable | PwPD | | Control | |
|---|---|---|---|---|
| | *M* (*SD*) | Range | *M* (*SD*) | Range |
| MWD | 284 (56) | 186–423 | 284 (33) | 229–363 |
| MIWI | 71 (35) | 22–155 | 62 (23) | 22–137 |
| Passage-relative MWD | 43 (35) | 7–139 | — | — |
| Passage-relative MIWI | 29 (21) | 4–93 | — | — |
| Word-relative MWD | 62 (26) | 33–144 | — | — |
| Word-relative MIWI | 52 (24) | 15–75 | — | — |

*Note.* All values are in milliseconds. PwPD = people with Parkinson's disease; MWD = mean word duration; MIWI = mean interword interval.

**Table 2.** Pearson correlation matrix of acoustic variables.

| Acoustic variable | MIWI | Passage-relative MWD | Passage-relative MIWI | Word-relative MWD | Word-relative MIWI |
|---|---|---|---|---|---|
| MWD | .582 | .107 | .154 | .188 | .358 |
| MIWI | | −.027 | .680 | .152 | .854 |
| Passage-relative MWD | | | .134 | .945 | .080 |
| Passage-relative MIWI | | | | .304 | .875 |
| Word-relative MWD | | | | | .286 |

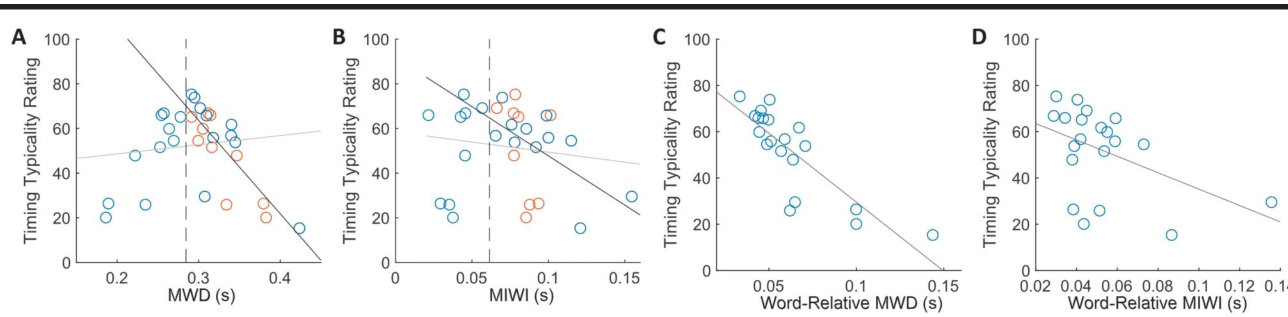
*Note.* MIWI = mean interword interval; MWD = mean word duration.

estimates of individual acoustic features in a single model were evaluated for significance. A backward model selection procedure yielded a model with four variables: MWD, MIWI, passage-relative MWD, and word-relative MIWI. The overall regression was statistically significant, $R^2 = .93$, $F(4, 15) = 46.92$, $p < .001$, and it was found that MWD ($\beta = 108.77$, $SE = 30.89$, $p = .003$), passage-relative MWD ($\beta = -403.02$, $SE = 39.21$, $p < .001$), and word-relative MIWI ($\beta = -542.65$, $SE = 113.85$, $p < .001$) significantly explained variance in timing typicality, whereas MIWI did not ($\beta = 116.17$, $SE = 90.50$, $p = .22$).

## Discussion

The present study aimed to determine acoustic measures that would account for most of the variance in perceptual ratings of how typical the timing of an individual with PD's speech was during a passage reading task. Six measures were examined—MWD (the inverse of articulation rate), passage-relative MWD, word-relative MWD, MIWI (the mean interval between word boundaries comprising pauses, disfluencies, and errors), passage-relative MIWI, and word-relative MIWI. By comparing a linear model that included all six variables with specific subsets of variables, all three hypotheses were confirmed: Deviation measures (computed relative to control speakers) improved estimation of timing typicality compared to absolute measures on their own, IWI measures (MIWI, passage-relative MIWI, and word-relative MIWI) significantly contributed to timing typicality above and beyond word duration measures, and deviation measures at the word level (word-relative MWDs and MIWI) improved the amount of variance in timing typicality accounted for relative to the absolute and passage-level measures alone. The raw acoustic measures MWD and MIWI, however, had no explanatory value, indicating that ratings of typicality related more to deviations from a reference, as hypothesized. Furthermore, an optimal model was derived using a stepwise regression procedure and evaluated to determine significant effects of individuals' acoustic measures. MWD, passage-relative MWD, and word-relative MIWI were all found to significantly contribute to explaining variability in timing typicality (accounting for the effects of the other variables) in a model accounting for 93% of the variance in timing typicality ratings. Each of the three results of the main hypothesis tests are discussed in greater detail below.

**Figure 1.** Relationship between acoustic variables and timing typicality in people with Parkinson's disease (PwPD). (A) Mean word duration (MWD) plotted against timing typicality. The dashed vertical line shows the reference MWD value from 40 healthy control speakers. Blue circles represent individual MWD values. Orange circles show the same PwPD as the blue circles left of the reference line but reflected about this line to show passage-relative MWD values. The light gray line indicates the least squares fit line for MWD, and the dark gray line indicates the least squares fit line for passage-relative MWD. (B) Mean interword interval (MIWI) plotted against timing typicality. The dashed vertical line shows the reference MIWI value from 40 healthy control speakers. Blue circles represent individual MIWI values. Orange circles show the same PwPD as the blue circles left of the reference line but reflected about this line to show relative passage-relative MIWI values. The light gray line indicates the least squares fit line for MIWI, and the dark gray line indicates the least squares fit line for passage-level MIWI. (C) MWD-relative MWD plotted against timing typicality. Blue circles represent individual speakers. The solid line indicates the least squares fit line. (D) Word-relative MIWI plotted against timing typicality. Blue circles represent individual speakers. The solid line indicates the least squares fit line.



The first result—that the deviation measures provided a significantly better fit of timing typicality than the absolute measures alone—was as expected. Because the definition of timing typicality did not have directionality (e.g., "atypical speech can be rushed, halting, or too evenly paced"), deviation from the average value of control speakers allowed for speech with either shorter or longer word durations to be rated as less typical. The significant effects of passage-relative MWD and word-relative MIWI in the optimal model suggest that deviation in both word durations and IWIs contributed to reduced timing typicality ratings. That is, deviation in both articulation rate and interword timing components (pausing/disfluencies/reading corrections) independently contributed to the sense of atypical timing. This result is also consistent with the second result, that IWI-based measures significantly explained timing typicality ratings above and beyond word duration measures. Atypical pausing characteristics and disfluencies are both found in PD (Darling-White & Huber, 2020; Hammen & Yorkston, 1996; Knowles et al., 2021; Logemann et al., 1978; Lowit et al., 2018; Martínez-Sánchez et al., 2016; McRae et al., 2002; Nishio & Niimi, 2001; Skodda & Schlegel, 2008; Tjaden & Wilding, 2011), and this study confirms that both are salient characteristics in perceiving speech timing in this population. The third result, that calculating absolute deviations at the word level improved explanation of variability in timing typicality ratings, also suggested that timing variability was a salient feature of atypical timing in PD. Based on the results of the optimal model, this effect was mainly due to variation in IWIs across the passage. To better understand what word-relative MIWI captures, it is worth discussing how the word-relative MIWI and passage-relative MIWI are related. The passage-relative MIWI measure is equivalent to the word-relative MIWI if the absolute value of word-level IWI deviation is not taken. When the absolute value is taken, the word-relative MIWI captures average pausing/disfluency differences if most or all syllables have either longer or shorter IWIs than the reference (overall increased/decreased interword time). However, when IWIs are near average for a given speaker, it captures variability—greater spread around the reference values leads to greater word-relative MIWI values. The interpretation of this variability component is not straightforward. It could reflect more overall variability across IWIs (i.e., shorter IWIs are shorter, long IWIs are longer), less overall variability across IWIs (i.e., long IWIs are shorter, short IWIs are longer). However, it could also reflect something else, potentially some aspect of variability previously captured in perceptual studies, such as inappropriate silences, short rushes, and variable rate (Darley et al., 1969; Plowman-Prine et al., 2009), or acoustic studies, such as reduced breath pauses at typical boundaries and more breath pauses at atypical boundaries (Darling-White & Huber, 2020) or acceleration across the passage (Skodda & Schlegel, 2008). Indeed, the variability component could be capturing several different timing characteristics that vary across individuals. Future work is needed to determine whether this variability is due to commonalities across speakers due to specific features of the sample (e.g., differences between syntactically appropriate and syntactically inappropriate word boundaries), temporal variability (e.g., acceleration across phrases or the whole passage), or idiosyncrasies in timing characteristics (e.g., some individuals showing inappropriate pauses and short rushes of speech).

A small number of acoustic timing measures were able to account for 93% of the variance in timing typicality ratings ($R^2$ = .93), indicating that acoustic timing measures are highly related to the perceptual ratings. This provides support for the assertion that ratings of timing typicality reflect actual speech timing. However, other factors that impact prosody such as pitch and loudness variation could have indirectly affected perceptions of speech timing despite the instructions for listeners to explicitly ignore these factors. It has also been suggested that articulatory imprecision contributes to the perception of increased speaking rate in PwPD (Netsell et al., 1975; Tjaden, 2000), although evidence for this is minimal. Because pitch, loudness, and articulation variables may be correlated with one or more of the acoustic timing measures in the present study, future work examining measures of all of these speech characteristics will be needed to confirm their relative contributions to the perception of timing typicality.

## *Limitations and Future Directions*

The models that best explained timing typicality included acoustic timing measures that considered the average production of the reading passage in a sample of typical speakers. This introduces a challenge for future potential applications of using these acoustic measures to replace perceptual ratings of timing typicality because large samples of recordings from age- and sex-matched control speakers are not usually available. One area for future research will be determining how many typical speakers are required to attain similar results and how closely they need to be matched to the PwPD. Since older speakers tend to speak more slowly (e.g., Jacewicz et al., 2009; Ramig, 1983), matching PwPD and typical speakers by age may be important when computing norm-referenced acoustic measures. Determining the minimum number of typical speakers needed to attain similar results would help to evaluate the feasibility of using these acoustic measures in practice. One possibility for improving feasibility would be to have a freely available database of typical speakers of different ages reading a standardized passage using standardized instructions, from which a sample matched to the characteristics of a clinical or research sample could automatically be derived.

A limitation of using word-relative MIWI for clinical and research purposes is that, although it may be an informative component of auditory–perceptual ratings of timing typicality, it can only be calculated for a standard reading passage where individual word boundary timing deviations can be derived. On one hand, both research and clinical settings already use standard passages during speaker evaluation, so it should not be burdensome to record such a sample for comparison. On the other hand, reading passages do not provide the most ecological contexts for assessing speech. Therefore, despite word-relative MIWI explaining a significant amount of variance in typicality ratings, continuing to evaluate passage-relative measures in future work will be important for determining acoustic measures that are sensitive to timing typicality in more spontaneous or variable speaking conditions. One avenue for doing this will be evaluating passage-level measures of duration variability or systematic fluctuation in word and interword timing. The next steps in the assessment of these measures include evaluating acoustic/perceptual relations in a new sample of speakers and in other reading passages to determine how generalizable the current results are. In addition, determining the extent to which changes in acoustic ratings for individuals across time predict changes in perceptual ratings would be important for applying such acoustic measures to within-speaker changes due to therapy or disease progression.

Finally, the method of sampling PwPD—namely, using a sample from an earlier study in which ratings of intelligibility and naturalness were not correlated (see the Methods section)—may have made the results less generalizable to the broader population of PwPD. In particular, this sample contained speakers with mostly mild to moderate overall dysarthria severity. However, prior work has found that in PD, there can be large variation in overall severity despite high intelligibility in the earlier stages (Tjaden et al., 2014), potentially indicating a reduced correlation between intelligibility and naturalness. In addition, potential differences in regional dialect between the listeners and the speakers may have biased ratings of timing typicality based on the listeners' own expectations of speech timing. Unfortunately, while all speakers and listeners were native speakers of a North American dialect of English, we do not have further information on their specific regional dialects.

## Conclusions

The present study proposed several acoustic measures and evaluated how well they could explain variation in auditory–perceptual ratings of timing typicality. A combination of absolute word and interword timing measures provided an exceptional fit to the perceptual ratings ($R^2$ = .93). Acoustic timing measures may therefore offer a less time-consuming and less expensive alternative for estimating timing typicality in PwPD than conducting an auditory–perceptual experiment with 20 listeners. Future work should examine the relationship between these acoustic measures and perceptual ratings of speech timing in a larger speaker sample, using varying speech stimuli, and in relation to changes due to disease progression and treatment to determine the utility of these measures in research and clinical settings.

## Data Availability Statement

The data sets generated and/or analyzed during the current study are available from the corresponding author on reasonable request.

## Acknowledgments

This work was supported by R01 DC020867 (to C.E.S. and D.D.M.) from the National Institute on Deafness and Other Communication Disorders. The authors wish to thank Alex Estrada for their assistance collecting auditory–perceptual data, as well as Julia Toto, Kaitlyn Siedman, and Megan Cushman for their assistance with the recruitment.

## References

**American Speech-Language-Hearing Association.** (2005). *Guidelines for manual pure-tone threshold audiometry* [Guidelines]. https://doi.org/10.1044/policy.GL2005-00014

**Boersma, P., & Weenink, D.** (2022). *Praat: Doing phonetics by computer* (Version 6.2.20) [Computer software].

**Chiu, Y.-F., Neel, A., & Loux, T.** (2021). Exploring the acoustic perceptual relationship of speech in Parkinson's disease. *Journal of Speech, Language, and Hearing Research, 64*(5), 1560–1570. https://doi.org/10.1044/2021_JSLHR-20-00610

**Darley, F. L., Aronson, A. E., & Brown, J. R.** (1969). Differential diagnostic patterns of dysarthria. *Journal of Speech and Hearing Research, 12*(2), 246–269. https://doi.org/10.1044/jshr.1202.246

**Darling-White, M., & Huber, J. E.** (2020). The impact of Parkinson's disease on breath pauses and their relationship to speech impairment: A longitudinal study. *American Journal of Speech-Language Pathology, 29*(4), 1910–1922. https://doi.org/10.1044/2020_ajslp-20-00003

**Fairbanks, G.** (1960). *Voice and articulation drillbook* (2nd ed.). Harper & Row.

**Flasskamp, A., Kotz, S. A., Schlegel, U., & Skodda, S.** (2012). Acceleration of syllable repetition in Parkinson's disease is more prominent in the left-side dominant patients. *Parkinsonism & Related Disorders, 18*(4), 343–347. https://doi.org/10.1016/j.parkreldis.2011.11.021

**Frankford, S. A., Estrada, A., & Stepp, C. E.** (2024). Contributions of speech timing and articulatory precision to listener perceptions of intelligibility and naturalness in Parkinson's disease. *Journal of Speech, Language, and Hearing Research, 67*(9), 2951–2963. https://doi.org/10.1044/2024_JSLHR-23-00802

**Giladi, N., Shabtai, H., Rozenberg, E., & Shabtai, E.** (2001). Gait festination in Parkinson's disease. *Parkinsonism & Related Disorders, 7*(2), 135–138. https://doi.org/10.1016/s1353-8020(00)00030-4

**Hammen, V. L., & Yorkston, K. M.** (1996). Speech and pause characteristics following speech rate reduction in hypokinetic dysarthria. *Journal of Communication Disorders, 29*(6), 429–445. https://doi.org/10.1016/0021-9924(95)00037-2

**Ho, A. K., Iansek, R., Marigliani, C., Bradshaw, J. L., & Gates, S.** (1999). Speech impairment in a large sample of patients with Parkinson's disease. *Behavioural Neurology, 11*(3), 131–137. https://doi.org/10.1155/1999/327643

**Jacewicz, E., Fox, R. A., O'Neill, C., & Salmons, J.** (2009). Articulation rate across dialect, age, and gender. *Language Variation and Change, 21*(2), 233–256. https://doi.org/10.1017/S0954394509990093

**Knowles, T., Adams, S. G., & Jog, M.** (2021). Speech rate mediated vowel and stop voicing distinctiveness in Parkinson's disease. *Journal of Speech, Language, and Hearing Research, 64*(11), 4096–4123. https://doi.org/10.1044/2021_JSLHR-21-00160

**Koo, T. K., & Li, M. Y.** (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine, 15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

**Logemann, J. A., Fisher, H. B., Boshes, B., & Blonsky, E. R.** (1978). Frequency and cooccurrence of vocal tract dysfunctions in the speech of a large sample of Parkinson patients. *Journal of Speech and Hearing Disorders, 43*(1), 47–57. https://doi.org/10.1044/jshd.4301.47

**Lowit, A., Marchetti, A., Corson, S., & Kuschmann, A.** (2018). Rhythmic performance in hypokinetic dysarthria: Relationship between reading, spontaneous speech and diadochokinetic tasks. *Journal of Communication Disorders, 72,* 26–39. https://doi.org/10.1016/j.jcomdis.2018.02.005

**Maffia, M., De Micco, R., Pettorino, M., Siciliano, M., Tessitore, A., & De Meo, A.** (2021). Speech rhythm variation in early-stage Parkinson's disease: A study on different speaking tasks. *Frontiers in Psychology, 12,* Article 668291. https://doi.org/10.3389/fpsyg.2021.668291

**Martens, H., Van Nuffelen, G., Dekens, T., Hernández-Díaz Huici, M., Kairuz Hernández-Díaz, H. A., De Letter, M., & De Bodt, M.** (2015). The effect of intensive speech rate and intonation therapy on intelligibility in Parkinson's disease. *Journal of Communication Disorders, 58,* 91–105. https://doi.org/10.1016/j.jcomdis.2015.10.004

**Martínez-Sánchez, F., Meilán, J. J. G., Carro, J., Gómez Íñiguez, C., Millian-Morell, L., Pujante Valverde, I. M., López-Alburquerque, T., & López, D. E.** (2016). Speech rate in Parkinson's disease: A controlled study. *Neurología (English Edition), 31*(7), 466–472. https://doi.org/10.1016/j.nrleng.2014.12.014

**McGraw, K. O., & Wong, S. P.** (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods, 1*(1), 30–46. https://doi.org/10.1037/1082-989X.1.1.30

**McRae, P. A., Tjaden, K., & Schoonings, B.** (2002). Acoustic and perceptual consequences of articulatory rate change in Parkinson disease. *Journal of Speech, Language, and Hearing Research, 45*(1), 35–50. https://doi.org/10.1044/1092-4388(2002/003)

**Moreau, C., Ozsancak, C., Blatt, J., Derambure, P., Destee, A., & Defebvre, L.** (2007). Oral festination in Parkinson's disease: Biomechanical analysis and correlation with festination and freezing of gait. *Movement Disorders, 22*(10), 1503–1506. https://doi.org/10.1002/mds.21549

**Netsell, R., Daniel, B., & Celesia, G. G.** (1975). Acceleration and weakness in parkinsonian dysarthria. *Journal of Speech and Hearing Disorders, 40*(2), 170–178. https://doi.org/10.1044/jshd.4002.170

**Nishio, M., & Niimi, S.** (2001). Speaking rate and its components in dysarthric speakers. *Clinical Linguistics & Phonetics, 15*(4), 309–317. https://doi.org/10.1080/02699200010024456

**Pettorino, M., Busà, M. G., & Pellegrino, E.** (2016). Speech rhythm in Parkinson's disease: A study on Italian. *Interspeech, 2016,* 1958–1961. https://doi.org/10.21437/interspeech.2016-74

**Plowman-Prine, E. K., Okun, M. S., Sapienza, C. M., Shrivastav, R., Fernandez, H. H., Foote, K. D., Ellis, C., Rodriguez, A. D., Burkhead, L. M., & Rosenbek, J. C.** (2009). Perceptual characteristics of Parkinsonian speech: A comparison of the pharmacological effects of levodopa across speech and non-speech motor systems. *NeuroRehabilitation, 24*(2), 131–144. https://doi.org/10.3233/NRE-2009-0462

**Ramig, L. A.** (1983). Effects of physiological aging on speaking and reading rates. *Journal of Communication Disorders, 16*(3), 217–226. https://doi.org/10.1016/0021-9924(83)90035-7

**Rusz, J., Tykalova, T., Novotny, M., Zogala, D., Sonka, K., Ruzicka, E., & Dusek, P.** (2021). Defining speech subtypes in de novo Parkinson disease: Response to long-term levodopa therapy. *Neurology, 97*(21), e2124–e2135. https://doi.org/10.1212/wnl.0000000000012878

**Skodda, S., Flasskamp, A., & Schlegel, U.** (2010). Instability of syllable repetition as a marker of disease progression in Parkinson's disease: A longitudinal study. *Movement Disorders, 26*(1), 59–64. https://doi.org/10.1002/mds.23382

**Skodda, S., & Schlegel, U.** (2008). Speech rate and rhythm in Parkinson's disease. *Movement Disorders, 23*(7), 985–992. https://doi.org/10.1002/mds.21996

**Tjaden, K.** (2000). A preliminary study of factors influencing perception of articulatory rate in Parkinson disease. *Journal of Speech, Language, and Hearing Research, 43*(4), 997–1010. https://doi.org/10.1044/jslhr.4304.997

**Tjaden, K., Lam, J., & Wilding, G.** (2013). Vowel acoustics in Parkinson's disease and multiple sclerosis: Comparison of clear, loud, and slow speaking conditions. *Journal of Speech, Language, and Hearing Research, 56*(5), 1485–1502. https://doi.org/10.1044/1092-4388(2013/12-0259)

**Tjaden, K., Sussman, J. E., & Wilding, G. E.** (2014). Impact of clear, loud, and slow speech on scaled intelligibility and speech severity in Parkinson's disease and multiple sclerosis. *Journal of Speech, Language, and Hearing Research, 57*(3), 779–792. https://doi.org/10.1044/2014_JSLHR-S-12-0372

**Tjaden, K., & Wilding, G.** (2011). Speech and pause characteristics associated with voluntary rate reduction in Parkinson's disease and multiple sclerosis. *Journal of Communication Disorders, 44*(6), 655–665. https://doi.org/10.1016/j.jcomdis.2011.06.003

**Tysnes, O.-B., & Storstein, A.** (2017). Epidemiology of Parkinson's disease. *Journal of Neural Transmission, 124*(8), 901–905. https://doi.org/10.1007/s00702-017-1686-y

**van Brenk, F., Stipancic, K., Kain, A., & Tjaden, K.** (2022). Intelligibility across a reading passage: The effect of dysarthria and cued speaking styles. *American Journal of Speech-Language Pathology, 31*(1), 390–408. https://doi.org/10.1044/2021_AJSLP-21-00151

**Walsh, B., & Smith, A.** (2012). Basic parameters of articulatory movements and acoustics in individuals with Parkinson's disease. *Movement Disorders, 27*(7), 843–850. https://doi.org/10.1002/mds.24888

**Yuan, J., & Liberman, M.** (2008). Speaker identification on the SCOTUS corpus. *Journal of the Acoustical Society of America, 123*(Suppl. 5), Article 3878. https://doi.org/10.1121/1.2935783