The Effects of Electromagnetic Articulography Sensors on Speech in Individuals with and without Parkinson’s Disease

Austin Thompson1, Micah Hirsch2, and Yunjung Kim3

Abstract

1. Department of Communication Sciences and Disorders, University of Houston
2. Department of Speech, Language, and Hearing Sciences, Boston University
3. School of Communication Science and Disorders, Florida State University

**Conflicts of Interest**:  
The authors have no relevant financial or non-financial information to disclose.

**Corresponding Author**:  
Austin Thompson, PhD, CCC-SLP  
athomp27@central.uh.edu

**Authorship Contributions** (CRediT taxonomy - https://casrai.org/credit/)  
*Author Roles*: 1conceptualization, 2data curation, 3formal analysis, 4funding acquisition, 5investigation, 6methodology, 7project administration, 8resources, 9software, 10supervision, 11validation, 12visualization, 13writing – original draft, 14writing – reviewing & editing

AT: 1, 2, 3, 4, 5, 6, 9, 11, 12, 14  
MH: 3, 5, 6, 7, 9, 11, 12, 13, 14  
YK: 1, 2, 4, 5, 6, 7, 8, 9, 10, 11, 12, 14

**Ethical Approval**: This study was approved by the Florida State University’s Institutional Review Board (FSU IRB: 00002525).

**Keywords**: Dysarthria; Acoustics; electromagnetic articulography

# Abstract

**Purpose:** This study examined how wearing electromagnetic articulography (EMA) sensors affects acoustic and perceptual speech outcomes in people with Parkinson’s disease (PwPD) with dysarthria and neurologically healthy control speakers. Additionally, the study explored potential after-effects on acoustic and perceptual measures following approximately 45 minutes of wearing EMA sensors in both groups. Finally, we investigated whether wearing EMA sensors or after-sensor effects differentially impacted the two groups.

**Methods:** Thirty-four speakers (21 Controls and 13 PwPD) read “The Caterpillar” passage at three time points: (1) Before Sensors, (2) With Sensors, and (3) After Sensors. We analyzed changes in acoustic (articulation rate, articulatory-acoustic vowel space [AAVS], first and second spectral moment coefficients for fricatives) and perceptual (speech intelligibility, naturalness) measures across two key contrasts: sensor effects (With Sensors - Before Sensors) and after-sensor effects (After Sensors - Before Sensors).

**Results:** Bayesian linear mixed-effects models showed sensor effects (With Sensors - Before Sensors), with EMA sensors reducing intelligibility and naturalness and altering fricative spectral moments in both groups. Additionally, Control speakers exhibited a faster articulation rate with sensors. Notably, PwPD were more negatively impacted by sensor effects in terms of intelligibility ratings. After-sensor effects (After Sensors - Before Sensors) were also observed: Control speakers spoke faster following sensor removal, while PwPD demonstrated increased AAVS and were perceived as more natural. However, there was no compelling evidence that after-sensor effects differed between groups.

**Conclusion:** EMA sensors primarily impact sibilant fricative production and perceptions of intelligibility and naturalness in PwPD and Control speakers. PwPD experience greater sensor-related reductions in intelligibility, which should be carefully considered when using speech data collected with EMA to assess perceptual measures in clinical populations. Finally, PwPD exhibited increased naturalness and greater spectral distinctiveness following sensor removal, which we speculate may stem from increased passage familiarity and reduced cognitive demand.

# Introduction

Historically, our understanding of motor speech disorders, like dysarthria, has relied on perceptual and acoustic data. These methods are well-suited for measuring the phonatory, resonatory, and prosodic deficits associated with dysarthria. However, using perceptual and acoustic data to make inferences about the underlying articulatory movement is complicated by motor equivalence, which refers to the idea that multiple articulatory gestures can produce the same acoustic signal (Brunner & Hoole, 2012; Hughes & Abbs, 1976; Perkell et al., 1993; Perrier & Fuchs, 2015). For this reason, kinematic analysis methods are beneficial because they allow researchers to directly examine articulatory movement.

Among various kinematic methods, such as ultrasound and palatography (Hardcastle et al., 1991; Klein et al., 2013; McAuliffe et al., 2006a, 2006b), electromagnetic articulography (EMA) systems are among the most commonly used kinematic methods (Berry, 2011; Kim et al., 2024; Savariaux et al., 2017). By providing data on articulatory working space, displacement, speed, movement variability, and interarticulator coordination (Chu et al., 2020; Lee et al., 2017; Masapollo & Nittrouer, 2023; Mefferd, 2015; Rong et al., 2012; Teplansky et al., 2023; Thompson & Kim, 2024), EMA systems offer valuable kinematic data, which complements and triangulates the relatively more extensive literature on acoustic and perceptual data. Kinematic data—when combined with perceptual and acoustic measures—can offer a more nuanced understanding of speech motor control, particularly in populations with neurological speech impairments.

While EMA methods offer advantages for tracking articulatory movement, they also pose distinct challenges. EMA systems use small (e.g., 2 × 2 mm for the Wave System; NDI, Canada) strategically placed sensors on the lips, tongue, and jaw, each connected to the recording device by thin wires (e.g., 0.4 mm in diameter) (Rebernik et al., 2021). The mere presence of these sensors introduces a continuous somatosensory perturbation (constant tactile contact of the sensors and wires against the articulators) that speakers must accommodate. Unlike transient mechanical perturbations applied to the jaw (Lametti et al., 2012; Nasir & Ostry, 2006), EMA sensors impose a steady‐state alteration in oral input rather than a discrete displacement of articulator trajectories. Further, the novelty of the tactile sensation introduced by the sensors may draw attentional resources away from speech planning, potentially altering motor performance and making speech feel more effortful and less natural. Thus, understanding how sensor placement affects speech production is especially important in clinical populations, such as dysarthria secondary to Parkinson’s disease, who may be particularly sensitive to these somatosensory and cognitive demands.

The purpose of the current study was to examine the impact of EMA sensors on people with Parkinson’s disease (PwPD). Parkinson’s disease is a progressive neurodegenerative disorder that causes hypokinetic dysarthria in approximately 90% of cases (Ho et al., 1999; Moya-Galé & Levy, 2019). Hypokinetic dysarthria is characterized by a reduced range of articulatory gestures (Mefferd, 2015; Mefferd & Dietrich, 2019; Thompson & Kim, 2024), as well as perceptual qualities such as vocal breathiness, monopitch, monoloudness, short rushes of speech, and imprecise articulation (Darley et al., 1969a, 1969b; Duffy, 2020). These deficits typically cause decreased intelligibility and naturalness (Anand & Stepp, 2015; De Bodt et al., 2002; Plowman-Prine et al., 2009), which can have a negative impact on communicative participation and quality of life (Borrie et al., 2022; Spencer et al., 2020). Acoustically, PwPD often demonstrate smaller acoustic working space (e.g., smaller acoustic vowel space area (aVSA) or articulatory-acoustic vowel space [AAVS]) and reduced spectral contrastivity in fricatives, indicating diminished articulatory excursions and overall reduced articulatory precision (Bang et al., 2013; Lam & Tjaden, 2016; McRae et al., 2002, 2002; Tjaden et al., 2013; Whitfield & Mehta, 2019). Given these articulatory deficits in PwPD, EMA is a well-suited method for directly capturing and assessing kinematic movement in PwPD. However, it is not well understood how the presence of EMA sensors might impact speech production in PwPD.

## Sensor Effects

Investigations into the impact of EMA sensors on speech have predominantly been motivated by assessing the external validity of EMA findings. In other words, if EMA sensors systematically alter speech production or its perception, then caution should be exercised when generalizing results from EMA studies to study findings without kinematic data. A few studies have directly investigated the impact of kinematic sensors on speech production in neurologically healthy speakers, as summarized below.

Sensor effects have primarily been examined for fricative consonant production, specifically /s/ and /ʃ/. These sibilant fricatives require precise lingual-alveolar constriction and are, therefore, likely to be impacted by the presence of lingual sensors, which are often placed .5-2 cm from the tongue tip (Rebernik et al., 2021). Fricatives are typically studied using spectral moment analyses (Forrest et al., 1988), with spectral moments one (M1) and two (M2) representing the mean (spectral center of gravity) and standard deviation (spectral standard deviation) of spectral energy, respectively. M1 provides information about the place of articulation and has an inverse relationship to the size of the front cavity of the oral constriction. Therefore, M1 values are typically higher for /s/ compared to /ʃ/ (Jongman et al., 2000). In contrast, the articulatory basis of M2 is less well understood, but likely distinguishes sibilant from non-sibilant consonants and/or place of articulation for fricatives (Jongman et al., 2000; Koenig et al., 2013; Petrović, 2020). Therefore, M2 may still be relevant when considering sensor effects.

There is evidence suggesting that the presence of EMA sensors impact spectral moment measures for sibilant fricative production in healthy speakers. Dromey et al. (2018) examined EMA sensor effects on /s/ and /ʃ/ production and observed group-level sensor effects in neurologically healthy speakers, specifically decreased M1 values for /s/ and increased M2 values for /ʃ/. Two explanations for these findings have been proposed. One possibility is that the sensors and their attached wires, particularly when located near the primary constriction site, act as a mechanical perturbation, disrupting airflow and directly altering the resulting acoustic signal. Alternatively, speakers may adjust their articulatory patterns to compensate for sensor placement—an explanation that is not mutually exclusive with the mechanical perturbation hypothesis. Such adjustments could lead to a less forward constriction for /s/ (hence, a larger front cavity) and likely compensatory articulatory changes for /ʃ/. In contrast, Weismer & Bunton (1999) studied the effect of X-ray microbeam pellet placement[[1]](#footnote-21) on /s/ and /ʃ/ production in neurologically healthy speakers. They found about 20% of the speakers showed increased M1 values with pellets on compared to pellets off, indicating either a disruption of airflow or altered articulatory behavior characterized by a smaller front cavity and a more forward constriction. Although the findings of these two studies differ in terms of the direction of sensor-related effects on M1, they both suggest that sensor presence may affect articulatory behavior, acoustic signal characteristics, or both, resulting in altered spectral moment measures for sibilant fricatives.

To understand whether these sensor effects primarily result from airflow disruption or compensatory articulatory behavior, it may be useful to investigate phonemes involving less constriction, such as vowels, which are theoretically less susceptible to airflow disruptions from sensor placement. However, only a few studies have examined the impact of EMA sensors on vowel articulation and working space in healthy speakers, and their findings have been mixed. Weismer & Bunton (1999) found that some speakers had higher first formant frequency (F1) values and lower second formant frequency (F2) values during vowel production with the pellets on, suggesting a greater mouth opening (likely due to greater jaw movement) and more retracted tongue position, likely to avoid contact between the lingual pellets and the alveolar ridge. In a recent study, Tienkamp et al. (2024) found neurologically healthy speakers have reduced AAVS following EMA sensor placement, indicating a reduced articulatory working space with EMA sensors. The differences between these findings could be attributed to methodological variations, particularly the use of wired EMA sensors versus wireless X-ray microbeam pellets. Another factor to consider is that both studies used repeated stimuli (sentence repetition in Weismer & Bunton (1999), and repeated passage reading in Tienkamp et al. (2024)), but neither accounted explicitly for possible practice effects, which could confound their results; speakers tend to speak faster on subsequent readings of the same stimuli, and faster speech is typically characterized by reduced working space (Turner et al., 1995; Weismer et al., 2000). Nevertheless, both studies suggest that lingual sensors may lead to alterations in vowel articulation, whether through direct somatosensory perturbation or compensatory motor adjustments.

Additionally, there is some evidence that EMA sensors can lead to perceptually degraded speech. Weismer & Bunton (1999) found no consistent impact of pellets on perceptual judgments of articulatory precision, while Dromey et al. (2018) found reduced ratings of articulatory precision with EMA sensors on compared to before EMA sensor placement. For intelligibility, Meenakshi et al. (2014) found that listeners’ forced-choice intelligibility judgments for various VCV stimuli were significantly lower with sensors compared to without sensors. Taken together, the small number of studies and the conflicting findings underscores the need for further research to determine how EMA sensors influence different perceptual outcomes. Moreover, the effects of EMA sensors on some crucial perceptual constructs, like speech naturalness, remains unexplored.

Given that EMA is well-suited to characterize the articulatory deficits of clinical populations, it is crucial to understand how these populations are affected by the presence of sensors. This understanding has important implications for the validity of between-group comparisons. If sensor effects influence all speakers similarly, then researchers can assume that sensor-induced variability does not confound comparisons between speakers with and without motor speech disorders. However, if individuals with motor speech disorders are disproportionately affected by sensors, then comparisons of articulatory movements across groups could be misleading if these unique sensor effects are not appropriately accounted for.

Furthermore, PwPD might be uniquely affected by the somatosensory perturbation caused by the EMA sensors. As previously stated, compensating for the presence of EMA sensors likely requires auditory and somatosensory acuity. However, PwPD have been shown to have auditory (Chen & Watson, 2017; Mollaei et al., 2016) and somatosensory deficits (Chen & Watson, 2017; Hammer & Barlow, 2010; Schneider et al., 1986). For instance, Chen & Watson (2017) found that PwPD had reduced tactile acuity of the tongue tip, which was linked to reduced spectral contrasts between /s/ and /ʃ/. Extending this logic, it is plausible that reduced auditory or somatosensory sensitivity in PwPD may constrain both their awareness of sensor-related perturbations and the extent to which they can implement effective compensatory adjustments compared to Control speakers.

To date, only Tienkamp et al. (2024) have investigated EMA sensor effects in PwPD. In their study, 46 speakers (23 control speakers and 23 PwPD) completed readings both before and after five sensors were applied: one on the jaw, two on the lips, and two on the tongue (1 cm from the tongue tip and 5 cm anterior to the /k/ constriction). Both groups showed a comparable reduction in AAVS following sensor placement, suggesting that sensors reduced working space and that PwPD were not differentially affected. However, because participants read the same passage multiple times, practice effects may have increased articulation rate and thus compressed vowel space (Turner et al., 1995; Weismer et al., 2000). Without controlling for articulation rate, it is unclear whether the smaller AAVS reflects sensor effects or repeated-reading practice effects. Further research should control for articulation rate to isolate the specific impact of EMA sensors on vowel and fricative articulation.

Furthermore, no studies have examined how speech produced with EMA sensors in PwPD is perceived by listeners. This area of investigation has important methodological implications. EMA is often used to assess therapeutic strategies in clinical populations by capturing both kinematic data and listener-based perceptual ratings in response cues like speaking louder, slower, or more clearly (Kearney et al., 2017; Thompson & Kim, 2024). While relative within-speaker changes across conditions (e.g., habitual vs. clear speech) may remain interpretable, between-group comparisons may be biased by differential EMA effects. Evidence from healthy speakers shows that sensors may reduce perceptual ratings (Dromey et al., 2018; Meenakshi et al., 2014), and it is plausible that these negative perceptual effects may be larger or qualitatively different in clinical populations with altered sensorimotor control. If unaccounted for, such sensor effects could confound between-group perceptual comparisons.

## After-Sensor Effects

Examining speech after sensor removal (i.e., after-sensor effects) may provide additional insights beyond sensor effects alone. To our knowledge, no studies have investigated the after-effects of EMA sensor removal. However, research on other somatosensory perturbations, such as mechanical jaw loads (Nasir & Ostry, 2006; Tremblay et al., 2003) and dental prostheses (Hamlet & Stone, 1976), suggests that after-effects typically manifest as compensatory responses in the opposite direction of the perturbation. For instance, studies that have applied a mechanical load to the jaw have found that speakers adapt their jaw movements to counteract the load to maintain their unperturbed acoustic output, and when the load is removed, their jaw may initially deviate in the opposite direction (Nasir & Ostry, 2006; Tremblay et al., 2003). Given that EMA sensors introduce a somatosensory perturbation that may alter vowel and fricative production, it is reasonable to hypothesize that there may be a period of readjustment following sensor removal as the articulators return to their natural movement patterns.

Investigating after-effects in PwPD may be particularly beneficial, given their characteristic hypokinetic dysarthria. If EMA sensors restrict a speaker’s range of motion, as suggested by findings of reduced acoustic distinctiveness and perceptions of precision (Dromey et al., 2018; Tienkamp et al., 2024), speakers may compensate by making larger articulatory gestures while wearing sensors to achieve their pre-sensor acoustic output. If these compensatory adjustments persist after sensor removal, they could possibly yield perceptual benefits. This would be particularly beneficial for PwPD, given that hypokinetic dysarthria is characterized by reduced amplitude of articulatory movements, vowel centralization, and decreased spectral contrastivity in fricatives (Mefferd & Dietrich, 2019; Thompson & Kim, 2024; Tjaden & Wilding, 2004; Weismer et al., 2001).

Research on after-effects following speech-related somatosensory perturbations in PwPD is limited. However, insights can be drawn from research on sensorimotor adaptation, which often uses auditory feedback paradigms. These studies have shown that both PwPD and control speakers exhibit some degree of maintained adaptation to F1 formant perturbations even after the perturbation is removed (Abur et al., 2018; Miller et al., 2023; Mollaei et al., 2013; Purcell & Munhall, 2006; Villacorta et al., 2007). However, it is important to note that after-effects are not typically the focus of these studies, and individual variability is observed. Beyond speech, after-effects have been documented in PwPD and healthy individuals following sensorimotor perturbations in various motor tasks, such as walking (Bultitude et al., 2012; Roemmich et al., 2014; Sorrento et al., 2018), pointing (Buch et al., 2003; Contreras-Vidal & Buch, 2003), and throwing (Martin et al., 1996). However, unlike these studies, which primarily manipulate visual or auditory feedback, EMA sensors introduce a somatosensory perturbation, warranting further investigation into their unique after-effects.

## The Current Study

This study is a secondary analysis of data originally collected to examine perceptual, acoustic, and kinematic speech characteristics in PwPD with hypokinetic dysarthria and neurologically healthy control speakers. Here, we investigated the impact of EMA sensors on PwPD and control speakers, focusing on both sensor presence (sensor effects) and post-sensor removal (after-sensor effects). An additional aim is to determine whether EMA sensors impact PwPD differently compared to control speakers, thereby assessing the external validity of EMA data. To explore these effects, we used spectral moment analyses (M1 and M2) to evaluate sibilant fricative production (/s/ and /ʃ/), analyzed AAVS to examine acoustic working space, and gathered perceptual ratings of intelligibility and naturalness. Our research questions are: (1) Do EMA sensors affect acoustic and perceptual measures in PwPD and control speakers (Before Sensors vs. With Sensors)? If so, do the effects differ between groups? and (2) Are there after-sensor effects in PwPD and control speakers (Before Sensors vs. After Sensors)? If so, do the effects differ between groups?

Based on prior studies (Dromey et al., 2018; Tienkamp et al., 2024; Weismer & Bunton, 1999), we hypothesize that both groups will show sensor effects, reflected by a decrease in perceptual and acoustic measures, except for M2, with sensors on compared to before sensor placement. For M2, we expect an increase with sensors, indicating a wider spread of spectral values. Because the lingual sensors sit directly at the constriction site for /s/ and /ʃ/, we predict that fricative spectral moments (M1 and M2) will be more disrupted by EMA sensor placement than AAVS in both groups. Further, given evidence of somatosensory deficits in PwPD (Chen & Watson, 2017; Hammer & Barlow, 2010; Schneider et al., 1986), we hypothesize that PwPD may be less or differentially affected by the somatosensory perturbation introduced by the EMA sensors. As a result, they may demonstrate smaller and less effective compensatory changes in their motor plans in response to the EMA sensors. Similarly, we predict after-sensor effects for both groups, reflected by increases in most measures (except for M2) following sensor removal. However, we expect these compensatory carryover effects to be smaller in magnitude for PwPD, given their auditory and somatosensory deficits, as well as their underlying dysarthria.

# Method

The data for this study were collected as part of a larger kinematic and acoustic study in the Florida State University (FSU) Motor Speech Laboratory. The procedures outlined below were approved by the Florida State University’s Institutional Review Board (FSU IRB: 00002525). Informed consent was obtained from all participants before they participated in the study.

## Speaker Data

### Speakers

A total of 34 speakers were included in the study, including 13 PwPD and 21 Control speakers. [Table 1](#tbl-speakerDemo) provides a summary of the demographic information for both groups, and detailed participant-level descriptions are available in Supplemental Tables [1](#supptbl-PDspeakerDemo) and [2](#supptbl-controlSpeakerDemo). While efforts were made to match the two groups in terms of age and sex, there were some discrepancies. The Control group included more female speakers (12 female, 9 male) and was younger on average (M = 64.33 years) compared to the PwPD group (4 female, 9 male; M = 71.00 years). The PwPD varied in years since diagnosis, ranging from 2 to 15 years (M = 7.11 years), and exhibited a wide range of hypokinetic dysarthria severity, from mild to profound. All PwPD were evaluated in their on-medication state, and none had undergone deep brain stimulation.

#### Table 1

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 1: Listener group demographics.   | **Characteristic** | **All Speakers**1 | **Control**1 | **People with Parkinson’s Disease (PwPD)**1 | | --- | --- | --- | --- | | Sex |  |  |  | | Female | 16 (47%) | 12 (57%) | 4 (31%) | | Male | 18 (53%) | 9 (43%) | 9 (69%) | | Age | 66.88 (10.99) | 64.33 (10.81) | 71.00 (10.38) | | Time Since Diagnosis (years) | - | - | 6.62 (4.48) | | Severity2 |  |  |  | | Mild | - | - | 3 (23%) | | Moderate | - | - | 5 (38%) | | Severe | - | - | 3 (23%) | | Profound | - | - | 2 (15%) | | 1 n (%); Mean (SD) | | | | | 2 Severity labels were previously determined by Thompson & Kim (2024) based on the intelligibility measures using the surrogate-severity measures outlined in Stipancic et al. (2022). Specifically, intelligibility values > 94% are ‘Normal’, 85% - 94% are ‘Mild’, 70% - 84% are ‘Moderate’, 45% - 69% are ‘Severe’, and < 45% are ‘Profound’. | | | | |

### Data Collection

Acoustic and kinematic data were collected simultaneously in a sound-attenuating booth using the Wave system and Wavefront (NDI, n.d.) with an AKG C1000S microphone placed approximately 30 cm from the speaker recording the speech stimuli. The acoustic data had a sampling rate of 20 kHz and 16-bit resolution. Kinematic data were not analyzed in the current study, as kinematic data were not available before sensor application and after sensor removal. Therefore, speech outcomes were limited to acoustic and perceptual measures. Five five-degrees-of-freedom (5DOF) sensors, 2 mm in diameter, were affixed to various articulators, including two lingual sensors, the tongue front (affixed medially 2 cm from the tongue tip) and tongue back (affixed medially 3 cm from the tongue front sensor), two labial sensors (affixed to the upper and lower vermilion lip border, respectively), and a jaw sensor (adhered to the labial surface of the central lower incisors). Additionally, a six-degrees-of-freedom (6DOF) reference sensor for head movements was attached to the bridge of a pair of glasses. The sensors were adhered to the articulators using PeriAcryl Oral Tissue Adhesive, a non-toxic dental surgical glue. All speakers completed each speech task with all six sensors attached. For a few participants, sensors may have become unattached during a speech task. However, in these cases, the sensors were reattached, and the recording was restarted.

The speakers read “The Caterpillar” passage (Patel et al., 2013) at three different time points: (1) before sensor placement (Before Sensors), (2) at least 10 minutes after sensor placement (With Sensors), allowing participants adequate time to adapt to the sensors (Dromey et al., 2018), and (3) after sensor removal (After Sensors). At each time point, speakers were instructed to read the passage in their everyday conversational voice. “The Caterpillar” was originally selected for the larger study because it was designed with motor speech disorders in mind, featuring increasing word complexity, repeated words, and prosodically demanding sentences, making it well-suited for dysarthria assessment. For the current study, we selected this passage specifically because it poses greater linguistic and prosodic demands than simple word or sentence repetition, which we theorized would make it more sensitive to sensor effects.

Following the initial recording (Before Sensors), researchers attached EMA sensors to the participants’ target articulators, a process typically lasting five to ten minutes, depending on individual differences in sensor application ease. Immediately after sensor placement, participants completed a standardized 10-minute conversational practice period with the examiner. During this practice, participants were instructed to speak as naturally as possible and ignore the presence of the sensors. Previous research suggests that this 10-minute interval provides sufficient time for speakers to adapt to sensor placement, as perceptual speech clarity ratings typically stabilize after approximately 10 minutes, with minimal additional adaptation occurring beyond this point (Dromey et al., 2018). Further support is provided by Tienkamp et al. (2024), who found that articulatory-acoustic measures after a period of sensor habituation did not differ significantly from measures taken immediately after sensor placement, suggesting limited additional adaptation beyond the initial interval.

Following this adaptation period, the kinematic experiment began, involving several speech tasks. Although multiple tasks were administered between primary recordings, these were not analyzed for the current study. However, the With Sensors reading of “The Caterpillar” passage consistently occurred within the first three tasks, except in one instance where a participant’s lingual sensor became detached during the initial reading, which required sensor reapplication and a rereading of the task later in the experiment.

The entire data collection session lasted approximately two hours, with the three readings of “The Caterpillar” passage consistently recorded in the same order. The average time between the Before Sensors and With Sensors recordings was 26.26 minutes (SD = 10.55), which included sensor application and the 10-minute practice period. While the exact duration between sensor application and the With Sensors recording was not precisely controlled across participants, the standardized 10-minute practice period ensured all speakers had an equal minimum opportunity to adapt. The time between the Before Sensors and After Sensors recordings was 56.37 minutes (SD = 7.34). Lastly, the average time between the last recording with sensors and the After Sensors recording was 5.21 minutes (SD = 2.22), which included a few minutes to remove sensors and allow participants to remove residual glue from their tongue and lips.

#### Table 2

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 2: Speaker stimuli and selected segments used to calculate articulation rate.   | Section | Passage Text1 | | --- | --- | | Section 1 | **Do you like amusement parks?** Well, I sure do. To amuse myself, I went twice last spring. My most MEMORABLE moment was riding on the Caterpillar, which is a gigantic rollercoaster high above the ground. When I saw how high the Caterpillar rose into the bright blue sky, I knew it was for me. | | Section 2 | After waiting in line for thirty minutes, I made it to the front where the man measured my height to see if I was tall enough. I gave the man my coins, asked for change, and jumped on the cart. Tick, tick, tick, the Caterpillar climbed slowly up the tracks. **It went SO high I could see the parking lot. Boy, was I SCARED!** I thought to myself, ‘There’s no turning back now.’ | | Section 3 | People were so scared they screamed as we swiftly zoomed fast, fast, and faster along the tracks. As quickly as it started, the Caterpillar came to a stop. Unfortunately, it was time to pack the car and drive home. That night I dreamt of the wild ride on the Caterpillar. **Taking a trip to the amusement park and riding on the Caterpillar was my MOST memorable moment ever!** | | 1 Bold text indicates the three segments that were used to calculate the articulation rate. | | |

### Acoustic Measures

Four acoustic measures were analyzed at each of the three time points: Before Sensors, With Sensors, and After Sensors. The target speech segments were manually segmented using the TextGrid function in Praat (Boersma & Weenink, 2021). Acoustic analyses were conducted in the R statistical environment [R Core Team (2023); Version 4.3.2] using the rPraat package [Bořil & Skarnitzl (2016); Version 1.3.2-1], which is an interface for using Praat in R.

**Articulation Rate (syl/s).** Articulation rate was calculated as the number of syllables per second measured from breath groups. Breath groups were identified in three sentences from “The Caterpillar” passage (bolded in [Table 2](#tbl-stimuli)), chosen to capture variability in length, complexity, and prosody (i.e., declarative, exclamative, and interrogative). Breath group boundaries were identified by (1) audible breaths during the sentence or (2) silent pauses longer than 150 ms.

**Articulatory-Acoustic Vowel Space (AAVS; mel2).** The AAVS was calculated based on the methods described in Whitfield & Goberman (2014) and Whitfield & Mehta (2019), with minor modifications to be more directly comparable to the methods of Tienkamp et al. (2024). Using Praat, we extracted the formant trajectory trace for the entire passage reading by generating Linear Predictive Coding (LPC) values for F1 and F2 at every five milliseconds of the passage reading (Burg method; window length = .025 s; time step = .005 s; max number of formants: 5; formant ceiling: 5000 Hz [male], 5500 Hz [female]). Then, voiceless segments were filtered out of the data to obtain only the periodic, voiced segments. Voicing was determined by identifying intervals with F0 data exceeding 20 ms (Whitfield & Mehta, 2019). Next, we applied a two-step process to filter outliers. Local outliers were removed using a median absolute deviation filter, eliminating data points exceeding 2.5 times the median absolute deviation (Tienkamp et al., 2024). For bivariate outliers, formant data were low-pass filtered at 10 Hz, and Mahalanobis distances were computed for each F1–F2 pair. Pairs exceeding 2 SDs from the centroid were excluded (Whitfield & Mehta, 2019).

The cleaned formant data were then transformed into mel values to facilitate direct comparison with findings from Tienkamp et al. (2024). To derive the AAVS, the covariance matrix of F1 and F2 was first calculated to capture both the individual variability of each formant and their co-variation. The overall spread of the data in the F1–F2 plane was then determined by computing the determinant of this matrix, known as the generalized variance. Finally, taking the square root of the generalized variance provided a measure comparable to a bivariate standard deviation. Higher AAVS values indicate more peripheral F1 and F2 values and, consequently, a larger articulatory–acoustic working space.

**Spectral Moment Coefficients (kHz).** M1 and M2 were calculated for the speaker’s production of /s/ and /ʃ/ (Forrest et al., 1988). One token of each fricative was manually segmented from “The Caterpillar” passage (“saw” for /s/ and “sure” for /ʃ/). Fricative boundaries were identified using the waveform and wideband spectrogram view in Praat. The onsets were identified at the point when high-frequency, aperiodic energy first appeared on the spectrogram. The offsets were identified at the point of the first glottal pulse of the following vowel. Consistent with the methods of Dromey et al. (2018), Praat’s “To Spectrum…” function was used to generate a spectrum from the entire fricative segment. The default spectrogram window length of 5 ms with a Gaussian window was used, resulting in a spectral estimate weighted toward the middle 50% of the fricative duration. M1, which represents the weighted average frequency of the fricative’s spectrum, was obtained using Praat’s “Get centre of gravity…” function. M2, which measures how much the frequencies deviate from the center of gravity (M1), was calculated using the “Get standard deviation…” function. The final spectral measures were expressed in kHz.

## Perceptual Data

### Listeners

A total of 79 listeners were recruited from undergraduate communication science and disorders courses at Florida State University to provide perceptual ratings. The full demographic information for the listeners is provided in Supplemental Table [3](#supptbl-listenerDemo). To summarize, the listeners were women between 18 and 23 years old, predominantly white, and not Hispanic or Latino. While listeners’ hearing was not formally tested, none reported a history of hearing or communication disorders.

### Data Collection & Perceptual Measures

To obtain perceptual ratings of intelligibility and naturalness, listeners completed an online perceptual experiment programmed using Gorilla [Anwyl-Irvine et al. (2020); <http://www.gorilla.sc/>]. Listeners heard the raw, non-intensity-normalized audio samples and rated intelligibility and naturalness using a continuous 100-point horizontally oriented visual analog scale (VAS).

We chose not to intensity normalize the audio files to preserve the natural speech characteristics of our speakers, particularly for PwPD and dysarthria. Hypophonia, or reduced vocal intensity, is a hallmark feature of hypokinetic dysarthria, and intensity normalization would have amplified speakers’ voices, which may have artificially diminished the severity of the speakers’ dysarthria. However, we controlled for recording variability by ensuring a consistent speaker-to-microphone distance and maintaining identical microphone gain settings across all recordings. We also chose not to introduce multitalker babble or background noise, as our goal was to isolate the effects of EMA sensors on perceptual judgments in unaltered speech. This approach aligns with prior studies investigating EMA sensor effects (Dromey et al., 2018; Meenakshi et al., 2014; Weismer & Bunton, 1999). With these methodological decisions, we aimed to ensure that listener ratings reflect the combined perceptual consequences of both the dysarthria and the EMA sensors, without artificial enhancements or external confounds.

The left and right ends of the VAS corresponded to rating values of 0 and 100, respectively, which were not visible to the listener. Instead, the listeners were presented with left and right endpoints labeled as “cannot understand anything” and “understand everything” for intelligibility ratings (Tjaden et al., 2014), and “highly unnatural” and “highly natural” for naturalness ratings (Anand & Stepp, 2015). Listeners always completed a block of ratings for intelligibility first, then a block of ratings for naturalness.

Prior to making the ratings in each block, listeners were provided with instructions about the perceptual measure and how to use the VAS. Based on the consent form detailing the study, the listeners were made aware that they may hear individuals with motor speech disorders. However, the group membership (Control or PwPD) and time point (Before Sensors, With Sensors, and After Sensors) were masked to the listeners.

To minimize stimuli familiarization effects, the passage recordings were split into three sections for the listeners to rate ([Table 2](#tbl-stimuli)). Additionally, the recording was only played once before listeners made their ratings. To minimize speaker familiarization effects, listeners rated half of the speakers (n = 17) on intelligibility and the other half on naturalness. Speakers were not repeated between the intelligibility and naturalness blocks. Finally, we made efforts to minimize familiarization effects statistically by entering trial order into our models as a control variable (see the [Statistical Analysis](#sec-stats) section). These procedures yielded 12–16 independent ratings per speaker for each measure (intelligibility and naturalness) at every time point (Before Sensors, With Sensors, and After Sensors).

Additionally, four previously rated recordings were randomly selected for the listener to rate again to calculate intra-listener reliability. All passage sections were presented in a randomized order. Therefore, listeners completed a total of 21 ratings in each block: 17 passage recordings (one for each of the 17 speakers) and 4 randomly selected passage sections for intra-listener reliability. The perceptual rating task took approximately 15 to 20 minutes to complete.

## Reliability

To ensure measurement reliability, the inter- and intra-measurer reliability was assessed for the acoustic measurements made by the researchers. For inter-measurer reliability, each examiner measured 20% of the other examiner’s data, and the two sets of measurements were evaluated by calculating the mean absolute error (MAE) and standard deviation (SDAE) and using intraclass correlation coefficients (ICC) through the *icc* function in the *irr* package (Gamer et al., 2019).

Inter-measurer reliability was excellent for articulation rate (MAE = .15 syl/s, SDAE = .22 syl/s; ICC(A,2) = .97), AAVS (MAE = 4.60 mel², SDAE = 12.60 mel²; ICC(A,2) = 1.00), M1 (MAE = .13 kHz, SDAE = .18 kHz; ICC(A,2) = .99), and M2 (MAE = .13 kHz, SDAE = .18 kHz; ICC(A,2) = .99). Similarly, intra-measurer reliability was assessed by having each examiner re-measure 20% of the data at least one month after the original measurement. The first and second sets of measurements were also evaluated using ICC analyses. Intra-measurer reliability was excellent for articulation rate (MAE = .05 syl/s, SDAE = .15 syl/s; ICC(A,2) = .99), AAVS (MAE = .72 mel², SDAE = 1.74 mel²; ICC(A,2) = 1.00), M1 (MAE = .09 kHz, SDAE = .16 kHz; ICC(A,2) = .99), and M2 (MAE = .09 kHz, SDAE = .16 kHz; ICC(A,2) = .99).

Intra-listener reliability was assessed for the perceptual ratings provided by listeners. During the perceptual experiment, four previously rated recordings were randomly presented a second time to each listener. Agreement between the initial and repeated ratings was evaluated using MAE, SDAE, and ICC. Reliability was excellent for intelligibility (MAE = 5.12%, SDAE = 8.48%; ICC(A,2) = .91) and good for naturalness (MAE = 12.83%, SDAE = 15.00%; ICC(A,2) = .87).

Finally, inter-listener reliability was assessed by examining the mean (M) and standard deviation (SD) of the intelligibility and naturalness ratings for each speaker and time point (Hustad et al., 2015; Thompson & Kim, 2024)[[2]](#footnote-38). Because not every listener rated every speaker, ICC was not appropriate. The SD of ratings ranged from 0.48 to 32.60 for intelligibility and 9.87 to 36.74 for naturalness. Lower SD values were observed for speakers at the extremes of the scales, while higher SD values were observed for speakers in the mid-range ([Figure 1](#fig-1)). These SD ranges are consistent with previous studies using VAS methods for perceptual ratings (Thompson & Kim, 2024).

#### Figure 1

|  |
| --- |
| Figure 1: The mean and standard deviation of ratings across speakers and time points. |

## Statistical Analysis

To address our research questions, we utilized a Bayesian hierarchical modeling approach. This approach allowed us to account for our modest sample size, quantify uncertainty in effect estimates, and interpret EMA sensor effects and after-sensor effects using descriptive probabilities. For a comprehensive overview of Bayesian mixed-effects modeling for speech data, see Nalborczyk et al. (2019). Statistical analyses were conducted in the R statistical environment [R Core Team (2023); Version 4.3.2] using Stan modeling language (Carpenter et al., 2017) via the *brms* package (version 2.21.0) (Bürkner, 2018) and the *emmeans* package (Length, 2023).

We constructed eight Bayesian mixed-effects models, one for each acoustic and perceptual outcome measure. The outcomes for the eight models were articulation rate, AAVS, M1 and M2 for /s/ and /ʃ/, and intelligibility and naturalness ratings, respectively. For each model, the interaction between speaker group (Group: Control [reference level] and PwPD) and time point (Time Point: Before Sensors [reference level], With Sensors, and After Sensors) was included as a fixed effect. Additionally, speaker sex (Sex: male [reference level] and female) and speaker age (Age) were entered as covariates into the models to try to statistically control for the differences in sex and age across the two speaker groups. Articulation rate was also entered as a covariate in all models except the one predicting articulation rate, to account for possible rate-related differences across time points. Each model included Speaker ID as a random intercept to account for speaker variability. All models were specified with weakly informative priors, meaning we assumed no effect of Group, Time Point, Group × Time Point interaction, or any of the covariates on the target measures. Specifically, the models were specified with regularizing Gaussian priors for the intercept and slope coefficients (μ = 0, σ = 100), and a Cauchy distribution for the standard deviation parameter (μ = 0, σ = 100). For each outcome measure, model specifications were adapted based on the distribution characteristics of the data:

**Articulation Rate.** Articulation rate approximated a normal distribution. Therefore, the model was built using a Gaussian family function. In addition to the general model specifications described above, the random effect structure for the articulation rate model also included random intercepts for each phrase for each speaker (see the bolded phrases in [Table 2](#tbl-stimuli)) to allow intercepts to vary across the target phrases.

**AAVS.** The distribution of the AAVS measure was lognormal; thus, a lognormal link function was used. In addition to the general model specification, passage duration was also included as a fixed effect covariate, as longer passage durations would indicate slower speech rates, which is known to produce enlarged working spaces (Tjaden & Wilding, 2004).

**Spectral Moment Measures (M1 & M2).** The spectral moment values for /s/ and /ʃ/ approximated a normal distribution. Therefore, the models were built using Gaussian distributions. Beyond the general model specification described above, no additional fixed or random effects were entered into the models.

**Perceptual Measures.** Given that the intelligibility and naturalness distributions were bounded between 0 and 100, with a clustering of values near the lower and upper limits, we rescaled these variables to a 0 to 1 range, allowing us to model the data appropriately using a Beta distribution. Additionally, Trial Order was entered into the models to control for any potential impact that familiarization and order effects had on the perceptual ratings. Finally, random effects for the perceptual models included random intercepts for passage section (see [Table 2](#tbl-stimuli)) per speaker, as well as random intercepts for Listener ID, to account for listener variability. The perceptual models used the same Gaussian and Cauchy priors as described above. Additionally, we employed gamma priors for the Beta distribution’s shape parameters with α = 1, β = .5.

Following the construction of these models, we answered our research questions by examining the pairwise comparisons between Before Sensors and With Sensors (RQ1, Sensor Effects) and Before Sensors and After Sensors (RQ2, After-Sensor Effects) for each speaker group using the *emmeans* package (Length, 2023). Additionally, interaction contrasts between groups were examined to determine whether these sensor effects (With Sensors - Before Sensors × Group) or after-sensor effects (After Sensors - Before Sensors × Group) differed between Control and PwPD groups.

The Markov chain Monte Carlo algorithm was used to implement the Bayesian models. Four sampling chains with 4000 iterations were run for each model, with a burn-in period of 1000 iterations. To assess the robustness of an effect, we report the 95% credible interval and probability of direction (pd, not to be confused with “Parkinson’s disease”) for each parameter. For the pairwise comparisons across time points, we report the 95% Highest Probability Density (HPD) interval. The 95% credible interval indicates that we are 95% confident that the true parameter value lies within the specified range. The 95% HPD interval, on the other hand, represents the range containing 95% of the most probable values of the parameter, based on the posterior distribution. The pd value reflects the proportion of the posterior distribution that falls on the same side of zero as the median (i.e., the probability that the effect is in a consistent direction). We interpret effects as robust when the 95% intervals do not include zero and the pd exceeds 95%. In practical terms, a “robust” effect is analogous to a “statistically significant” effect with p < .05 in frequentist statistics, though the underlying logic and interpretation differ.

# Results

[Table 3](#tbl-summaryMeasures) presents summary statistics for the acoustic and perceptual measures across all speakers and by speaker sex. The main findings from the Bayesian models – specifically, sensor effects (With Sensors – Before Sensors) and after-sensor effects (After Sensors – Before Sensors) – are presented in [Table 4](#tbl-pairwise) and Figures [2](#fig-articRate) – [7](#fig-Nat). Contrasts that were statistically robust and potentially meaningful are highlighted in bold in the table and indicated in the figure legends. Due to space constraints, full model summaries for the eight Bayesian models are provided in the Supplementary Information (see Supplemental Tables [4](#supptbl-articRateAAVS) – [7](#supptbl-perceptual)). All reported estimates reflect posterior distributions derived from the specified models and should be interpreted as conditional on the data and model assumptions. Data preparation, visualization, and analysis code are available on our OSF project page (<https://osf.io/n7kse/>).

#### Table 3

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 3: Summary of measures for Control speakers and PwPD across all speakers and by speaker sex.   | Measure | **Control** | | | **People with Parkinson’s disease (PwPD)** | | | | --- | --- | --- | --- | --- | --- | --- | | Before Sensors | With Sensors | After Sensors | Before Sensors | With Sensors | After Sensors | | All Speakers | | | | | | | | Articulation Rate (syl/s) | 4.73 (.86) | 4.87 (.78) | 4.98 (.69) | 4.74 (.80) | 4.80 (.73) | 4.87 (.84) | | Articulatory Acoustic Vowel Space (mel²) | 27603.88 (7643.89) | 26143.51 (7435.93) | 28194.31 (8998.68) | 18777.24 (4370.62) | 18433.20 (4395.77) | 20735.47 (4918.72) | | M1 for /s/ (kHz) | 5.61 (.79) | 4.81 (.91) | 5.46 (.74) | 4.90 (1.18) | 4.53 (1.02) | 4.97 (1.09) | | M2 for /s/ (kHz) | 2.71 (.34) | 2.82 (.32) | 2.78 (.38) | 2.71 (.31) | 2.91 (.17) | 2.77 (.38) | | M1 for /ʃ/ (kHz) | 4.09 (.55) | 3.94 (.56) | 4.25 (.66) | 3.80 (.77) | 3.74 (.70) | 4.02 (1.02) | | M2 for /ʃ/ (kHz) | 2.53 (.22) | 2.65 (.15) | 2.44 (.21) | 2.50 (.19) | 2.62 (.22) | 2.47 (.19) | | Intelligibility (VAS) | 95.54 (10.01) | 92.98 (13.01) | 96.45 (9.33) | 87.89 (21.13) | 80.16 (25.52) | 85.96 (24.41) | | Naturalness (VAS) | 72.95 (26.63) | 64.97 (27.74) | 75.30 (25.04) | 54.27 (31.79) | 43.81 (30.71) | 59.21 (30.42) | | Female Speakers | | | | | | | | Articulation Rate (syl/s) | 4.61 (.79) | 4.66 (.63) | 4.77 (.69) | 4.64 (.75) | 4.75 (.77) | 4.83 (.72) | | Articulatory Acoustic Vowel Space (mel²) | 31369.97 (7555.95) | 30340.51 (6715.33) | 33045.91 (9068.73) | 20430.07 (2681.58) | 21605.86 (1813.22) | 23667.13 (2039.21) | | M1 for /s/ (kHz) | 5.97 (.66) | 5.05 (.86) | 5.56 (.85) | 5.88 (.22) | 5.55 (.43) | 5.86 (.72) | | M2 for /s/ (kHz) | 2.80 (.39) | 2.90 (.38) | 2.91 (.40) | 2.80 (.45) | 2.95 (.15) | 3.07 (.47) | | M1 for /ʃ/ (kHz) | 4.23 (.48) | 4.11 (.52) | 4.38 (.45) | 4.23 (.52) | 4.12 (.48) | 4.70 (.40) | | M2 for /ʃ/ (kHz) | 2.46 (.23) | 2.63 (.14) | 2.43 (.25) | 2.47 (.28) | 2.45 (.21) | 2.42 (.17) | | Intelligibility (VAS) | 95.33 (11.08) | 92.97 (13.77) | 95.45 (11.52) | 92.77 (16.10) | 90.65 (17.01) | 95.98 (8.53) | | Naturalness (VAS) | 74.35 (27.85) | 64.24 (29.10) | 73.55 (26.13) | 69.00 (29.42) | 59.91 (30.24) | 76.50 (24.42) | | Male Speakers | | | | | | | | Articulation Rate (syl/s) | 4.88 (.94) | 5.14 (.89) | 5.26 (.60) | 4.78 (.83) | 4.82 (.72) | 4.90 (.90) | | Articulatory Acoustic Vowel Space (mel²) | 22582.43 (4234.24) | 20547.51 (3817.65) | 21725.52 (2641.07) | 18042.64 (4897.33) | 17023.14 (4525.66) | 19432.51 (5340.80) | | M1 for /s/ (kHz) | 5.13 (.73) | 4.49 (.92) | 5.32 (.56) | 4.47 (1.18) | 4.08 (.86) | 4.57 (1.00) | | M2 for /s/ (kHz) | 2.59 (.24) | 2.72 (.18) | 2.60 (.29) | 2.67 (.25) | 2.88 (.19) | 2.64 (.27) | | M1 for /ʃ/ (kHz) | 3.90 (.60) | 3.72 (.57) | 4.06 (.87) | 3.61 (.81) | 3.57 (.74) | 3.72 (1.08) | | M2 for /ʃ/ (kHz) | 2.61 (.19) | 2.67 (.17) | 2.46 (.17) | 2.51 (.15) | 2.69 (.19) | 2.49 (.20) | | Intelligibility (VAS) | 95.84 (8.23) | 93.00 (12.01) | 97.76 (4.94) | 85.69 (22.76) | 75.13 (27.38) | 81.50 (27.67) | | Naturalness (VAS) | 71.19 (25.02) | 65.99 (25.82) | 77.48 (23.52) | 47.65 (30.67) | 37.17 (28.47) | 51.76 (29.79) | |

## Sensor Effects

Our first research question examined the impact of sensors (i.e., With Sensors – Before Sensors) on the various acoustic and perceptual measures in PwPD and Control speakers. Additionally, we were interested in understanding if the impact of sensors was comparable between the two groups (i.e., Group × With Sensors – Before Sensors; labeled “PwPD - Control” within the table and figures). The findings are presented in the left column of [Table 4](#tbl-pairwise) and visualized in the middle panel of Figures [2](#fig-articRate)-[7](#fig-Nat) for each measure.

**Articulation Rate.** Control speakers demonstrated a robust increase in articulation rate when wearing sensors (pd = 98%), whereas PwPD speakers exhibited a smaller and less certain effect (pd = 78%; middle panel of [Figure 2](#fig-articRate)). Despite these differing patterns, the between-group difference in sensor effects was not robustly supported (pd = 76%), suggesting no clear evidence of a difference in how PwPD and Control speakers responded to wearing sensors.

**AAVS.** Control speakers had a 96% probability of a reduced AAVS with sensors compared to before sensors; however, the 95% HPD interval crossed zero, indicating uncertainty and a lack of robust evidence for this effect (middle panel of [Figure 3](#fig-AAVS)). In contrast, PwPD speakers showed even weaker evidence of reduced vowel space, with only a 71% probability. Additionally, the difference between groups in sensor effects was not robustly supported (pd = 80%), providing no clear evidence that PwPD and Control speakers responded differently to wearing sensors.

#### Figure 2

|  |
| --- |
| *Note*. (a) Posterior predictions show the predicted articulation rate at each time point (Before Sensors, With Sensors, After Sensors) for each group, adjusted for age and sex. (b) Sensor effects: the marginal effects show the change in articulation rate from Before Sensors to With Sensors for each group, along with the difference in that change between groups. (c) After-sensor effects: the marginal effects show the change from Before Sensors to After Sensors for each group, and the group difference in that change. Estimates are from a Bayesian multilevel model. Effects were considered robust if the posterior probability of direction (pd) was greater than 95% and the 95% highest probability density (HPD) interval did not include zero.  PwPD = People with Parkinson’s disease; pd = probability of direction.  Figure 2: Predicted articulation rate and sensor-related changes in speakers with and without Parkinson’s disease (PwPD). |

#### Figure 3

|  |
| --- |
| *Note*. (a) Posterior predictions show the predicted articulatory-acoustic vowel space (AAVS) at each time point (Before Sensors, With Sensors, After Sensors) for each group, adjusted for age, sex, and articulation rate. (b) Sensor effects: the marginal effects show the change in AAVS from Before Sensors to With Sensors for each group, along with the difference in that change between groups. (c) After-sensor effects: the marginal effects show the change from Before Sensors to After Sensors for each group, and the group difference in that change. Estimates are from a Bayesian multilevel model. Effects were considered robust if the posterior probability of direction (pd) was greater than 95% and the 95% highest probability density (HPD) interval did not include zero.  PwPD = People with Parkinson’s disease; pd = probability of direction.  Figure 3: Predicted articulatory-acoustic vowel space (AAVS) and sensor-related changes in speakers with and without Parkinson’s disease (PwPD). |

**M1.** For /s/, both Control and PwPD speakers showed robust reductions in M1 values when wearing sensors (100% and 98% probabilities, respectively; middle panel of [Figure 4](#fig-M1)). Although there was a high probability of a between-group difference in the magnitude of this reduction (pd = 97%), the 95% HPD interval crossed zero, indicating this difference was not robustly supported. In contrast, for /ʃ/, neither Control nor PwPD speakers showed robust sensor effects (83% and 63% probabilities, respectively). The between-group difference for /ʃ/ was similarly weak (pd = 63%), providing no evidence that the groups differed meaningfully in their response to wearing sensors.

**M2.** For /s/, there was a 96% probability that Control speakers had higher M2 values when wearing sensors compared to before sensors, but this effect was not robust due to the 95% HPD interval crossing zero (middle panel of [Figure 5](#fig-M2)). In contrast, PwPD speakers demonstrated a robust increase in M2 values with sensors (99% probability). However, there was no clear evidence that the groups differed in their response to wearing sensors. For /ʃ/, neither Control nor PwPD speakers showed robust increases in M2 with sensors (83% and 62% probabilities, respectively), and there was no evidence of a between-group difference.

**Intelligibility.** Both Control and PwPD speakers were perceived as less intelligible when wearing sensors compared to before sensors, with robust effects in each group (99% and 100% probabilities, respectively; middle panel of [Figure 6](#fig-Int)). Notably, the negative impact of sensors on intelligibility was robustly greater for PwPD compared to Control speakers (100% probability).

**Naturalness.** Both Control and PwPD speakers were perceived as robustly less natural when wearing sensors compared to before sensors (100% probability for both groups; middle panel of [Figure 7](#fig-Nat)). However, unlike intelligibility ratings, there was no clear evidence of a between-group difference in the sensor effects for naturalness (pd = 60%).

#### Table 4

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 4: Pairwise comparisons.   |  | **Research Question 1** Sensor Effects (With Sensors - Before Sensors) | | | | **Research Question 2** After-Sensor Effects (After Sensors - Before Sensors) | | | | | --- | --- | --- | --- | --- | --- | --- | --- | --- | | Group | Estimate1 | 95% HPD | | pd | Estimate2 | 95% HPD | | pd | | LL | UL | LL | UL | | Articulation Rate (syl/s) | | | | | | | | | | Control | .14 | .01 | .26 | 98% | .24 | .12 | .37 | 100% | | PwPD | .06 | -.09 | .22 | 78% | .14 | -.02 | .29 | 96% | | PwPD - Control | -.07 | -.27 | .13 | 76% | -.11 | -.31 | .10 | 84% | | Articulatory Acoustic Vowel Space (mel²) | | | | | | | | | | Control | -1438.13 | -3037.31 | 281.19 | 96% | 348.55 | -1358.32 | 2176.68 | 65% | | PwPD | -443.77 | -2132.03 | 1096.85 | 71% | 2107.60 | 339.68 | 3789.71 | 99% | | PwPD - Control | 998.23 | -1267.34 | 3335.97 | 80% | 1746.76 | -606.12 | 4104.68 | 92% | | M1 for /s/ (kHz) | | | | | | | | | | Control | -.80 | -1.09 | -.52 | 100% | -.16 | -.44 | .14 | 86% | | PwPD | -.37 | -.72 | -.02 | 98% | .06 | -.29 | .43 | 64% | | PwPD - Control | .43 | -.02 | .87 | 97% | .22 | -.23 | .67 | 83% | | M1 for /ʃ/ (kHz) | | | | | | | | | | Control | -.11 | -.34 | .11 | 83% | .21 | -.02 | .45 | 96% | | PwPD | -.04 | -.33 | .23 | 63% | .26 | -.03 | .54 | 96% | | PwPD - Control | .06 | -.29 | .42 | 63% | .05 | -.31 | .41 | 60% | | M2 for /s/ (kHz) | | | | | | | | | | Control | .10 | -.02 | .22 | 96% | .05 | -.07 | .18 | 79% | | PwPD | .20 | .04 | .34 | 99% | .06 | -.09 | .21 | 77% | | PwPD - Control | .09 | -.09 | .28 | 83% | .01 | -.19 | .20 | 52% | | M2 for /ʃ/ (kHz) | | | | | | | | | | Control | -.11 | -.33 | .12 | 83% | .21 | -.02 | .44 | 96% | | PwPD | -.04 | -.33 | .23 | 62% | .26 | -.02 | .54 | 96% | | PwPD - Control | .06 | -.31 | .41 | 64% | .05 | -.29 | .42 | 60% | | Intelligibility (VAS) | | | | | | | | | | Control | -1.13 | -2.35 | -.12 | 99% | .79 | -.15 | 1.83 | 96% | | PwPD | -5.68 | -8.83 | -3.07 | 100% | -.64 | -2.27 | .92 | 80% | | PwPD - Control | -4.54 | -7.87 | -1.67 | 100% | -1.44 | -3.33 | .32 | 95% | | Naturalness (VAS) | | | | | | | | | | Control | -9.82 | -14.42 | -5.69 | 100% | 1.64 | -2.21 | 5.87 | 80% | | PwPD | -10.82 | -16.79 | -4.73 | 100% | 6.89 | 1.28 | 12.77 | 99% | | PwPD - Control | -.98 | -8.76 | 6.33 | 60% | 5.22 | -1.72 | 12.17 | 93% | | 1 Negative estimates indicate Before Sensors > With Sensors. | | | | | | | | | | 2 Negative estimates indicate Before Sensors > After Sensors. | | | | | | | | |   *Note*. pd = probability of direction; HPD = Highest Probability Density; LL = lower limit; UL = upper limit. Bold values indicate robust effects. |

## After-Sensor Effects

Our second research question examined after‐sensor effects (i.e., After Sensors – Before Sensors) to understand the impact that wearing EMA sensors for a 45-minute recording session had on various acoustic and perceptual measures in PwPD and Control speakers. Additionally, we were interested in understanding if these effects were comparable between the two groups (i.e., Group × Before Sensors – After Sensors; labeled “PwPD - Control” within the table and figures). The findings are presented in the right column of [Table 4](#tbl-pairwise) and visualized in the right panel of Figures [2](#fig-articRate)-[7](#fig-Nat) for each measure.

**Articulation Rate.** Control speakers showed a robust increase in articulation rate following sensor removal compared to before sensors (100% probability), whereas PwPD speakers exhibited a smaller, less certain increase (pd = 96%; HPD crossing zero; right panel of [Figure 2](#fig-articRate)). The between-group difference in after-sensor effects was not robust (pd = 84%), indicating no clear evidence that PwPD and Control speakers differed meaningfully in articulation rate following sensor removal.

**AAVS.** Control speakers showed no robust difference in AAVS following sensor removal compared to before sensors (right panel of [Figure 3](#fig-AAVS)). In contrast, there was a 99% probability that PwPD speakers robustly increased their AAVS following sensor removal. However, the between-group difference in these after-sensor effects was not robustly supported (pd = 92%), suggesting no clear evidence that the groups differed meaningfully in their AAVS responses following sensor removal.

#### Figure 4

|  |
| --- |
| *Note*. (a) Posterior predictions show the predicted fricative spectral center of gravity (M1) at each time point (Before Sensors, With Sensors, After Sensors) for each group, adjusted for age and sex. (b) Sensor effects: the marginal effects show the change in M1 from Before Sensors to With Sensors for each group, along with the difference in that change between groups. (c) After-sensor effects: the marginal effects show the change from Before Sensors to After Sensors for each group, and the group difference in that change. Estimates are from a Bayesian multilevel model. Effects were considered robust if the posterior probability of direction (pd) was greater than 95% and the 95% highest probability density (HPD) interval did not include zero.  PwPD = People with Parkinson’s disease; pd = probability of direction.  Figure 4: Predicted fricative spectral center of gravity (M1) and sensor-related changes in speakers with and without Parkinson’s disease (PwPD). |

#### Figure 5

|  |
| --- |
| *Note*. (a) Posterior predictions show the predicted fricative spectral standard deviation (M2) at each time point (Before Sensors, With Sensors, After Sensors) for each group, adjusted for age and sex. (b) Sensor effects: the marginal effects show the change in M2 from Before Sensors to With Sensors for each group, along with the difference in that change between groups. (c) After-sensor effects: the marginal effects show the change from Before Sensors to After Sensors for each group, and the group difference in that change. Estimates are from a Bayesian multilevel model. Effects were considered robust if the posterior probability of direction (pd) was greater than 95% and the 95% highest probability density (HPD) interval did not include zero.  PwPD = People with Parkinson’s disease; pd = probability of direction.  Figure 5: Predicted fricative spectral standard deviation (M2) and sensor-related changes in speakers with and without Parkinson’s disease (PwPD). |

**M1.** For /s/, neither the Control speakers nor PwPD exhibited robust after-sensor effects, and the between-group difference in after-sensor effects was also not robust (right panel of [Figure 4](#fig-M1)). Similarly, for /ʃ/, neither Control speakers nor PwPD showed robust after-sensor effects, and the between-group difference for /ʃ/ after-sensor effects was likewise not robust.

**M2.** For /s/, neither the Control speakers nor PwPD showed robust after-sensor effects, and the between-group difference in after-sensor effects was similarly not robust (right panel of [Figure 5](#fig-M2)). For /ʃ/, neither Control speakers nor PwPD exhibited robust after-sensor effects, and the between-group difference in after-sensor effects was not robust.

**Intelligibility.** There was a 96% probability that Control speakers increased intelligibility ratings following sensor removal compared to before sensors, though this effect was uncertain and not robust (right panel of [Figure 6](#fig-Int)). Similarly, there was an 80% probability that PwPD speakers increased intelligibility ratings following sensor removal, but this effect was also not robustly supported. Furthermore, the between-group difference was not robust (pd = 95%), indicating no clear evidence that PwPD and Control speakers differed meaningfully in their intelligibility after sensor removal.

**Naturalness.** Control speakers showed no robust difference in naturalness following sensor removal (pd = 80%; right panel of [Figure 7](#fig-Nat)). In contrast, there was a 99% probability that PwPD speakers were perceived as more natural after sensors compared to before sensors, indicating a robust improvement. However, the between-group difference in these after-sensor effects was not robust (pd = 93%), providing no clear evidence that PwPD and Control speakers differed meaningfully in their naturalness ratings following sensor removal.

#### Figure 6

|  |
| --- |
| *Note*. (a) Posterior predictions show the predicted intelligibility ratings at each time point (Before Sensors, With Sensors, After Sensors) for each group, adjusted for age and sex. (b) Sensor effects: the marginal effects show the change in ratings from Before Sensors to With Sensors for each group, along with the difference in that change between groups. (c) After-sensor effects: the marginal effects show the change from Before Sensors to After Sensors for each group, and the group difference in that change. Estimates are from a Bayesian multilevel model. Effects were considered robust if the posterior probability of direction (pd) was greater than 95% and the 95% highest probability density (HPD) interval did not include zero.  PwPD = People with Parkinson’s disease; pd = probability of direction.  Figure 6: Predicted intelligibility ratings and sensor-related changes in speakers with and without Parkinson’s disease (PwPD). |

#### Figure 7

|  |
| --- |
| *Note*. (a) Posterior predictions show the predicted naturalness ratings at each time point (Before Sensors, With Sensors, After Sensors) for each group, adjusted for age and sex. (b) Sensor effects: the marginal effects show the change in ratings from Before Sensors to With Sensors for each group, along with the difference in that change between groups. (c) After-sensor effects: the marginal effects show the change from Before Sensors to After Sensors for each group, and the group difference in that change. Estimates are from a Bayesian multilevel model. Effects were considered robust if the posterior probability of direction (pd) was greater than 95% and the 95% highest probability density (HPD) interval did not include zero.  PwPD = People with Parkinson’s disease; pd = probability of direction.  Figure 7: Predicted naturalness ratings and sensor-related changes in speakers with and without Parkinson’s disease (PwPD). |

# Discussion

This study evaluated the impact of EMA sensors on various acoustic and perceptual speech outcomes in PwPD and Control speakers, with an additional goal of determining whether sensor presence affects both groups equally. We also examined after-effects following sensor removal to identify potential carry-over effects from sensor wear. Three major findings emerged: (1) sensor effects were observed in fricative /s/ production and perceptual ratings of intelligibility and naturalness; (2) after-sensor effects were observed for acoustic working space and ratings of naturalness for PwPD; and (3) sensor and after-sensor effects were largely similar across speaker groups, with the exception of the intelligibility ratings. In the following sections, we contextualize these findings within previous research and discuss their potential implications.

## Sensors Effects May Impact Sibilants, Intelligibility, and Naturalness

To examine sensor effects, we explored the contrast between the Before Sensors and With Sensors time points across the various acoustic and perceptual measures. We found meaningful sensor effects for the spectral moments of /s/ and perceptual ratings of intelligibility and naturalness. For articulation rate, Control speakers robustly increased their rate with sensors on, whereas PwPD speakers showed a similar, though less certain and non-robust, effect. However, the posterior predictions in the left panel of [Figure 2](#fig-articRate) reveal that articulation rate increased with each subsequent reading of the passage, regardless of sensor status. This pattern suggests that the increase is more likely due to passage familiarization than to a sensor-induced effect, as even larger increases were observed between the Before Sensors and After Sensors time points when speakers were most familiar with the passage. Thus, it is unlikely that EMA sensors induce a rate-increasing effect.

For fricatives, both PwPD and Control speakers produced lower M1 values for /s/ with sensors on compared to before sensors, consistent with Dromey et al. (2018). The magnitude of this change was similar for both speaker groups as indicated by non-robust interaction effects. These findings suggest that the presence of the sensors, particularly the tongue front sensor, obstructed the lingual-alveolar constriction required for /s/ production. Specifically, the tongue front sensor, located approximately 2 cm from the tongue tip, likely acted as a mechanical obstruction, elongating the front cavity and shifting the constriction location, resulting in an acoustic profile more similar to /ʃ/, as reflected by the lower M1 values. We also observed increased spectral moment variability while wearing sensors, particularly in PwPD, as indicated by increased M2 values for /s/; a smaller, non-robust increase was observed in Control speakers. Unlike M1 changes, the articulatory correlates of M2 are less understood. However, one possibility is that the sensor wires obstructed the air stream and increased the spectral variability. This could explain why both the present study and Dromey et al. (2018) observed sensor effects for M2, while Weismer & Bunton (1999), who used wireless gold pellets, did not. Taken together, these results suggest that EMA sensors, and their associated wires, meaningfully impact the acoustic signal for fricative sibilants, particularly /s/. Moreover, even after a 10-minute practice period, speakers cannot fully compensate for these disruptions during fricative production.

Our analysis did not reveal robust sensor effects on AAVS, which contrasts with recent findings showing reduced AAVS when sensors are worn (Tienkamp et al., 2024). Several factors may explain why our findings differ from Tienkamp et al. (2024). First, they had a larger sample size, particularly for PwPD, which could increase the power to detect sensor effects. Second, the studies involved different languages. Our participants spoke American English, whereas Tienkamp et al. (2024) examined Dutch speakers. Language-specific characteristics might influence AAVS in ways that complicate direct comparisons. Third, articulation rate increased across subsequent readings in our study, potentially contributing to a more centralized vowel space (Turner et al., 1995; Weismer et al., 2000). In our study, we aimed to control for this by including articulation rate as a covariate in our AAVS model. However, although Tienkamp et al. (2024) also employed repeated readings, they did not control for articulation or speech rate, which could have confounded their findings. Further research is needed to clarify how EMA sensors influence the working space across various languages and speaking tasks.

In our sample, EMA sensors negatively impacted listener perceptions of speech intelligibility and naturalness in both PwPD and Control speakers. Regarding intelligibility, our findings align with Meenakshi et al. (2014), as both studies found that EMA sensors reduce intelligibility despite differences in measurement approaches. Similarly, while Dromey et al. (2018) did not assess intelligibility or naturalness, their finding of reduced perceptual judgements of articulatory precision with sensors is consistent with our results. These findings suggest that even after a 10-minute adaptation period, EMA sensors introduce perceptual degradation. Researchers using EMA data for perceptual ratings should consider that intelligibility and naturalness may be affected. Rather than viewing this as a limitation of EMA, we emphasize it as a methodological factor that should be accounted for in study designs. To mitigate these effects, we recommend collecting speech samples for perceptual ratings before applying EMA sensors.

## After-Sensor Effects May Impact Articulatory-Acoustic Vowel Space and Naturalness Ratings

To examine after-sensor effects, we compared acoustic and perceptual measures between the Before Sensors and After Sensors time points. For most measures, there was no clear evidence of after-sensor effects. For articulation rate, Control speakers robustly increased their rate between the Before Sensors and After Sensors time points, whereas PwPD speakers showed a similar, though less certain and non-robust, effect. However, like the articulation rate findings for sensor effects, this rate increase is likely the result of passage familiarization rather than true after-sensor effects. Future research investigating after-effects should account for such familiarization effects, as they may influence certain speech outcome measures like articulation rate.

In our data, we observed that PwPD increased their AAVS by 10.63% between the Before Sensors and After Sensors time points, even after controlling for the increases in articulation rate. The finding that PwPD increased AAVS following sensor removal, while Control speakers did not, is difficult to account for within the current study design. One potential explanation, though speculative, relates to cognitive resources. PwPD are known to experience cognitive deficits even at early stages of the disease, and these cognitive deficits are associated with impaired motor performance (Monastero et al., 2018). Therefore, it is likely that with each subsequent reading of the passage, the speakers became increasingly familiar with the speech task, thus reducing its cognitive demand. In addition to passage familiarity, the continuous somatosensory perturbation introduced by the sensors likely drew attentional resources away from speech planning, thus adding a layer of cognitive complexity. For PwPD, these demands may have been especially burdensome. By the After Sensors time point, speakers were most familiar with the task and, in the case of PwPD, were no longer contending with the added cognitive load of sensor interference. This combination may have enabled a more efficient motor performance, resulting in an increased acoustic working space for PwPD relative to their Before Sensors reading of the passage. Future research is needed to further explore the interaction between cognitive load and sensor effects in PwPD.

Additionally, PwPD were rated as 11.12% more natural at the After Sensors time point compared to Before Sensors. This perceptual improvement may be a consequence of the increased AAVS, aligning with Whitfield & Goberman (2014), who found that increased vowel space was associated with perceptual gains in PwPD (i.e., speech clarity improvements). However, after-sensor effects were not observed for intelligibility ratings, suggesting that while the articulatory adjustments led to increased naturalness, they did not meaningfully enhance intelligibility. One possible explanation is that, with repeated readings, PwPD became more familiar with the passage and began producing more natural suprasegmental features—such as improved prosody, pitch variation, or vocal intensity—that were not captured by the acoustic measures in this study (Klopfenstein, 2015). Naturalness ratings may have been more sensitive to these global prosodic changes than intelligibility ratings. In contrast, Control speakers, who did not show the same improvement in naturalness from Before Sensors to After Sensors, likely had intact prosody at baseline and thus had less room for perceptual gains through practice.

It remains unclear how long these articulatory and perceptual benefits of after-sensor effects may last. In our study, the after-sensor recording was conducted immediately after sensor removal, which took an average of 5.21 minutes. Research on motor adaptation suggests that while return to baseline performance (i.e., de-adaptation) is gradual, it generally occurs more quickly than the initial adaptation process (Bastian, 2008; Davidson & Wolpert, 2004). However, research in sensorimotor adaptation in response to altered auditory feedback suggests that the rate of de-adaptation can vary, and may not be faster than the rate of adaptation (Kitchen et al., 2022; Mitsuya et al., 2011). Dromey et al. (2018) found that after a 10-minute adaptation period to EMA sensors, no further perceptual or acoustic improvements were observed. Based on this, it is reasonable to hypothesize that after-sensor effects may not persist beyond 10 minutes. However, future research should systematically investigate the duration of after-sensor effects using methods similar to those employed in Dromey et al. (2018) to assess sensor effects.

## Effects Were Similar Across Speaker Groups, Except for Intelligibility

Our study aimed to determine whether sensor effects and after-sensor effects differentially impacted PwPD and Control speakers. To investigate potential group differences, we examined interactions between speaker group and the contrasts of interest (RQ1: With Sensors - Before Sensors × Group; RQ2: After Sensors - Before Sensors × Group; labeled “PwPD - Control” within the table and figures). Among all the measures analyzed, the only robust group difference emerged in the sensor effects on intelligibility.

Although both groups experienced reduced intelligibility while wearing sensors, the magnitude of this effect was robustly greater for PwPD. Control speakers showed only an estimated 1.14% decrease in intelligibility ratings with sensors, whereas PwPD showed an estimated 5.81% decrease. This suggests that PwPD are more susceptible to sensor effects when it comes to intelligibility. Interestingly, PwPD and Control speakers did not differ in how sensors impacted articulatory working space (AAVS) or fricative production (M1 & M2), which contradicts our initial hypothesis that PwPD would be less, or differentially affected by EMA sensors. This indicates that, although sensors affect articulation similarly across groups, the perceptual consequences are more pronounced for PwPD. One possible explanation is that the interaction between sensor effects and dysarthria has a compounding impact on intelligibility, amplifying the perceptual deficits already present in PwPD.

The finding that sensor effects for intelligibility ratings were more pronounced for PwPD compared to Control speakers has important methodological implications. Specifically, studies using acoustic data recorded with EMA sensors to obtain perceptual measures should anticipate some degree of perceptual degradation caused by the sensors, particularly for PwPD with hypokinetic dysarthria, who appear to be more affected than neurologically healthy speakers. Researchers should interpret such measurements with caution. Alternatively, to ensure that perceptual judgments reflect dysarthria rather than a combination of dysarthria and sensor effects, researchers could collect speech samples for perceptual analysis prior to sensor placement.

Finally, the current study was designed to examine sensor- and after-sensor effects at a group level. However, individual speaker variability within both groups suggests that some individuals may be more sensitive to sensor presence (see individual trend lines in Supplemental Figure [1](#suppfig-rawData)). One possibility is that speakers vary in their auditory and/or somatosensory acuity, which may influence how strongly they are perturbed by the sensors and how successfully they adapt to them. Since EMA sensors function as somatosensory perturbations that can also alter the acoustic speech signal, future research should aim to disentangle the relative contributions of auditory and somatosensory feedback in driving compensatory speech behaviors.

## Limitations and Future Directions

There are a few limitations to note in this study. First, the sample size was modest, and the two groups were not perfectly matched for age and sex. Although statistical controls were applied to account for these differences, future research should aim for more balanced group designs to minimize potential confounding effects. In addition, because the study focused primarily on older adults, reflecting the clinical population of interest, the observed sensor effects may not generalize to younger speakers. Future work should examine how age influences responses to sensors.

Second, this study utilizes data collected as part of a broader investigation into perceptual, acoustic, and kinematic characteristics in PwPD and control speakers. As such, the design lacked the level of experimental control typically dedicated to perturbation studies. For instance, we did not collect multiple baseline, adaptation, or wash-out trials, which are commonly used to confirm effects and track behavioral changes over time. This limits our ability to precisely characterize the time course of adaptation and recovery. Additionally, PwPD were tested in their medication-on state to align with the applied and ecologically valid focus of the larger study on dysarthria as it presents in everyday communication. However, different findings may have emerged had we tested speakers in the medication-off state, as is sometimes done in perturbation studies (Hammer & Barlow, 2010; Mollaei et al., 2016; but see Chen & Watson, 2017).

Third, the nature of EMA sensor perturbations introduces inherent variability that differs from other perturbation paradigms, such as auditory or mechanical jaw perturbations. EMA sensor application and removal are time-consuming and can vary between participants, making it difficult to standardize adaptation and post-adaptation periods. For example, because the sensor application process differed across participants, the amount of time each speaker wore the sensors at the With Sensors time point varied, although all received a minimum of 10 minutes to adapt. Similarly, between the final recording with sensors and the After Sensors recording, participants underwent the sensor removal process, which involved variable durations (M = 5.21 minutes; SD = 2.22) and differing levels of verbal interaction with research staff. As a result, the After Sensors recording may reflect different stages of de-adaptation across participants. Therefore, observed after-sensor effects should be interpreted with caution.

Finally, a key motivation for this study was to explore how somatosensory deficits in PwPD might influence their ability to compensate for the perturbations introduced by EMA sensors. However, we did not collect direct measures of somatosensory function. Assessing somatosensory abilities could provide a clearer understanding of individual differences in adaptation to EMA sensors. It is reasonable to hypothesize that PwPD with more severe somatosensory deficits may be less capable of adapting to and compensating for EMA sensors.

# Conclusion

The purpose of this study was to examine the impact of EMA sensors on speech production and perception in PwPD and Control speakers, with a focus on sensor effects and after-sensor effects. Our findings indicate that EMA sensors primarily affected sibilant fricative production and perceptual ratings of intelligibility and naturalness in both groups. However, PwPD experienced a greater decline in intelligibility ratings when wearing sensors compared to Control speakers. Notably, at least five minutes after sensor removal, PwPD demonstrated enlarged AAVSs and were perceived to be more natural compared to before sensor application. We speculate that these after-sensor effects may reflect reduced cognitive load as PwPD became more familiar with the task, though this interpretation requires further empirical testing. Overall, these findings highlight important methodological considerations for interpreting EMA data collected from both PwPD and Control speakers.

# Acknowledgments

This study was supported by two NIDCD grants: an F31 awarded to A. Thompson (NIH DC020121) and an R01 awarded to Y-J. Kim (NIH DC020468). The Korea Health Industry Development Institute (KHIDI) grant awarded to Y-J. Kim also partly supported the study (HI22 C0736). We would also like to thank Mia Carter, a master’s student and research assistant in the Motor Speech Disorders Lab at FSU, for her assistance with acoustic segmentation for this project.

# Data Availability Statement

Deidentified datasets, data preprocessing scripts, data analysis scripts, and other supplementary materials relevant to this study are available on this study’s project page on the Open Science Framework (OSF) platform (<https://osf.io/n7kse/>).

# References

Abur, D., Lester-Smith, R. A., Daliri, A., Lupiani, A. A., Guenther, F. H., & Stepp, C. E. (2018). Sensorimotor adaptation of voice fundamental frequency in Parkinson’s disease. *PLOS ONE*, *13*(1), e0191839. <https://doi.org/10.1371/journal.pone.0191839>

Anand, S., & Stepp, C. E. (2015). Listener perception of monopitch, naturalness, and intelligibility for speakers with parkinson’s disease. *Journal of Speech, Language, and Hearing Research*, *58*(4), 1134–1144. <https://doi.org/10.1044/2015_JSLHR-S-14-0243>

Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods*, *52*(1), 388–407. <https://doi.org/10.3758/s13428-019-01237-x>

Bang, Y.-I., Min, K., Sohn, Y. H., & Cho, S.-R. (2013). Acoustic characteristics of vowel sounds in patients with Parkinson disease. *NeuroRehabilitation*, *32*(3), 649–654. <https://doi.org/10.3233/NRE-130887>

Bastian, A. J. (2008). Understanding sensorimotor adaptation and learning for rehabilitation. *Current Opinion in Neurology*, *21*(6), 628–633. <https://doi.org/10.1097/WCO.0b013e328315a293>

Berry, J. J. (2011). Accuracy of the NDI wave speech research system. *Journal of Speech, Language, and Hearing Research*, *54*(5), 1295–1301. <https://doi.org/10.1044/1092-4388(2011/10-0226)>

Boersma, P., & Weenink, D. (2021). *PRAAT: Doing phonetics by computer*.

Bořil, T., & Skarnitzl, R. (2016). *Tools rPraat and mPraat* (P. Skarnitzl, A. Horák, I. Kopeček, & K. Pala, Eds.; pp. 367–374). Springer International Publishing.

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

Brunner, J., & Hoole, P. (2012). Motor equivalent strategies in the production of German /∫/ under perturbation. *Language and Speech*, *55*(4), 457–476. <https://doi.org/10.1177/0023830911434098>

Buch, E. R., Young, S., & Contreras-Vidal, J. L. (2003). Visuomotor adaptation in normal aging. *Learning & Memory*, *10*(1), 55–63. <https://doi.org/10.1101/lm.50303>

Bultitude, J. H., Rafal, R. D., & Tinker, C. (2012). Moving forward with prisms: Sensory-motor adaptation improves gait initiation in parkinson’s disease. *Frontiers in Neurology*, *3*. <https://doi.org/10.3389/fneur.2012.00132>

Bürkner, P.-C. (2018). Advanced Bayesian Multilevel Modeling with the R Package brms. *The R Journal*, *10*(1), 395–411. <https://journal.r-project.org/archive/2018/RJ-2018-017/index.html>

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P., & Riddell, A. (2017). Stan: A Probabilistic Programming Language. *Journal of Statistical Software*, *76*, 1–32. <https://doi.org/10.18637/jss.v076.i01>

Chen, Y.-W., & Watson, P. J. (2017). Speech production and sensory impairment in mild parkinson’s disease. *The Journal of the Acoustical Society of America*, *141*(5), 3030–3041. <https://doi.org/10.1121/1.4980138>

Chu, S. Y., Barlow, S. M., Lee, J., & Wang, J. (2020). Effects of utterance rate and length on the spatiotemporal index in parkinson’s disease. *International Journal of Speech-Language Pathology*, *22*(2), 141151. <https://doi.org/10.1080/17549507.2019.1622781>

Contreras-Vidal, J. L., & Buch, E. R. (2003). Effects of Parkinson’s disease on visuomotor adaptation. *Experimental Brain Research*, *150*(1), 25–32. <https://doi.org/10.1007/s00221-003-1403-y>

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

Darley, F. L., Aronson, A. E., & Brown, J. R. (1969b). Clusters of deviant speech dimensions in the dysarthrias. *Journal of Speech and Hearing Research*, *12*(3), 462–496. <https://doi.org/10.1044/jshr.1203.462>

Davidson, P. R., & Wolpert, D. M. (2004). Scaling down motor memories: De-adaptation after motor learning. *Neuroscience Letters*, *370*(2-3), 102–107. <https://doi.org/10.1016/j.neulet.2004.08.003>

De Bodt, M. S., Hernández-Díaz Huici, M. E., & Van De Heyning, P. H. (2002). Intelligibility as a linear combination of dimensions in dysarthric speech. *Journal of Communication Disorders*, *35*(3), 283–292. <https://doi.org/10.1016/S0021-9924(02)00065-5>

Dromey, C., Hunter, E., & Nissen, S. L. (2018). Speech adaptation to kinematic recording sensors: Perceptual and acoustic findings. *Journal of Speech, Language, and Hearing Research*, *61*(3), 593–603. <https://doi.org/10.1044/2017_JSLHR-S-17-0169>

Duffy, J. R. (2020). *Motor speech disorders: Substrates, differential diagnosis, and management* (4th ed.). Elsevier.

Forrest, K., Weismer, G., Milenkovic, P., & Dougall, R. N. (1988). Statistical analysis of word-initial voiceless obstruents: Preliminary data. *The Journal of the Acoustical Society of America*, *84*(1), 115–123. <https://doi.org/10.1121/1.396977>

Gamer, M., Lemon, J., Fellows, I., & Singh, P. (2019). *Irr: Various coefficients of interrater reliability and agreement*. <https://bioconductor.statistik.tu-dortmund.de/cran/web/packages/irr/irr.pdf>

Hamlet, S. L., & Stone, M. (1976). Compensatory vowel characteristics resulting from the presence of different types of experimental dental prostheses. *Journal of Phonetics*, *4*(3), 199–218. <https://doi.org/10.1016/S0095-4470(19)31244-6>

Hammer, M. J., & Barlow, S. M. (2010). Laryngeal somatosensory deficits in Parkinson’s disease: implications for speech respiratory and phonatory control. *Experimental Brain Research*, *201*(3), 401–409. <https://doi.org/10.1007/s00221-009-2048-2>

Hardcastle, W. J., Gibbon, F. E., & Jones, W. (1991). Visual display of tongue-palate contact: Electropalatography in the assessment and remediation of speech disorders. *British Journal of Disorders of Communication*, *26*(1), 4174. <https://doi.org/10.3109/13682829109011992>

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), 327643. <https://doi.org/10.1155/1999/327643>

Hughes, O. M., & Abbs, J. H. (1976). Labial-mandibular coordination in the production of speech: implications for the operation of motor equivalence. *Phonetica*, *33*(3), 199–221. <https://doi.org/10.1159/000259722>

Hustad, K. C., Oakes, A., & Allison, K. (2015). Variability and diagnostic accuracy of speech intelligibility scores in children. *Journal of Speech, Language, and Hearing Research*, *58*(6), 1695–1707. <https://doi.org/10.1044/2015_JSLHR-S-14-0365>

Jongman, A., Wayland, R., & Wong, S. (2000). Acoustic characteristics of english fricatives. *The Journal of the Acoustical Society of America*, *108*(3), 1252–1263. <https://doi.org/10.1121/1.1288413>

Kearney, E., Giles, R., Haworth, B., Faloutsos, P., Baljko, M., & Yunusova, Y. (2017). Sentence-level movements in parkinson’s disease: Loud, clear, and slow speech. *Journal of Speech, Language, and Hearing Research*, *60*(12), 3426–3440. <https://doi.org/10.1044/2017_JSLHR-S-17-0075>

Kim, Y., Kent, R. D., & Thompson, A. (2024). *Instrumental analysis of articulation* (M. J. Ball, N. Müller, & E. Spencer, Eds.; 2nd ed., pp. 505–522). John Wiley & Sons.

Kitchen, N. M., Kim, K. S., Wang, P. Z., Hermosillo, R. J., & Max, L. (2022). Individual sensorimotor adaptation characteristics are independent across orofacial speech movements and limb reaching movements. *Journal of Neurophysiology*, *128*(3), 696–710. <https://doi.org/10.1152/jn.00167.2022>

Klein, H. B., McAllister, B. T., Davidson, L., & Grigos, M. I. (2013). A multidimensional investigation of children’s /r/ productions: Perceptual, ultrasound, and acoustic measures. *American Journal of Speech-Language Pathology*, *22*(3), 540–553. <https://doi.org/10.1044/1058-0360(2013/12-0137)>

Klopfenstein, M. (2015). Relationship between acoustic measures and speech naturalness ratings in parkinson’s disease: A within-speaker approach. *Clinical Linguistics & Phonetics*, *29*(12), 938–954. <https://doi.org/10.3109/02699206.2015.1081293>

Koenig, L. L., Shadle, C. H., Preston, J. L., & Mooshammer, C. R. (2013). Toward improved spectral measures of /s/: Results from adolescents. *Journal of Speech, Language, and Hearing Research*, *56*(4), 1175–1189. <https://doi.org/10.1044/1092-4388(2012/12-0038)>

Lam, J., & Tjaden, K. (2016). Clear speech variants: An acoustic study in parkinson’s disease. *J Speech Lang Hear Res*, *59*(4), 631–646. <https://doi.org/10.1044/2015_JSLHR-S-15-0216>

Lametti, D. R., Nasir, S. M., & Ostry, D. J. (2012). Sensory preference in speech production revealed by simultaneous alteration of auditory and somatosensory feedback. *Journal of Neuroscience*, *32*(27), 9351–9358. <https://doi.org/10.1523/JNEUROSCI.0404-12.2012>

Lee, J., Littlejohn, M. A., & Simmons, Z. (2017). Acoustic and tongue kinematic vowel space in speakers with and without dysarthria. *Int J Speech Lang Pathol*, *19*(2), 195–204. <https://doi.org/10.1080/17549507.2016.1193899>

Length, R. (2023). *Emmeans: Estimated marginal means, aka least-squares means*. <https://CRAN.R-project.org/package=emmeans.>

Martin, T. A., Keating, J. G., Goodkin, H. P., Bastian, A. J., & Thach, W. T. (1996). Throwing while looking through prisms: II. Specificity and storage of multiple gazethrow calibrations. *Brain*, *119*(4), 1199–1211. <https://doi.org/10.1093/brain/119.4.1199>

Masapollo, M., & Nittrouer, S. (2023). Interarticulator speech coordination: timing is of the essence. *Journal of Speech, Language, and Hearing Research*, *66*(3), 901–915. <https://doi.org/10.1044/2022_JSLHR-22-00594>

McAuliffe, M. J., Ward, E. C., & Murdoch, B. E. (2006a). Speech production in parkinson’s disease: I. An electropalatographic investigation of tongue-palate contact patterns. *Clinical Linguistics & Phonetics*, *20*(1), 1–18. <https://doi.org/10.1080/02699200400001044>

McAuliffe, M. J., Ward, E. C., & Murdoch, B. E. (2006b). Speech production in parkinson’s disease: II. Acoustic and electropalatographic investigation of sentence, word and segment durations. *Clinical Linguistics & Phonetics*, *20*(1), 19–33. <https://doi.org/10.1080/0269-9200400001069>

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)>

Meenakshi, N., Yarra, C., Yamini, B. K., & Ghosh, P. K. (2014). *Interspeech 2014*. 935–939. <https://doi.org/10.21437/Interspeech.2014-243>

Mefferd, A. S. (2015). Articulatory-to-acoustic relations in talkers with dysarthria: A first analysis. *Journal of Speech, Language, and Hearing Research*, *58*(3), 576–589. <https://doi.org/10.1044/2015_JSLHR-S-14-0188>

Mefferd, A. S., & Dietrich, M. S. (2019). Tongue- and jaw-specific articulatory underpinnings of reduced and enhanced acoustic vowel contrast in talkers with parkinson’s disease. *Journal of Speech, Language, and Hearing Research*, *62*(7), 2118–2132. <https://doi.org/10.1044/2019_JSLHR-S-MSC18-18-0192>

Miller, H. E., Kearney, E., Nieto, -C. A., Falsini, R., Abur, D., Acosta, A., Chao, S.-C., Dahl, K. L., Franken, M., Heller, M. E. S., Mollaei, F., Niziolek, C. A., Parrell, B., Perrachione, T., Smith, D. J., Stepp, C. E., Tomassi, N., & Guenther, F. H. (2023). Do not cut off your tail: A mega-analysis of responses to auditory perturbation experiments. *Journal of Speech, Language, and Hearing Research*, *66*(11), 4315–4331. <https://doi.org/10.1044/2023_JSLHR-23-00315>

Mitsuya, T., MacDonald, E. N., Purcell, D. W., & Munhall, K. G. (2011). A cross-language study of compensation in response to real-time formant perturbation. *The Journal of the Acoustical Society of America*, *130*(5), 2978–2986. <https://doi.org/10.1121/1.3643826>

Mollaei, F., Shiller, D. M., Baum, S. R., & Gracco, V. L. (2016). Sensorimotor control of vocal pitch and formant frequencies in parkinson’s disease. *Brain Research*, *1646*, 269–277. <https://doi.org/10.1016/j.brainres.2016.06.013>

Mollaei, F., Shiller, D. M., & Gracco, V. L. (2013). Sensorimotor adaptation of speech in Parkinson’s disease. *Movement Disorders*, *28*(12), 1668–1674. <https://doi.org/10.1002/mds.25588>

Monastero, R., Cicero, C. E., Baschi, R., Davì, M., Luca, A., Restivo, V., Zangara, C., Fierro, B., Zappia, M., & Nicoletti, A. (2018). Mild cognitive impairment in Parkinson’s disease: the Parkinson’s disease cognitive study (PACOS). *Journal of Neurology*, *265*(5), 1050–1058. <https://doi.org/10.1007/s00415-018-8800-4>

Moya-Galé, G., & Levy, E. S. (2019). Parkinson’s disease-associated dysarthria: Prevalence, impact and management strategies. *Research and Reviews in Parkinsonism*, *Volume 9*, 9–16. <https://doi.org/10.2147/jprls.S168090>

Nalborczyk, L., Batailler, C., L, œvenbruck H., Vilain, A., & B, ürkner P.-C. (2019). An introduction to bayesian multilevel models using brms: A case study of gender effects on vowel variability in standard indonesian. *Journal of Speech, Language, and Hearing Research*, *62*(5), 1225–1242. <https://doi.org/10.1044/2018_JSLHR-S-18-0006>

Nasir, S. M., & Ostry, D. J. (2006). Somatosensory precision in speech production. *Current Biology*, *16*(19), 1918–1923. <https://doi.org/10.1016/j.cub.2006.07.069>

NDI. (n.d.). *Wave system (NDI, canada).* Northern Digital Inc.

Patel, R., Connaghan, K., Franco, D., Edsall, E., Forgit, D., Olsen, L., Ramage, L., Tyler, E., & Russell, S. (2013). "The caterpillar": A novel reading passage for assessment of motor speech disorders. *Am J Speech Lang Pathol*, *22*(1), 1–9. <https://doi.org/10.1044/1058-0360(2012/11-0134)>

Perkell, J. S., Matthies, M. L., Svirsky, M. A., & Jordan, M. I. (1993). Trading relations between tongue-body raising and lip rounding in production of the vowel /u/: A pilot ‘‘motor equivalence’’ study. *The Journal of the Acoustical Society of America*, *93*(5), 2948–2961. <https://doi.org/10.1121/1.405814>

Perrier, P., & Fuchs, S. (2015). *Motor equivalence in speech production* (pp. 223–247). John Wiley & Sons, Ltd. <https://doi.org/10.1002/9781118584156.ch11>

Petrović, B. T. (2020). Spectral moments of fricative consonants in serbian (an account of female speakers’ production). *Филолог Часопис За Језик, Књижевност и Културу*, *21*, 95–133. <https://www.ceeol.com/search/article-detail?id=888874>

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>

Purcell, D. W., & Munhall, K. G. (2006). Adaptive control of vowel formant frequency: Evidence from real-time formant manipulation. *The Journal of the Acoustical Society of America*, *120*(2), 966–977. <https://doi.org/10.1121/1.2217714>

R Core Team. (2023). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>

Rebernik, T., Jacobi, J., Jonkers, R., Noiray, A., & Wieling, M. (2021). A review of data collection practices using electromagnetic articulography. *Laboratory Phonology*, *12*(1). <https://doi.org/10.5334/labphon.237>

Roemmich, R. T., Nocera, J. R., Stegemöller, E. L., Hassan, A., Okun, M. S., & Hass, C. J. (2014). Locomotor adaptation and locomotor adaptive learning in parkinson’s disease and normal aging. *Clinical Neurophysiology : Official Journal of the International Federation of Clinical Neurophysiology*, *125*(2), 313–319. <https://doi.org/10.1016/j.clinph.2013.07.003>

Rong, P., Loucks, T., Kim, H., & Hasegawa-Johnson, M. (2012). Relationship between kinematics, F2 slope and speech intelligibility in dysarthria due to cerebral palsy. *Clinical Linguistics & Phonetics*, *26*(9), 806822. <https://doi.org/10.3109/02699206.2012.706686>

Savariaux, C., Badin, P., Samson, A., & Gerber, S. (2017). A comparative study of the precision of carstens and northern digital instruments electromagnetic articulographs. *Journal of Speech, Language, and Hearing Research*, *60*(2), 322–340. <https://doi.org/10.1044/2016_JSLHR-S-15-0223>

Schneider, J. S., Diamond, S. G., & Markham, C. H. (1986). Deficits in orofacial sensorimotor function in Parkinson’s disease. *Annals of Neurology*, *19*(3), 275–282. <https://doi.org/10.1002/ana.410190309>

Sorrento, G. U., Archambault, P. S., & Fung, J. (2018). Adaptation and post-adaptation effects of haptic forces on locomotion in healthy young adults. *Journal of NeuroEngineering and Rehabilitation*, *15*(1), 20. <https://doi.org/10.1186/s12984-018-0364-0>

Spencer, K. A., Friedlander, C., & Brown, K. A. (2020). Predictors of health-related quality of life and communicative participation in individuals with dysarthria from parkinson’s disease. *International Journal of Neurodegenerative Disorders*, *3*(1), 1–7. <https://doi.org/10.23937/2643-4539/1710014>

Teplansky, K. J., Wisler, A., Goffman, L., & Wang, J. (2023). The impact of stimulus length in tongue and lip movement pattern stability in amyotrophic lateral sclerosis. *Journal of Speech, Language, and Hearing Research*. <https://doi.org/10.1044/2023_JSLHR-23-00079>

Thompson, A., & Kim, Y. (2024). Acoustic and kinematic predictors of intelligibility and articulatory precision in parkinson’s disease. *Journal of Speech, Language, and Hearing Research*, *67*(10), 3595–3611. <https://doi.org/10.1044/2024_JSLHR-24-00153>

Tienkamp, T. B., Rebernik, T., Jacobi, J., Wieling, M., & Abur, D. (2024). *ISSP 2024 - 13th International Seminar on Speech Production*. 91–94. <https://doi.org/10.21437/issp.2024-24>

Tjaden, K., Kain, A., & Lam, J. (2014). Hybridizing conversational and clear speech to investigate the source of increased intelligibility in speakers with parkinson’s disease. *Journal of Speech, Language, and Hearing Research*, *57*(4), 1191–1205. <https://doi.org/10.1044/2014_JSLHR-S-13-0086>

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., & Wilding, G. E. (2004). Rate and loudness manipulations in dysarthria: Acoustic and perceptual findings. *Journal of Speech, Language, and Hearing Research*, *47*(4), 766–783. <https://doi.org/10.1044/1092-4388(2004/058)>

Tremblay, S., Shiller, D. M., & Ostry, D. J. (2003). Somatosensory basis of speech production. *Nature*, *423*(6942), 866–869. <https://doi.org/10.1038/nature01710>

Turner, G. S., Tjaden, K., & Weismer, G. (1995). The influence of speaking rate on vowel space and speech intelligibility for individuals with amyotrophic lateral sclerosis. *Journal of Speech, Language, and Hearing Research*, *38*(5), 1001–1013. <https://doi.org/10.1044/jshr.3805.1001>

Villacorta, V. M., Perkell, J. S., & Guenther, F. H. (2007). Sensorimotor adaptation to feedback perturbations of vowel acoustics and its relation to perception. *The Journal of the Acoustical Society of America*, *122*(4), 2306–2319. <https://doi.org/10.1121/1.2773966>

Weismer, G., & Bunton, K. (1999). Influences of pellet markers on speech production behavior: Acoustical and perceptual measures. *The Journal of the Acoustical Society of America*, *105*(5), 2882–2894. <https://doi.org/10.1121/1.426902>

Weismer, G., Jeng, J.-Y., Laures, J. S., Kent, R. D., & Kent, J. F. (2001). Acoustic and intelligibility characteristics of sentence production in neurogenic speech disorders. *Folia Phoniatrica Et Logopaedica*, *53*(1), 1–18. <https://doi.org/10.1159/000052649>

Weismer, G., Laures, J. S., Jeng, J. Y., Kent, R. D., & Kent, J. F. (2000). Effect of speaking rate manipulations on acoustic and perceptual aspects of the dysarthria in amyotrophic lateral sclerosis. *Folia Phoniatr Logop*, *52*(5), 201–219. <https://doi.org/10.1159/000021536>

Whitfield, J. A., & Goberman, A. M. (2014). Articulatoryacoustic vowel space: Application to clear speech in individuals with Parkinson’s disease. *Journal of Communication Disorders*, *51*, 19–28. <https://doi.org/10.1016/j.jcomdis.2014.06.005>

Whitfield, J. A., & Mehta, D. D. (2019). Examination of clear speech in parkinson disease using measures of working vowel space. *Journal of Speech, Language, and Hearing Research*, *62*(7), 2082–2098. <https://doi.org/10.1044/2019_JSLHR-S-MSC18-18-0189>

# Supplemental Information

## Group, Sex, Age, and Order Effects

Although group, sex, age, and trial order effects within the Before Sensors time point were not the primary focus of this study, their inclusion in our models as interaction terms (Group × Time Point) or covariates (sex, age, trial order) allowed us to examine these variables.

**Group Effects.** PwPD demonstrated robustly smaller AAVSs (β = -.29, pd = 100%, CI = [-.48, -.08]) compared to control speakers at the baseline Before Sensors condition, after adjusting for age, sex, and articulation rate. Conversely, PwPD and control speakers showed comparable articulation rates (β = .04, pd = 57%, CI = [-.39, .46]), M1 (/s/: β = -.30, pd = 83%, CI = [-.93, .31]; /ʃ/: β = -.05, pd = 57%, CI = [-.56, .48]) and M2 values (/s/: β = .06, pd = 70%, CI = [-.18, .31]; /ʃ/: β = -.05, pd = 59%, CI = [-.55, .46]), intelligibility ratings (β = -.27, pd = 81%, CI = [-.90, .35]) and naturalness ratings (β = -.61, pd = 96%, CI = [-1.31, .10]) at the baseline Before Sensors condition.

**Sex Effects.** Female speakers demonstrated robustly larger AAVSs (β = .31, pd = 100%, CI = [.14, .49]) and higher M1 and M2 values for /s/ (M1: β = .80, pd = 100%, CI = [.27, 1.33]; M2: β = .24, pd = 99%, CI = [.03, .46]) compared to male speakers, after adjusting for age, group, and articulation rate. Male and female speakers did not robustly differ in their articulation rates (β = -.33, pd = 95%, CI = [-.71, .06]), M1 or M2 values for /ʃ/ (M1: β = .36, pd = 94%, CI = [-.11, .81]; M2: β = .35, pd = 94%, CI = [-.09, .80]), intelligibility ratings (β = .30, pd = 85%, CI = [-.27, .87]), or naturalness ratings (β = .32, pd = 84%, CI = [-.33, .95]).

**Age Effects.** Older speakers were perceived to be less intelligible (β = -.03, pd = 99%, CI = [-.06, -.004]) and natural (β = -.03, pd = 98%, CI = [-.06, -.003]), and demonstrated reduced M1 values for /s/ (β = -.03, pd = 99%, CI = [-.06, -.004]) in the Before Sensors time point, after adjusting for group, sex, and articulation rate. In contrast, age did not robustly impact articulation rate (β = -.02, pd = 97%, CI = [-.04, .001]), AAVS (β = -.003, pd = 73%, CI = [-.01, .01]), M1 for /ʃ/ (β = -.02, pd = 97%, CI = [-.04, .0003]), or M2 measures (/s/: β = -.0003, pd = 53%, CI = [-.01, .01]; /ʃ/: β = -.02, pd = 97%, CI = [-.04, .00003]).

**Trial Order Effects.** For the perceptual measures, trial order had no robust impact on listeners’ intelligibility ratings (β = -.01, pd = 97%, CI = [-.02, .0004]). However, listeners systematically rated samples as less natural the later they were presented in the perceptual experiment (β = -.01, pd = 99%, CI = [-.03, -.002]). This tendency was consistent across both groups, regardless of the speaker’s sex or age.

## Figures

### Figure S1

|  |
| --- |
| *Note*. The raw data for each participant across all three time points (left panel), between Before Sensors and With Sensors (middle panel), and between Before Sensors and After Sensors (right panel).  Figure S1: Individual raw data across the target measures. |

## Tables

### Table S1

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  | Group | Age | Time Since Diagnosis (years) | Severity1 | | --- | --- | --- | --- | --- | | Male Speakers | | | | | | PDM01 | PD | 60 |  | Profound | | PDM02 | PD | 80 |  | Profound | | PDM03 | PD | 50 |  | Mild | | PDM04 | PD | 79 |  | Severe | | PDM05 | PD | 65 |  | Moderate | | PDM06 | PD | 82 |  | Moderate | | PDM07 | PD | 78 |  | Severe | | PDM08 | PD | 82 |  | Moderate | | PDM09 | PD | 74 |  | Moderate | | n | 9 | — | — | — | | M | — | 72.22 | — | — | | SD | — | 11.34 | — | — | | Female Speakers | | | | | | PDF01 | PD | 80 |  | Mild | | PDF02 | PD | 61 |  | Severe | | PDF03 | PD | 63 |  | Moderate | | PDF04 | PD | 69 |  | Mild | | n | 4 | — | — | — | | M | — | 68.25 | — | — | | SD | — | 8.54 | — | — | | N | 13 | — | — | — | | Total M | — | 71.00 | — | — | | Total SD | — | 10.38 | — | — | | 1 Severity labels were previously determined by Thompson & Kim (2024) based on the intelligibility measures using the surrogate-severity measures outlined in Stipancic et al. (2022). Specifically, intelligibility values > 94% are `Normal`, 85% - 94% are `Mild`, 70% - 84% are `Moderate`, 45% - 69% are `Severe`, and < 45% are `Profound`. | | | | |   Table S1: Speaker demographics for the speakers with Parkinson’s disease. |

### Table S2

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  | Group | Age | | --- | --- | --- | | Male Speakers | | | | CM01 | Control | 49 | | CM02 | Control | 51 | | CM03 | Control | 85 | | CM04 | Control | 47 | | CM05 | Control | 68 | | CM06 | Control | 70 | | CM07 | Control | 53 | | CM08 | Control | 74 | | CM09 | Control | 81 | | n | 9 | — | | M | — | 64.22 | | SD | — | 14.52 | | Female Speakers | | | | CF01 | Control | 66 | | CF02 | Control | 67 | | CF03 | Control | 54 | | CF04 | Control | 68 | | CF05 | Control | 56 | | CF06 | Control | 68 | | CF07 | Control | 65 | | CF08 | Control | 49 | | CF09 | Control | 65 | | CF10 | Control | 72 | | CF11 | Control | 76 | | CF12 | Control | 67 | | n | 12 | — | | M | — | 64.42 | | SD | — | 7.69 | | N | 21 | — | | Total M | — | 64.33 | | Total SD | — | 10.81 |   Table S2: Speakers demographics for the control speakers. |

### Table S3

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| | **Characteristic** | **N = 79**1 | | --- | --- | | Age | 20.62 (0.96) | | Gender |  | | Woman | 79 (100%) | | Race |  | | Asian American | 3 (4%) | | Biracial or Multiracial | 3 (4%) | | Black/African American | 1 (1%) | | Prefer not to say | 1 (1%) | | White/Caucasian | 71 (90%) | | Ethnicity |  | | Hispanic/Latino | 13 (16%) | | Not Hispanic/Latino | 64 (81%) | | Prefer not to say | 2 (3%) | | 1 Mean (SD); n (%) | |   Table S3: Listener demographics. |

### Table S4

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| | Predictors | Articulation Rate (syl/s) | | | AAVS (log of mel2) | | | | --- | --- | --- | --- | --- | --- | --- | | Estimates | pd | CI (95%) | Estimates | pd | CI (95%) | | Fixed Effects | | | | | | | | Intercept (Control × Before Sensors) | 6.04 | 1.00 | 4.79 – 7.26 | 10.23 | 1.00 | 9.44 – 11.01 | | PwPD | .04 | .57 | -.39 – .46 | -.29 | 1.00 | -.48 – -.08 | | Female | -.33 | .95 | -.71 – .06 | .31 | 1.00 | .14 – .49 | | Age | -.02 | .97 | -.04 – .001 | -.003 | .73 | -.01 – .01 | | Articulation Rate |  |  |  | -.01 | .58 | -.10 – .08 | | Control × With Sensors | .14 | .98 | .01 – .26 | -.06 | .96 | -.12 – .01 | | PwPD × With Sensors | -.07 | .76 | -.27 – .13 | .03 | .74 | -.07 – .14 | | Control × After Sensors | .24 | 1.00 | .12 – .37 | .01 | .65 | -.05 – .08 | | PwPD × After Sensors | -.10 | .84 | -.31 – .10 | .09 | .96 | -.01 – .19 | | Random Effects | | | | | | | | σ2 | .126 |  |  | .103 |  |  | | NSpeakerID | 34 |  |  | 34 |  |  | | NSpeakerID/Phrase | 102 |  |  |  |  |  | | Summary | | | | | | | | Observations | 305 |  |  | 102 |  |  | | Bayesian R2 | .798 |  |  | .905 |  |  |   Table S4: *Note.* PwPD = People with Parkinson’s disease |

### Table S5

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| | Predictors | M1 for /s/ (kHz) | | | M1 for /ʃ/ (kHz) | | | | --- | --- | --- | --- | --- | --- | --- | | Estimates | pd | CI (95%) | Estimates | pd | CI (95%) | | Fixed Effects | | | | | | | | Intercept (Control × Before Sensors) | 7.07 | 1.00 | 4.47 – 9.73 | 6.59 | 1.00 | 4.32 – 8.92 | | PwPD | -.30 | .83 | -.93 – .31 | -.05 | .57 | -.56 – .48 | | Female | .80 | 1.00 | .27 – 1.33 | .36 | .94 | -.11 – .81 | | Age | -.03 | .99 | -.06 – -.004 | -.02 | .97 | -.04 – 3e-04 | | Articulation Rate | -.003 | .51 | -.34 – .34 | -.29 | .97 | -.59 – .01 | | Control × With Sensors | -.80 | 1.00 | -1.08 – -.52 | -.11 | .83 | -.33 – .12 | | PwPD × With Sensors | .43 | .97 | -.02 – .88 | .06 | .63 | -.29 – .42 | | Control × After Sensors | -.16 | .86 | -.44 – .14 | .21 | .96 | -.03 – .45 | | PwPD × After Sensors | .22 | .83 | -.24 – .67 | .05 | .60 | -.31 – .41 | | Random Effects | | | | | | | | σ2 | .206 |  |  | .144 |  |  | | NSpeakerID | 34 |  |  | 34 |  |  | | Summary | | | | | | | | Observations | 102 |  |  | 103 |  |  | | Bayesian R2 | .799 |  |  | .745 |  |  |   Table S5: *Note.* PwPD = People with Parkinson’s disease |

### Table S6

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| | Predictors | M2 for /s/ (kHz) | | | M2 for /ʃ/ (kHz) | | | | --- | --- | --- | --- | --- | --- | --- | | Estimates | pd | CI (95%) | Estimates | pd | CI (95%) | | Fixed Effects | | | | | | | | Intercept (Control × Before Sensors) | 2.22 | 1.00 | 1.10 – 3.31 | 6.55 | 1.00 | 4.34 – 8.78 | | PwPD | .06 | .70 | -.18 – .31 | -.05 | .59 | -.55 – .46 | | Female | .24 | .99 | .03 – .46 | .35 | .94 | -.09 – .80 | | Age | -3e-04 | .53 | -.01 – .01 | -.02 | .97 | -.04 – 3e-05 | | Articulation Rate | .08 | .86 | -.07 – .23 | -.28 | .97 | -.58 – .004 | | Control × With Sensors | .10 | .96 | -.02 – .23 | -.11 | .83 | -.34 – .11 | | PwPD × With Sensors | .09 | .83 | -.10 – .28 | .07 | .64 | -.29 – .43 | | Control × After Sensors | .05 | .79 | -.08 – .18 | .21 | .96 | -.02 – .44 | | PwPD × After Sensors | .01 | .52 | -.19 – .20 | .05 | .60 | -.30 – .41 | | Random Effects | | | | | | | | σ2 | .039 |  |  | .022 |  |  | | NSpeakerID | 34 |  |  | 34 |  |  | | Summary | | | | | | | | Observations | 102 |  |  | 103 |  |  | | Bayesian R2 | .675 |  |  | .744 |  |  |   Table S6: *Note.* PwPD = People with Parkinson’s disease |

### Table S7

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| | Predictors | Intelligibility (logit scale) | | | Naturalness (logit scale) | | | | --- | --- | --- | --- | --- | --- | --- | | Estimates | pd | CI (95%) | Estimates | pd | CI (95%) | | Fixed Effects | | | | | | | | Intercept (Control × Before Sensors) | 5.52 | 1.00 | 3.23 – 7.80 | 3.21 | .99 | .62 – 5.87 | | PwPD | -.27 | .81 | -.90 – .35 | -.61 | .96 | -1.31 – .10 | | Female | .30 | .85 | -.27 – .87 | .32 | .84 | -.33 – .95 | | Age | -.03 | .99 | -.06 – -.004 | -.03 | .98 | -.06 – -.003 | | Articulation Rate | -.14 | .86 | -.38 – .11 | -.01 | .53 | -.29 – .27 | | Control × With Sensors | -.20 | .99 | -.38 – -.03 | -.44 | 1.00 | -.63 – -.25 | | PwPD × With Sensors | -.45 | 1.00 | -.72 – -.17 | .01 | .53 | -.29 – .32 | | Control × After Sensors | .16 | .96 | -.02 – .34 | .08 | .80 | -.11 – .28 | | PwPD × After Sensors | -.25 | .96 | -.53 – .02 | .21 | .90 | -.10 – .51 | | Trial Number | -.01 | .97 | -.02 – 4e-04 | -.01 | .99 | -.03 – -.002 | | Random Effects | | | | | | | | σ2 | -.371 |  |  | .221 |  |  | | NSpeakerID | 34 |  |  | 34 |  |  | | NSpeakerID/Passage | 102 |  |  | 102 |  |  | | NListenerID | 78 |  |  | 78 |  |  | | Summary | | | | | | | | Observations | 1326 |  |  | 1326 |  |  | | Bayesian R2 | .607 |  |  | .544 |  |  |   Table S7: *Note.* PwPD = People with Parkinson’s disease |

1. While the University of Wisconsin (UW) X-ray microbeam (XRMB) system is not an EMA system, this type of X-ray methodology was used prior to the widespread use of EMA (Barlow et al., 1983; Fujimura et al., 1973; Westbury, 1991). However, both methods allow for the investigation of fleshpoint data during speech movement. The XRMB used eight gold pellets attached to various articulators, including four lingual pellets. The pellets were 2-3 mm in diameter and did not require sensor wires like EMA devices. [↑](#footnote-ref-21)
2. Please note that the current study measured intelligibility using VAS, while Hustad et al. (2015) used orthographic transcriptions. However, the approach to describing reliability among perceptual ratings is consistent. [↑](#footnote-ref-38)