

This is the accepted version of an article published by the American Speech Language Hearing Association (ASHA) in the Journal of Speech, Language, and Hearing Research, © 2023, Published online February 24, 2023. It is available online at https://doi.org/10.1044/2022_JSLHR-22-00291 and is made available here with permission of ASHA Journals.

Impact of face masks on speech in Parkinson's disease: Effect of clear and loud speech styles

Thea Knowles^{1,2}, Gursharan Badh²

¹ Department of Communicative Sciences and Disorders, Michigan State University, East Lansing, MI

² Department of Communicative Disorders and Sciences, University at Buffalo, Buffalo, NY

Correspondence to:

Thea Knowles
205 Oyer Speech and Hearing Building,
Michigan State University, East Lansing MI 48823
thea@msu.edu

Conflict of Interest Statement: The authors have no conflicts of interest to disclose.

Funding: No funding was received for the current study.

Abstract

Purpose: Quantify the combined effects of face masks and effortful speech styles on speech intensity, spectral moments, and measures of spectral balance in talkers with Parkinson's disease.

Method: Fifteen people with Parkinson's disease and 15 healthy, older controls read aloud sentences in three face mask and three speech style conditions. Mask conditions included no mask, surgical masks, and KN95 masks. Speech styles included habitual, clear, and loud. Acoustic measures of intensity, spectral moments, and spectral balance were modelled as a function of speaker group, mask, and speech style.

Results: Overall, talkers with PD demonstrated lower concentrations of high frequency spectral energy in their speech. Face masks attenuated high frequency energy, while clear followed by loud speaking styles amplified high frequencies. Overall, the attenuation observed by face masks was preserved across speech styles, and both mask and speech patterns were observed to be similar across groups.

Discussion: Clear and loud speech styles were effective in compensating for the damping effects of masks in talkers with and without PD. However, given that people with PD demonstrated poorer overall spectral balance compared to controls, the gains afforded by speaking clearly or loudly may be limited when wearing a face mask.

Introduction

The use of face masks in the general population has increased dramatically since the onset of the COVID-19 pandemic. At the time of writing, face masks continue to be required in many health care settings. While an important means of personal protective equipment to prevent the spread of viral particles, face masks can make spoken communication difficult. This is likely the case because masks can affect audibility of the speech signal and reduce the availability of visual cues (see Oosthuizen et al. (2022) and Badh & Knowles (2022) for reviews). This difficulty may be enhanced in adverse listening conditions, such as in the presence of background noise or a distorted speech signal (Toscano & Toscano, 2021), and may be worse for certain masks over others (Corey et al., 2020).

People with communication disorders, such as those with dysarthria secondary to Parkinson's disease (PD), may be at a particular disadvantage when communicating while wearing a mask. Even without a mask, speech of many people with PD is often less intelligible compared to that of controls (Miller et al., 2007; Tjaden et al., 2014; Weismer et al., 2001). The speech of people with PD may be especially challenging to understand through a mask because some of the acoustic characteristics of parkinsonian speech are similar to the acoustic damping effects of face masks, meaning that the face masks may disproportionately affect people with PD compared to typical speakers or even other forms of communication disorders. Namely, relatively weak concentrations of higher frequency acoustic energy is characteristic both of parkinsonian speech (Cushnie-Sparrow, 2021; Dromey, 2003; Smith & Goberman, 2014) and face coverings (Atcherson et al., 2021; Corey et al., 2020; Goldin et al., 2020; Maryn et al., 2021). A common form of behavioral speech treatment in PD to improve speech intelligibility is to help individuals increase their vocal effort, i.e. speak more loudly or more clearly (Tjaden et al., 2014; Yorkston et al., 2007). This strategy of using effortful speech may be especially useful when wearing a mask in order to compensate for the damping effects of wearing a face covering.

This study examines the spectral acoustic consequences of wearing two types of face masks (surgical and KN95) worn by people with and without PD in three speaking conditions: habitual, clear, and loud. As will be described below, acoustic measures that characterize the shape and tendency of the spectral distribution of speech were chosen to identify the objective impact of face masks and speech styles on the acoustic profiles of speech in PD. Acoustic measures that are likely to be sensitive to speech intelligibility in PD were also measured to examine how masks impact speech transmission across the three speech styles. To the authors knowledge, this is the first study that examines the effects of face masks on communication of individuals with speech disorders. Furthermore, this study extends acoustic descriptions of clear and loud speech styles in order to contribute to a broader understanding of how spectral properties of speech may be modifiable. Findings will be discussed in the context of speech intelligibility and other perceptual outcomes.

Filtering effects of face masks on speech

Previous literature has consistently shown that most face coverings of the sort recommended by public health officials to protect against viral transmission of COVID-19 are associated with a low-pass filtering effect of spectral speech information (Badh & Knowles, 2022; Oosthuizen et al., 2022). Face masks have been shown to attenuate higher frequency information above 1 - 2 kHz across the long-term average spectrum of the speech signal (Atcherson et al., 2021; Corey et al., 2020; Goldin et al., 2020; Maryn et al., 2021). Attenuation from masks is also associated with lower, or more negative, spectral tilt (Magee

et al., 2020; Nguyen et al., 2021), which reflects a reduced ratio of the concentration of high frequency energy to low frequency energy in the speech signal. This spectral attenuation has been demonstrated to be less impacted for certain types of masks, such as surgical masks, compared to denser, more tightly woven masks, such as KN95 and N95 masks (Atcherson et al., 2020; Atcherson et al., 2021; Corey et al., 2020; Nguyen et al., 2021; Pörschmann et al., 2020). More tightly fitting masks may also restrict articulatory movements, such as jaw lowering or lip protrusion, which may have additional effects on the speech signal by modifying vocal tract configurations and in turn vowel and possibly consonant acoustics (Gutz et al., 2021; McKenna et al., 2022). McKenna et al. (2022) found talkers produced more vowel centralization when wearing N95 masks, captured by a decrease in their acoustic vowel articulation index. This finding suggests that speakers may be producing smaller articulatory movements, which the authors postulated may have occurred due to restricted oral movements when speakers wore N95 masks. However, Gutz et al. (2021) found the opposite when wearing KN95 masks, namely that speakers tended to increase their vowel space area with the masks on. This suggests speakers may be making compensatory speech adjustments when wearing masks and these adjustments may depend on the features of the masks themselves. Further acoustic and kinematic studies are needed to identify the underlying causes of vocal tract adjustments speakers may make as a result of mask wearing.

Characterization of this acoustic filtering imposed by face masks can be accomplished in several ways. One way is to pinpoint specific frequency bands or cut-offs at which attenuation is notable compared to when a talker is not wearing a mask (as demonstrated by Atcherson et al., 2021; Corey et al., 2020; Goldin et al., 2020; Maryn et al., 2021). Another is to characterize spectral balance measures that take into account the overall relative concentrations of low to high energy across the speech spectrum. For example, speech produced with a mask that demonstrates spectral attenuation above 1 kHz will have a relatively higher concentration of low frequency energy compared to that same signal were it not filtered through a mask. As noted above, face masks likely have a multifaceted impact on speech signals by not only acting as a filter to the acoustic source signal directly, but also potentially by impacting articulatory positioning and talkers' compensatory behaviors. This study, however, will focus on the acoustic spectral consequences of wearing a mask, further discussed below.

Perceptual outcomes of speech produced with face masks

Perceptually, speech produced in a face mask has often been associated with reduced intelligibility compared to when talkers are not wearing a mask, especially in adverse listening environments (Cohn et al., 2021; Toscano & Toscano, 2021). These effects have been shown to be at least in part due to the auditory signal and not only the absence of visual cues when a talkers' mouth is covered by a mask (Corey et al., 2020; Palmiero et al., 2016; Pörschmann et al., 2020). Speech produced in masks may also sound "muffled" (Guskaroska, 2021), which may at least in part be resultant from acoustic attenuation of higher frequency energy in the speech signal (Ostwald, 1960) such as that imposed by a face covering. Not all masks demonstrate equal amounts of detriment for listeners, however. For example, cloth masks and N95 masks have been shown to have a worse impact on speech intelligibility compared to surgical masks (Bottalico et al., 2020; Brown et al., 2021; Corey et al., 2020; Mendel et al., 2008; Thomas et al., 2011; Toscano & Toscano, 2021).

While perceptual outcomes are critical in understanding the functional impact of face masks on spoken communication, it is not clear the extent to which the acoustic impacts of face

masks play a role in perception. Characterizing the acoustic impact of face masks and potential speech strategies that could be used to compensate for them is an important first step to understanding possible perceptual outcomes. The focus of the current manuscript is thus to provide an objective characterization of acoustic profiles of face masks and speech styles.

Spectral acoustics of speech in PD

People with PD may be at a particular disadvantage when speaking while wearing a face mask due to speech characteristics associated with the disease. The majority of people with PD will develop hypokinetic dysarthria over the course of their disease (Logemann et al., 1978; Müller et al., 2001; Mutch et al., 1986), which is typically characterized by imprecise consonants, speech rate abnormalities, reduced speech loudness, and monopitch/monoloudness, among other features (Adams & Dykstra, 2009; Darley et al., 1969a, 1969b). Phonatory abnormalities are among the most common symptoms in PD and often include a breathy-hoarse voice quality (Logemann et al., 1978) often accompanied by reduced speech loudness (Adams & Dykstra, 2009). This cluster of symptoms is referred to as hypophonia (Adams & Dykstra, 2009). Acoustically, hypophonia is often characterized by reduced speech intensity, reduced harmonics-to-noise ratio reflecting greater noise in the signal, and reduced spectral balance (Adams & Dykstra, 2009; Cushnie-Sparrow, 2021; Tjaden, 2008). Measures of spectral balance vary across the literature, but overall characterize the distribution of energy across the speech spectrum and include, among others, spectral tilt, ratios of low-to-high spectral energy, and spectral moments. Spectral moments characterize the central shape and tendency of energy across the long-term average speech spectrum (LTAS). In contrast, spectral tilt and low-high spectral ratio measures both reflect differences in dB between bands of low and high frequency energy, as described above.

Steeper spectral tilt has been correlated with increased perceptions of breathiness in people with and without voice disorders (Hillenbrand et al., 1994; Hillenbrand & Houde, 1996). Flatter (less negative) spectral tilt has also been associated with better intelligibility and greater perceived loudness and vocal effort in typical speakers (Lu & Cooke, 2009). In PD, intensive voice treatment has been shown to be associated with changes in ratios of low-to-high spectral energy (Alharbi et al., 2019). A small body of literature suggests that spectral moments are useful for indexing perceptual measures of speech severity in the speech of people with speech or voice disorders, including those with PD (Dromey, 2003; Smith & Goberman, 2014; Tanner et al., 2005; Tjaden et al., 2010). First and second spectral moments, which reflect spectral means and standard deviations, have been shown to decrease following voice treatment in people with dysphonia, reflecting greater concentrations of higher frequency energy post-treatment (Tanner et al., 2005). Tanner et al. (2005) also found that these acoustic changes were related to decreased perceptions of dysphonia severity. Speech in people with PD has been found to have lower first and second moments (spectral means and standard deviations) and higher third and fourth moments (skewness and kurtosis), all of which support a pattern of lower concentrations of high frequency spectral energy compared to controls (Dromey, 2003; Smith & Goberman, 2014). In contrast, Tjaden et al. (2010) did not find overall group differences in spectral moments between talkers with PD and typical, age-matched controls, but did find that lower first moments and higher third moments were both associated with increases in perceived speech severity. Furthermore, these relationships were strongest in the speech of talkers with PD compared to talkers with multiple sclerosis and controls (Tjaden et al., 2010).

Given that face masks attenuate higher frequency components of the speech signal, and that higher frequency components of speech are already weakened in the speech of people with PD, it stands to reason that this population may have greater speech-related challenges when wearing a face mask. The burden imposed by masks stands to pose an additional barrier to communication for individuals with PD above and beyond typical speech challenges.

Effortful speech and application to PD

A common behavioral approach for improving speech outcomes in people with PD is to help individuals learn to use more effortful speech, such as speaking more loudly or more clearly. Both clear and loud speech are considered global behavioral treatment approaches for dysarthria because the changes elicited affect multiple speech subsystems, including respiration, phonation, articulation, and resonance (Dromey & Ramig, 1998). Loud speech, achieved by increasing respiratory-phonatory effort, is a common goal of standardized, evidence-based speech intervention programs targeting hypophonia (i.e., Lee Silverman Voice Treatment program; Ramig et al. (2004)). Clear speech, characterized by hyperarticulation and exaggeration, is another speech style that shows promise for increasing speech intelligibility (Martel et al., 2021; Tjaden et al., 2014), as demonstrated by experimental studies in which talkers are instructed to speak more clearly. While standardized clear speech intervention programs do not currently exist for people with PD, a small number of treatment studies have demonstrated that this may be an effective form of rehabilitation for dysarthria (Park et al., 2016; Shin et al., 2022; Whelan et al., 2022). Speaking more clearly and/or more loudly may additionally be advantageous in compensating for the effects of face masks for all speaker groups, but particularly those with speech and/or voice disorders.

In neurotypical speakers, both clear and loud speech styles have been associated with increases in vocal intensity and spectral balance (Fant, 1960; Knowles & Badh, 2022; Krause & Braid, 2004; Smiljanic, 2021; Ternström et al., 2006). Clear and loud speech both demonstrate increased spectral energy in mid-to-high frequencies. For example, loud speech has been shown to increase energy in the range of the first five formants, characterized by an overall flatter spectral slope compared to habitual speech (Fant, 1960; Ternström et al., 2006). This appears to be directly related to sound pressure level (Ternström et al., 2006). Ternström et al. (2006) demonstrated that at very loud speech, energy around the first formant is even greater. Studies of clear speech have demonstrated that increased energy in the 1 - 3 kHz range, corresponding to the range of the first three formants, is responsible for much of the intelligibility benefit of speaking clearly for individuals without speech disorders (Gilbert et al., 2014; Hazan et al., 2018; Hazan & Baker, 2011; Krause & Braid, 2004, 2009; Smiljanic, 2021). These increases in spectral energy appear to be driven by greater jaw opening and lingual movements that occur during loud and clear speech, which has been suggested in both typical talkers (Mefferd, 2017; Schulman, 1989) and those with PD (Kearney et al., 2017; Mefferd & Dietrich, 2019; Wong et al., 2017). Consistent with the literature in neurotypical talkers, louder speech produced by people with PD has been associated with increased concentrations of mid-to-high frequency energy (Alharbi et al., 2019). To date, spectral changes in clear speech have not explicitly studied in PD.

One study to date has quantified the effect of loud speech styles on spectral moments in PD (Neel, 2009), which supports the pattern observed in typical speakers. That is, when talkers with PD speak more loudly, this results in increased spectral means and standard deviations

and decreased skewness and kurtosis (Neel, 2009). Neel (2009) further examined differences in perceptual outcomes when speakers were told to speak more loudly compared to when their habitual speech was presented at a higher volume (amplified) for listeners. They found that loudly produced speech was more intelligible than amplified speech. The increases in higher frequency spectral information, as captured in the spectral moments analysis, were suggested as a possible explanation for these differences.

Face masks and effortful speech in young, typical speakers

Our previous study investigated the spectral acoustics of speech of young, typical people wearing face masks and speaking in habitual and effortful styles (Knowles & Badh, 2022). Specifically, speakers read aloud sentences in three face mask conditions: no mask, with a surgical mask, and with a KN95 mask. Each face mask condition was evaluated in three speech styles: habitual speech, clear speech, and loud speech. The present study extends these findings to older talkers with and without PD.

Outcomes from younger participants using habitual speaking styles supported previous findings that face masks are associated with poorer measures of spectral balance, characterized by lower first and second moments, higher third and fourth moments, as well as lower spectral tilt and energy in the 1 - 3 kHz range. Surgical masks were associated with less attenuation compared to KN95 masks, though smaller effect sizes were found between the two masks than when comparing them to no mask. Clear and loud speaking styles, as expected, were found to be associated with increases in higher frequency spectral energy, with loud speech exhibiting greater changes compared to clear speech. The overall pattern of the face mask effects was maintained in clear and loud speech styles, though effect sizes for the mask conditions were smaller compared to for the speech style conditions when all speech styles were pooled. While face masks continued to attenuate higher frequency information across clear and loud speech styles, younger people were successful in increasing higher frequency components of the speech signal when using these styles in such a way that was equal to or surpassed their baseline spectral measures in habitual, unmasked speech. In other words, speaking more effortfully was associated with a greater magnitude of change in measures of spectral balance compared to wearing face masks, effectively compensating for the attenuation of the masks.

As we age, physiological changes, such as increased laryngeal rigidity, can result in age-related voice changes (presbyphonia; see reviews in Martins et al., 2014; Kendall, 2007). Voice related changes can include increased vocal fatigue, hoarse voice quality, and difficulties with vocal projection (Gregory et al., 2012). Older adults with PD may experience these changes to their voice as a consequence of typical aging in addition to voice changes resulting from the disorder, as described above. Understanding the impact of wearing face masks on the speech of those with voice-related challenges, be they due to aging or disease, is an important step in being able to provide appropriate public health measures that can maximize public safety while minimizing communicative challenges in vulnerable populations.

Summary & Purpose

The present study extends the paradigm used in Knowles & Badh (2022) to explore the A) isolated effects of face masks and effortful speech styles on the spectral profiles of speech in PD and B) combined effects of face masks and effortful speaking styles on speech in PD. To summarize, face masks attenuate higher frequency spectral speech information, which can

in turn lead to challenging perceptual conditions. More effortful speech styles, such as clear and loud speech, have the opposite effect, and may be used to compensate for the effects of masks in typical speakers. People with PD, whose speech is characterized by a damping of energy in higher frequency spectral components, may be particularly disadvantaged by the additional attenuation for frequencies above 1 - 2 kHz imposed by face masks (Atcherson et al., 2021; Corey et al., 2020; Goldin et al., 2020; Maryn et al., 2021), though to date no literature has investigated the effects of face masks on people with PD.

Speaking clearly and/or loudly are behavioral strategies that are often implemented as a means of improving spoken communication in PD, and may be particularly beneficial to use when face masks must be worn. The purpose of this study was to identify the impact of face masks on the speech of people with PD and to determine whether clear and/or loud speech styles are successful at overcoming acoustic attenuation imposed by masks in older adults with and without PD. Three research questions were of interest.

The first two questions were designed to characterize the isolated effects of 1) face masks and 2) speech styles on speech in PD in order to describe the spectral profiles of these factors:

1. **Effects of face masks on spectral moments in habitual speech in people with and without PD:** What is the impact of face masks on spectral moments in speech produced by people with and without Parkinson's disease? We predict that people with PD will have reduced levels of high frequency spectral speech energy compared to controls (Cushnie-Sparrow, 2021; Dromey, 2003; Smith & Goberman, 2014) and that wearing a face mask while speaking will attenuate high frequency spectral information compared to not wearing a mask (e.g., Atcherson et al., 2021; Corey et al., 2020; Goldin et al., 2020; Knowles et al., 2021; Maryn et al., 2021). We predict that attenuation imposed by face masks will have a similar *magnitude* of effect for both speaker groups, as there is no apriori reason to assume the filtering effect of masks would differ across speakers.
2. **Effects of effortful speech without a face mask on spectral moments in people with and without PD:** What is the impact of clear and loud speaking styles on spectral moments of speech produced by people with and without PD, without a face mask? We predict that clear, followed by loud speech will lead to an increase in higher frequency spectral information compared to habitual speech. This predicted order of effects is exploratory given the lack of research on spectral density characterizations of clear and loud speech. It is motivated, however, by the known contribution of increased energy in the range of the first formants (Fant, 1960; Ternström et al., 2006), which is expected to be highest in loud speech given the overall increased intensity (e.g., Tjaden & Wilding, 2004).

The third question was designed to capture the *combined* effect of face masks and effortful speech styles on speech in PD in a way that characterized these effects on more targeted acoustic speech outcomes:

3. **Effects of PD, face masks, and effortful speech on acoustic measures of spectral balance:** What is the relationship between face masks and effortful speech styles (clear and loud) on acoustic measures of speech intensity and spectral balance (spectral tilt, mid-range frequency energy) of speech produced by people with and without PD? We predict that, when all mask and speech style data are pooled, higher frequency spectral information will be a) attenuated by face masks, as described in

Research Question 1 and b) increased by effortful speech styles, as described in Research Question 2. Based on our previous findings in younger talkers (Knowles & Badh, 2022), we anticipate that speaking effortfully will result in larger effect sizes than wearing a mask. Given the established effects of clear and loud speech styles on speech production in PD (Kearney et al., 2017; Mefferd & Dietrich, 2019; Tjaden et al., 2013; e.g., Tjaden & Wilding, 2004), we expect this our finding in younger talkers to replicate in older talkers with and without PD.

Methods

This study was approved by the Institutional Review Board at the University at Buffalo.

Participants

Participant demographics for the PD and older control (OC) participants are listed in Table 1. Sixteen participants with PD (11 men, 5 women; age range 62 - 79) and 18 controls (OC; 10 men, 8 women; age range 57 - 83) were recruited for this study. All participants were native speakers of North American English or Canadian English. Other than PD, participants reported no other neurological diagnoses. Participants completed the Montreal Cognitive Assessment (MoCA; Nasreddine et al., 2005). One PD and one OC participant (PD10, OC03) scored below 21/30 (the suggested cut-off for PD-dementia; Dalrymple-Alford et al., 2010) and were excluded as this could have impacted task performance. One control participant (OC01) had to leave the session early and did not complete the paradigm, and one control participant (OC10) had difficulty completing the modified speech conditions and was excluded.

In total, 15 PD (10 men, 5 women, age range 62 - 79) and 15 OC participants (10 men, 5 women, age range 57 - 80) were included in the final analysis. Of these, one OC participant (OC14) reported having a lisp as a child which they received treatment for until age 7. Two PD participants (PD04, PD05) reported having seen a speech-language pathologist for treatment of speech symptoms related to PD more than two years ago. These participants were not excluded on these bases. Hypophonia severity and speech characteristics were identified by two speech-language pathologists. Speech characteristics correspond to perceptual features recommended by Duffy (2019) when rating motor speech disorders. PD participants also completed the Communicative Participation Item Bank (CPIB; Baylor et al., 2013) as an index of the degree to which they felt their condition interfered with their participation in various communicative contexts. Lower scores (out of a possible 30) indicate more perceived interference.

Table 1: Participant demographics.

ID	Age	Gender	MoCA (/30)	Years since PD diagnosis	CPIB (/30)	Hypophonia severity	Deviant perceptual characteristics
PD01	79	m	23	3	24	mild-moderate	imprecise consonants, breathy/hoarse, short phrases
PD02	70	m	26	8	26	mild	hoarse, short rushes of speech, increased rate
PD03	72	m	26	<1	20	mild	hoarse/wet, monopitch, imprecise articulation
PD04	63	m	28	3	21	mild	breathy/hoarse, increased pitch
PD05	67	m	29	12	11	mild	audible inspiration, short rushes of speech, increased rate overall, imprecise consonants
PD06	69	w	29	7	2	moderate	breathy/hoarse, strained
PD07	64	w	28	6	7	mild	short phrases, audible inspirations, monopitch, imprecise articulation
PD08	74	w	25	2	16	mild	imprecise articulation, monopitch
PD09	68	m	27	5	15	mild	breathy transient, reduced stress, prolonged phonemes, monopitch
PD11	65	m	24	8	15	mild	increased rate, breathy/hoarse, fast rushes, increased pitch
PD12	66	m	27	3	20	mild	breathy/hoarse, imprecise consonants, monopitch
PD13	62	m	24	1	5	mild	breathy, short phrases, loudness decay
PD14	73	w	21	1	24	very mild	breathy/hoarse
PD15	67	m	26	5	19	mild	breathy/hoarse, imprecise consonants, monopitch, decreased loudness
PD16	70	w	29	1	30	none	none noted
OC02	66	m	26		-	-	-
OC04	60	m	27		-	-	-
OC05	64	m	24		-	-	-
OC06	63	w	30		-	-	-
OC07	73	w	25		-	-	-
OC08	57	w	25		-	-	-
OC09	64	w	27		-	-	-
OC11	58	m	23		-	-	-
OC12	74	w	24		-	-	-

ID	Age	Gender	MoCA (/30)	Years since PD diagnosis	CPIB (/30)	Hypophonia severity	Deviant perceptual characteristics
OC13	60	m	28		-	-	-
OC14	58	m	27		-	-	-
OC15	80	m	27		-	-	-
OC16	67	m	24		-	-	-
OC17	74	m	26		-	-	-
OC18	77	m	27		-	-	-

Procedures

All participants read aloud phonetically balanced Harvard sentence lists (IEEE, 1969) in three face mask conditions and three speech style conditions (nine conditions in total). Face mask conditions included no mask, a standard three-ply surgical mask, and a KN95 mask. Face mask conditions were randomized across participants to avoid an order effect. Speech style conditions included habitual, clear, and loud. All participants began with their habitual speech style, completing all three mask conditions before moving on to the next speech style. The order of the clear and loud speech conditions were counterbalanced across all participants. For each speech style and mask combination, participants read two Harvard sentence lists, which were randomly selected from lists 1 - 18. For each altered speech style condition, participants were instructed on how to speak, and then given the opportunity to practice reading aloud sentences from a previous list until they felt they were ready to proceed with the task. Instructions for the clear speaking condition were to “speak clearly by over-enunciating your speech, similar to how you might speak to someone who is having difficulty hearing you, or someone who is learning English and is having difficulty understanding you.” Instructions for loud speech were to “speak at a volume that feels two times louder than your normal speaking voice.” Participants did not receive feedback on their performance.

Participants completed the experiment in a sound-treated room in approximately 30 minutes. Stimuli were presented on a computer monitor, and participants were seated at a 6 inch mouth-to-microphone distance from a table top microphone (Shure SM58). At the beginning of the experiment, calibration with a stable sound source was performed (Švec & Granqvist, 2018) in the following way. The researcher positioned a small loudspeaker under the chin of the participant and from it played a 1000 Hz tone of a fixed intensity. This was repeated three times and recorded. Real-time intensity of the tone was measured via a sound level meter (Galaxy Audio CM-170) positioned adjacent to the microphone. Following the recording sessions, the average intensity of the tone was measured off-line in Praat for each participant. The difference between the average real-time intensity and the off-line Praat intensity calculation was used as the “calibration factor.” The speech intensity of each utterance for each participant was adjusted by this calibration factor.

Acoustic measures

In preparation of acoustic measurement extraction, all utterance boundaries were demarcated and silences were removed using the Trim Silences function in Praat

(threshold: -35 dB; minimum silence duration of 100 ms; Boersma & Weenink 2021). Utterances in which participants produced major reading errors or disfluencies were discarded (~4%, nearly all for PD speakers and approximately evenly distributed across speaking conditions). A customized Praat script was used to automatically calculate acoustic measures of interest for each utterance following silence removal. Speech intensity was calculated from the utterance (after silences were removed), while all other measures were calculated from the spectrum or from the LTAS curves generated from each utterance.

Outcome measures were selected in order to capture measures of central tendency and spread of acoustic energy, as well as measures that may be sensitive to the filtering effects of the masks and speaking styles. Measures included spectral moments, spectral tilt, mid-range frequency energy, and speech intensity, described in more detail below.

Spectral moments characterize the central shape and tendency of energy across the speech spectrum and may be particularly useful for characterizing the effects of masks on the speech signal of persons with speech or voice disorders. The first spectral moment reflects spectral mean and is often referred to as the center of gravity (CoG), with higher values associated with a greater relative concentration of high frequency energy. The second spectral moment reflects the standard deviation of spectral mean. The third moment, skewness, is an indirect measure of tilt. Higher, more positive skewness levels reflect a lower overall concentration of high frequency energy. The fourth moment, kurtosis, reflects the “peakedness” of the spectrum, with higher values associated with more tightly concentrated values, and lower values associated with more even spread.

Spectral tilt, as described above in the introduction, captures the difference in energy between a low and high frequency band in the speech spectrum. In this study, spectral tilt reflects the difference in energy between 0 - 1 kHz and 1 - 10 kHz. Mid-range frequency energy reflects the mean energy in the 1 - 3 kHz range of the speech spectrum. Mid-range frequency energy between 1 - 3 kHz was chosen due to its importance in speech intelligibility, particularly as noted in clear speech literature (Hazan & Markham, 2004; Krause & Braida, 2004).

To summarize, outcome measures across the research questions were as follows:

- Research Questions 1 and 2: Spectral moments (M1 - M4; as described above)
 - M1 - Spectral means
 - M2 - Standard deviation of the spectral means
 - M3 - Spectral skewness
 - M4 - Spectral kurtosis
- Research Question 3: Spectral balance measures
 - Mean speech intensity (dB) across each utterance
 - Mid-range frequency energy: Mean energy in the 1 - 3 kHz range of the long-term average spectrum (LTAS)
 - Spectral tilt: Difference in energy between 0 - 1 kHz and 1 - 10 kHz of the LTAS

Statistical analyses

In brief, outcome measures were modelled as a function of speaker group, face mask, and speech style, as well as all possible interactions using linear mixed effects models in R (R Core Team, 2020). Gender was included as a co-variate. Models were constructed using the

lmer function from the *lmerTest* package and p-values were calculated using the Satterthwaite approximation (Kuznetsova et al., 2017). Three sets of models were built in order to answer the research questions, described below. All fixed effects and outcome measures for each set of models is reported in Table 2.

Coding of fixed effects

Speaker group and gender were sum-coded (OC = +1, PD = -1; Women = +1, Men = -1). A positive model estimate for group indicates a higher estimated value for the OC compared to the PD group. For gender a positive estimate indicates a higher value observed for women compared to men. Both face masks and speaking styles were contrast coded using Helmert contrasts, which permit the mean of the baseline level to be compared to the mean of the subsequent levels. A three-level contrast allows for two comparisons: Level 1 versus Levels 2 and 3, and Level 2 versus 3. The baseline for face masks was no mask. The first comparison (No Masks versus Masks) compares not wearing a mask with the average value from wearing both the surgical mask and KN95 mask (No mask: +2/3; Surgical mask: -1/3; KN95 mask: -1/3). The second comparison (Surgical versus KN95) compares the two masks to one another (Surgical mask: +1/2; KN95 mask: -1/2). Under this scheme, a positive model estimate for the No mask versus Masks comparison (for example) would indicate a higher value observed when no masks were worn. In other words, positive estimates indicate a higher value for the baseline condition in each comparison. Speaking style was coded in the same way as face masks, with Habitual set as the baseline. Comparisons included Habitual versus Clear/Loud (+2/3, -1/3, -1/3, respectively) and Clear versus Loud (+1/2, -1/2).

Model syntax

Research Question 1: Effects of PD and face masks on spectral moments in habitual speech

To address Research Question 1, we explored the effects of speaker group and masks in the habitual speech style condition. Four models were run, one for each spectral moment (M1 - M4). Initial model inspection revealed non-normal (skewed) residuals. Spectral moments were log-transformed and subsequent model inspection revealed that the data met assumptions of residual normality and heteroskedasticity. All models included fixed effects of group, mask condition (No mask, surgical, KN95), and their interaction, coded as described above. Participant gender was included as a covariate. Random effects included by-participant and by-item random intercepts, as well as by-participant random slopes for mask condition.

Research Question 2: Effects of PD and clear and loud speech styles without a mask on spectral moments

To address Research Question 2, we modelled the effects of speaker group and clear and loud versus habitual speech on the four spectral moments, resulting in four separate models. For these models, we included only speech produced in the no mask condition in order to characterize the unique effects of effortful speech on spectral moments. Fixed effects included group, speech condition (habitual, clear, loud), and their interaction, as well as gender. Random effects included by-participant and by-item random intercepts, as well as by-participant random slopes for speech condition. As with the models for Research

Question 1, spectral moments were log-transformed and model assumptions were confirmed using visual inspection of residuals.

Research Question 3: Effects of PD, face masks, and effortful speech styles on measures of spectral balance

Finally, to address Research Question 3, we pooled all the data and modelled the effects of group, masks, and speech styles combined. Outcome measures included speech intensity, spectral tilt, and mid-range frequency energy, resulting in three separate models. Fixed effects included group, mask condition, speech condition, and all possible interactions. Participant gender was again included as a covariate. Random effects included by-participant and by-item random intercepts. By-participant random slopes for mask and speech style were also included.

Effect size

In addition to reporting model estimates, we calculated effect sizes for each predictor using the method outlined by Westfall et al. (2014). Effect sizes were computed by dividing the estimate by the square root of the total variance of the random effects (i.e., the sum of the variance for each random effect term plus the total residual variance). We refer to general effect sizes as very small, small, medium, and large, which correspond roughly to traditional Cohen's d cutoffs of <0.2, 0.2 - 0.5, 0.5 - 0.8, and >0.8, respectively (Cohen, 1962). Effect sizes are not reported for non-statistically significant results ($p > 0.05$) and are instead reported in the results tables as "ns." This approach allows us to compare relative effect sizes of predictors within this study as an accompaniment to typical cut-offs of statistical significance. Extending use of this approach, which was designed for simple mixed effects model structures, may not be directly comparable to classic Cohen's d (Brysbaert & Stevens, 2018; Westfall et al., 2014). However, we use this approach with this caveat in mind as a benchmark for comparing relative effects in the present study.

Table 2: Summary of statistical models

Research Question	Data	Fixed effects	Outcome measures
1: Effects of PD and face masks on spectral moments in habitual speech	Habitual speech style condition; all face masks	Group * Face mask + Gender	Spectral moments (M1 - M4)
2: Effects of PD and clear and loud speech on spectral moments in speech without face masks	No Mask condition; all speech styles	Group* Speech style + Gender	Spectral moments (M1 - M4)
3: Effect of face masks and clear/loud speech on spectral acoustics	All face mask and speech style conditions	Group * Face mask * Speech style + Gender	Speech intensity and spectral balance measures (spectral tilt, mid-range frequency energy)

Results

An example of spectrograms and LTAS for one speaker are shown in Figure 1. Results are reported in Table 3 and Figures 2 - 5. A forest plot for the final model results for Research Question 3 appears in the Appendix, which was created with the *sjPlot* R package (Lüdtke, 2021). Results are organized in the following way. We report model results for each of the three research questions sequentially. Within each research question, we report main effects followed by interactions. Specifically, we first report the effects of speaker group and face masks in habitual speech on spectral moments (Research Question 1). We then report the effects of speaker group and speech style on spectral moments in speech produced without a mask (Research Question 2). Finally, we present pooled model results that explore the effects of face masks and speech styles on speech intensity and spectral balance measures (Research Question 3).

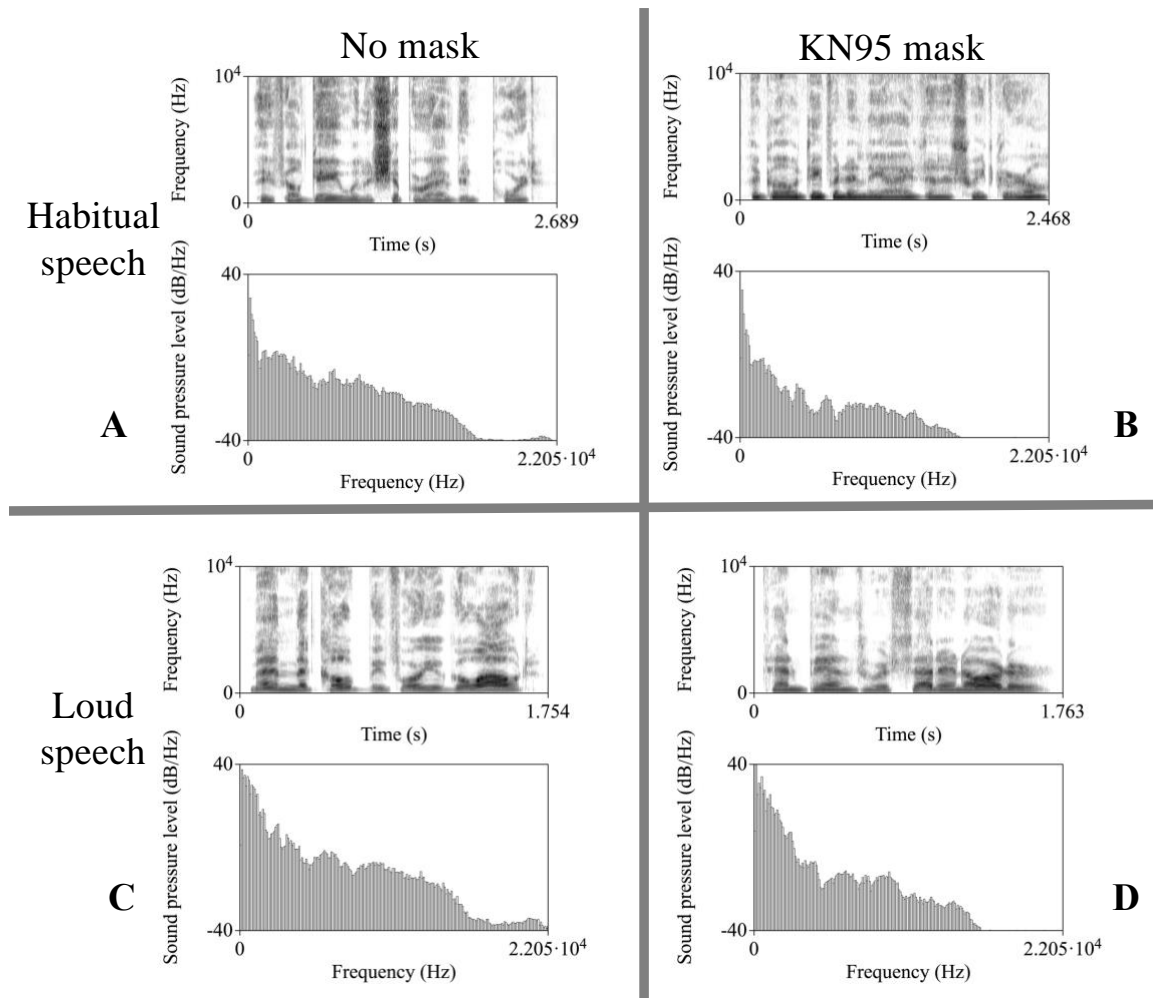


Figure 1 Examples of spectrograms and long-term average speech spectra for one talker (PD01) in (A) habitual speech, without a mask, (B) habitual speech, with the KN95 mask, (C) loud speech, without a mask, and (D) loud speech, with the KN95 mask.

Effects of group and masks on spectral acoustics in habitual speech.

Results for Research Question 1 addressing the impact of face masks on spectral acoustics of speech in habitual speech styles are reported in Table 3 and Figure 2.

Table 3: Model results for Research Question 1: effects of speaker group and face masks in habitual speech. Model estimates for each outcome measure are grouped by fixed effects terms. Abbreviations: PD = Parkinson's disease; OC = Older controls; NM = No masks; SM = Surgical mask; KN = KN95 mask; M1-Spectral mean = first spectral moment, spectral mean; M1-Spectral SD = second spectral moment, spectral standard deviation; M3-skew = third spectral moment, skewness; M4-kurt = fourth spectral moment, kurtosis; tilt = spectral tilt; mid = mid-range frequency energy 1-3 kHz; ns = Not statistically significant.

Measure	Contrast	Estimate	Std. Error	t	p	Effect size parameter	Effect size
M1-Spectral mean	(Intercept)	6.180	0.042	148.485	<0.001	19.957	large
	OC vs PD	0.128	0.038	3.382	0.002	0.413	small
	NM vs Masks	0.238	0.022	11.014	<0.001	0.768	medium
	SM vs KN	0.083	0.014	6.054	<0.001	0.269	small
	Women vs. Men	0.115	0.034	3.355	0.002	0.372	small
	OC vs PD:NM vs Masks	-0.020	0.022	-0.947	0.352	0.066	ns
	OC vs PD:SM vs KN	-0.018	0.014	-1.327	0.197	0.059	ns
M2-Spectral SD	(Intercept)	6.508	0.045	144.138	<0.001	17.529	large
	OC vs PD	0.057	0.042	1.365	0.184	0.153	ns
	NM vs Masks	0.460	0.023	20.268	<0.001	1.240	large
	SM vs KN	0.218	0.022	10.096	<0.001	0.588	medium
	Women vs. Men	0.147	0.038	3.875	0.001	0.397	small
	OC vs PD:NM vs Masks	-0.025	0.023	-1.096	0.282	0.067	ns
	OC vs PD:SM vs KN	-0.035	0.022	-1.609	0.12	0.094	ns
M3-skew	(Intercept)	1.872	0.053	35.461	<0.001	4.562	large
	OC vs PD	-0.108	0.047	-2.282	0.03	0.264	small
	NM vs Masks	-0.221	0.040	-5.599	<0.001	0.539	medium
	SM vs KN	-0.120	0.029	-4.114	<0.001	0.292	small
	Women vs. Men	-0.077	0.043	-1.781	0.086	0.189	ns

Measure	Contrast	Estimate	Std. Error	t	p	Effect size parameter	Effect size
M4-kurt	OC vs PD:NM vs Masks	0.031	0.040	0.784	0.439	0.076	ns
	OC vs PD:SM vs KN	0.035	0.029	1.204	0.239	0.085	ns
	(Intercept)	4.228	0.105	40.284	<0.001	5.047	large
	OC vs PD	-0.222	0.094	-2.347	0.027	0.265	small
	NM vs Masks	-0.685	0.077	-8.881	<0.001	0.818	large
	SM vs KN	-0.386	0.049	-7.832	<0.001	0.460	small
	Women vs. Men	-0.157	0.084	-1.857	0.075	0.187	ns
	OC vs PD:NM vs Masks	0.051	0.077	0.663	0.513	0.061	ns
	OC vs PD:SM vs KN	0.090	0.049	1.834	0.078	0.108	ns

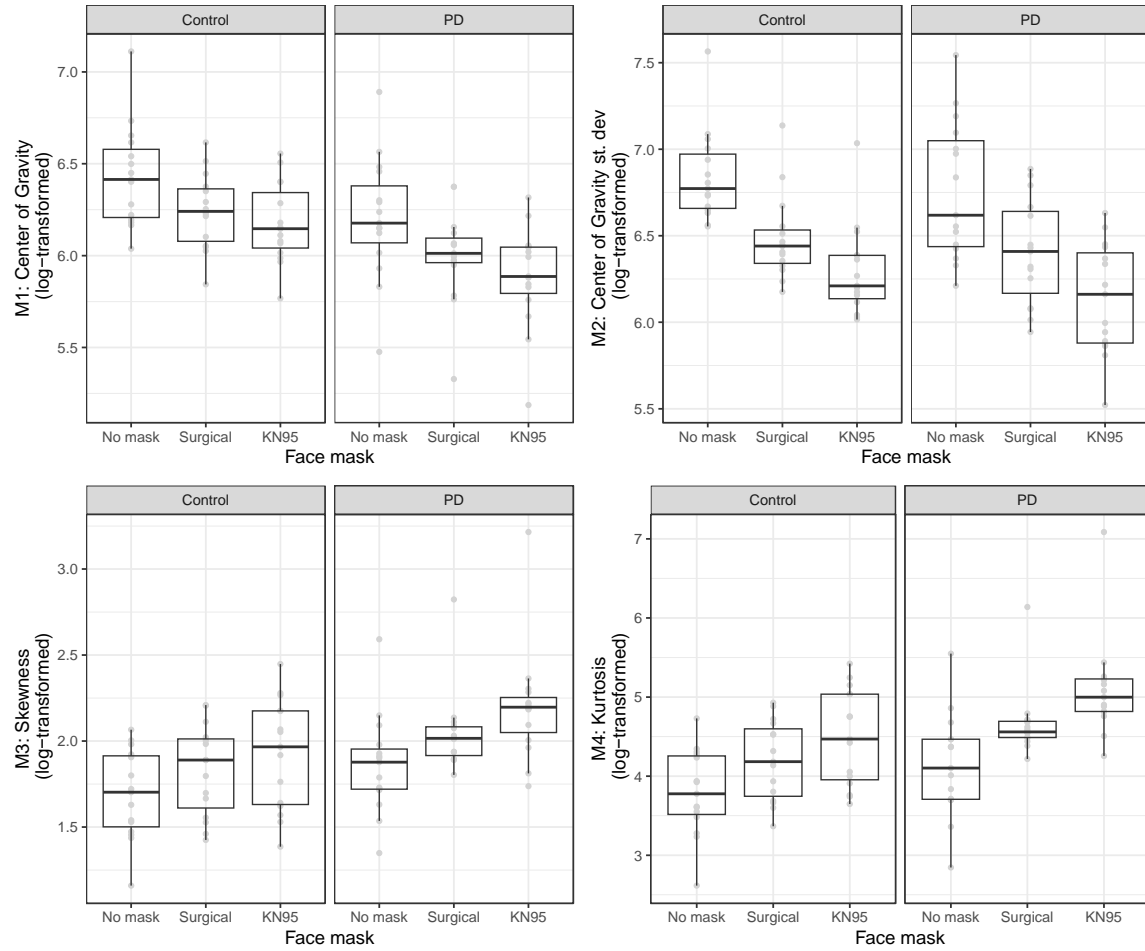


Figure 2 Effects of face mask and speaker group on spectral moments in habitual speaking styles. Outcome measures are reported on the log scale. PD = Parkinson's disease.

Main effect of group

Overall, the speech of the PD group was characterized by attenuated high frequency energy compared to controls. Specifically, in habitual speech, with all other effects held at their average values, a statistically significant main effect of group was found for M1, M3, and M4. No effect of group was found for M2. Small, positive model estimates indicated that controls had higher spectral means (M1: $\hat{\beta} = 0.128$, $p = 0.0021$) and lower skewness and kurtosis (M3 and M4, respectively; M3: $\hat{\beta} = -0.108$, $p = 0.031$; M4: $\hat{\beta} = -0.222$, $p = 0.0271$).

Main effect of mask

Wearing a mask attenuated higher frequency spectral energy, as evidenced by statistically significant main effects of face masks for all four spectral moments measures in habitual speech. The first comparison captured the difference between not wearing a mask compared to the average value of the surgical and KN95 masks (No Mask vs. Masks). Effects varied in magnitude. Negative estimates for this comparison indicate lower estimated values for speech produced with a mask compared to without, and reflect outcomes when group and gender are held at a constant, average value. Medium and large positive effects

were found for M1 and M2, respectively, reflecting higher spectral means and variability when participants were not wearing a mask (M1: $\hat{\beta} = 0.238, p < .001$; M2: $\hat{\beta} = 0.46, p < .001$). Medium and large negative effects were observed for M3 and M4, respectively, indicating lower skewness and kurtosis without a mask (M3: $\hat{\beta} = -0.221, p < .001$; M4: $\hat{\beta} = -0.685, p < .001$).

Statistically significant differences between the masks were also observed (SM vs. KN), though these effects were relatively smaller than the No Mask vs. Masks comparison. Surgical masks compared to KN95 masks were associated with higher M1 and M2 and lower M3 and M4, the same pattern observed for No Masks vs. Masks (M1: small effect, $\hat{\beta} = 0.083, p < .001$; M2: medium effect, $\hat{\beta} = 0.218, p < .001$; M3: small effect, $\hat{\beta} = -0.12, p < .001$; M4: small effect, $\hat{\beta} = -0.386, p < .001$). Findings support a greater damping effect of the KN95 mask compared to the surgical mask, indicating a hierarchy of effects which can be summarized by No mask > Surgical mask > KN95 mask.

Talker gender was also included as a covariate. Compared to men, women demonstrated significantly higher M1 and M2 (small effects, M1: $\hat{\beta} = 0.115, p = 0.0021$; M2: $\hat{\beta} = 0.147, p = 0.0011$). This is consistent with greater concentrations of low-frequency energy present in men's speech.

There were no group-by-mask interactions in habitual speech. Overall, this suggests that the pattern of effects of the masks was consistent across both speaker groups.

Effects of clear and loud speech on spectral moments

Results for Research Question 2 addressing the impact of speaker group and clear and loud speech on spectral moments in speech without a face mask are reported in Table 4 and Figure 3.

Table 4: Model results for Research Question 2: effects of speaker group and speech style without a mask. Model estimates for each outcome measure are grouped by fixed effects terms. Abbreviations: PD = Parkinson's disease; OC = Older controls; M1-Spectral mean = first spectral moment, spectral mean; M1-Spectral SD = second spectral moment, spectral standard deviation; M3-skew = third spectral moment, skewness; M4-kurt = fourth spectral moment, kurtosis; tilt = spectral tilt; mid = mid-range frequency energy 1-3 kHz; ns = Not statistically significant.

Measure	Contrast	Estimate	Std. Error	t	p	Effect size parameter	Effect size
M1-Spectral mean	(Intercept)	6.520	0.040	162.620	<0.001	17.618	large
	OC vs PD	0.110	0.037	3.009	0.006	0.298	small
	Habit vs Clear/Loud	-0.237	0.026	-9.198	<0.001	0.640	medium
	Clear vs Loud	-0.166	0.029	-5.615	<0.001	0.447	small
	Women vs. Men	0.189	0.037	5.080	<0.001	0.511	medium
	OC vs PD:Habit vs Clear/Loud	0.007	0.026	0.273	0.787	0.019	ns
	OC vs PD:Clear vs Loud	-0.017	0.029	-0.573	0.572	0.046	ns
M2-Spectral SD	(Intercept)	6.826	0.044	154.195	<0.001	17.517	large
	OC vs PD	0.036	0.040	0.906	0.373	0.094	ns
	Habit vs Clear/Loud	-0.012	0.026	-0.448	0.657	0.030	ns
	Clear vs Loud	0.010	0.021	0.491	0.627	0.026	ns
	Women vs. Men	0.162	0.038	4.226	<0.001	0.415	small
	OC vs PD:Habit vs Clear/Loud	0.008	0.026	0.292	0.772	0.020	ns
	OC vs PD:Clear vs Loud	-0.014	0.021	-0.697	0.492	0.037	ns
M3-skew	(Intercept)	1.615	0.048	33.993	<0.001	3.759	large
	OC vs PD	-0.082	0.042	-1.943	0.062	0.192	ns
	Habit vs Clear/Loud	0.164	0.035	4.737	<0.001	0.382	small
	Clear vs Loud	0.115	0.030	3.876	0.001	0.267	small

Measure	Contrast	Estimate	Std. Error	t	p	Effect size parameter	Effect size
	Women vs. Men	-0.075	0.044	-1.694	0.102	0.175	ns
	OC vs PD:Habit vs Clear/Loud	-0.009	0.035	-0.256	0.8	0.021	ns
	OC vs PD:Clear vs Loud	0.029	0.030	0.994	0.329	0.069	ns
M4-kurt	(Intercept)	3.609	0.102	35.277	<0.001	3.954	large
	OC vs PD	-0.176	0.092	-1.910	0.067	0.192	ns
	Habit vs Clear/Loud	0.241	0.065	3.700	0.001	0.264	small
	Clear vs Loud	0.145	0.055	2.615	0.014	0.158	very small
	Women vs. Men	-0.167	0.093	-1.788	0.085	0.183	ns
	OC vs PD:Habit vs Clear/Loud	-0.022	0.065	-0.331	0.743	0.024	ns
	OC vs PD:Clear vs Loud	0.056	0.055	1.021	0.316	0.062	ns

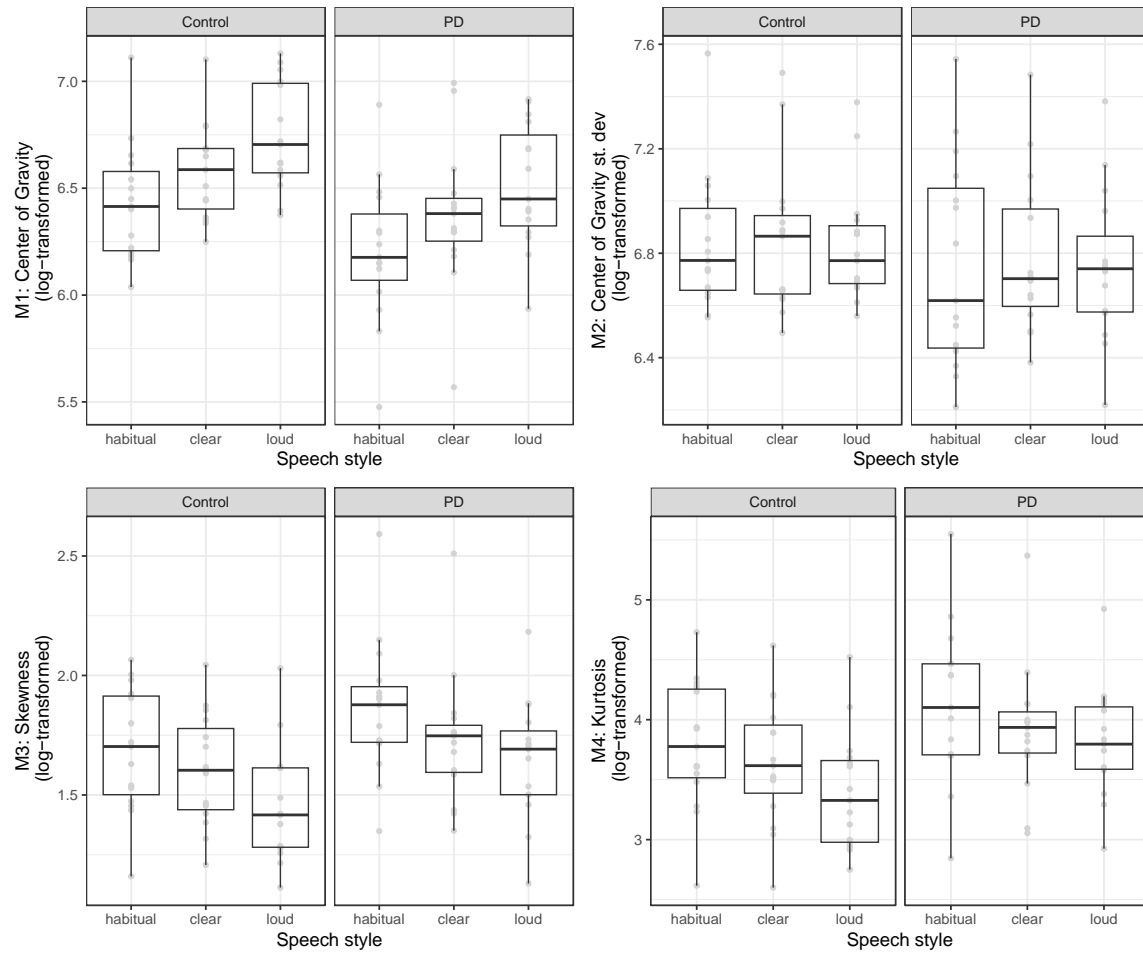


Figure 3 Effects of speech styles (habitual, clear, loud) and speaker group on spectral moments without a face mask. Outcome measures are reported on the log scale. PD = Parkinson's disease.

Main effect of group

Consistent with what was found in habitual speech, the PD group showed more attenuated higher frequency spectral energy compared to controls. This was captured by a small, statistically significant effect of speaker group for M1 ($\hat{\beta} = 0.11, p = 0.0061$), and reflects the No Mask data pooled across all three speech style conditions (habitual, clear, loud). No main effects of group were found for the other spectral moments.

Main effects of speech style

The first comparison captured the difference between habitual speech compared to clear and loud speech combined. A negative estimate indicates a lower value in habitual speech. Overall, clear and loud speech styles amplified higher frequency components of the spectrum, as evidenced by a medium effect on M1 ($\hat{\beta} = -0.237, p < .001$) and small effects on M3 and M4 (M3: $\hat{\beta} = 0.164, p < .001$; M4, kurtosis: $\hat{\beta} = 0.241, p = 0.0011$). No effect was observed for M2. This pattern was most evident for loud speech, which was captured by the Clear vs. Loud comparison. Here a negative estimate reflects a smaller value observed for clear compared to loud speech. Loud speech was associated with statistically significantly higher M1 and lower M3 and M4 values than clear speech (M1: small effect, $\hat{\beta} = -0.166, p < .001$; M3: small effect, $\hat{\beta} = 0.115, p = 0.0011$, M4: very small effect, $\hat{\beta} = 0.145, p = 0.0141$). Again, no effect was observed for M2.

Women had higher M1 and M2 than men (M1: medium effect, $\hat{\beta} = 0.189, p < .001$; M2: small effect, $\hat{\beta} = 0.162, p < .001$). There were no gender effects for M3 or M4. No group by speech style interactions were found.

Effects of masks and speaking style on speech intensity, spectral tilt, and mid-frequency energy

Results for Research Question 3 addressing the combined effects of masks and speech styles on measures of spectral balance including speech intensity, spectral tilt, and mid-range frequency energy are reported in Table 5 and Figures 4 and 5. A forest plot of model outcomes appears in the Appendix.

Table 5: Model results for Research Question 2 and 3: effects of speaker group, face mask, and speaking style. Model estimates for each outcome measure are grouped by fixed effects terms. Abbreviations: PD = Parkinson's disease; OC = Older controls; NM = No masks; SM = Surgical mask; KN = KN95 mask; tilt = spectral tilt; mid = mid-range frequency energy 1-3 kHz.; ns = Not statistically significant

Measure	Contrast	Estimate	Std. Error	t	p	Effect size parameter	Effect size
intensity	(Intercept)	81.443	0.753	108.139	<0.001	12.788	large
	OC vs PD	1.243	0.711	1.748	0.092	0.195	ns
	Habit vs Clear/Loud	-5.679	0.467	-12.166	<0.001	0.892	large
	Clear vs Loud	-3.761	0.656	-5.729	<0.001	0.591	medium
	NM vs Masks	0.392	0.105	3.742	0.001	0.061	very small
	SM vs KN	-0.302	0.100	-3.013	0.006	0.047	very small
	Women vs. Men	-1.590	0.713	-2.228	0.034	0.250	small
	OC vs PD:Habit vs Clear/Loud	-0.247	0.467	-0.529	0.601	0.039	ns
	OC vs PD:Clear vs Loud	-0.192	0.656	-0.293	0.772	0.030	ns
	OC vs PD:NM vs Masks	0.052	0.105	0.495	0.625	0.008	ns
	OC vs PD:SM vs KN	0.057	0.100	0.573	0.572	0.009	ns
	Habit vs Clear/Loud: NM vs Masks	0.096	0.086	1.121	0.262	0.015	ns
	Clear vs Loud:NM vs Masks	-0.296	0.099	-2.993	0.003	0.047	very small
	Habit vs Clear/Loud: SM vs KN	-0.481	0.099	-4.867	<0.001	0.076	very small
	Clear vs Loud:SM vs KN	0.324	0.114	2.841	0.005	0.051	very small

Measure	Contrast	Estimate	Std. Error	t	p	Effect size parameter	Effect size
	OC vs PD:Habit vs Clear/Loud: NM vs Masks	0.290	0.086	3.391	0.001	0.046	very small
	OC vs PD:Clear vs Loud:NM vs Masks	0.253	0.099	2.559	0.011	0.040	very small
	OC vs PD:Habit vs Clear/Loud: SM vs KN	-0.431	0.099	-4.357	<0.001	0.068	very small
	OC vs PD:Clear vs Loud:SM vs KN	0.206	0.114	1.808	0.071	0.032	ns
tilt	(Intercept)	-18.207	0.394	-46.168	<0.001	4.633	large
	OC vs PD	1.095	0.343	3.191	0.004	0.279	small
	Habit vs Clear/Loud	-2.681	0.293	-9.156	<0.001	0.682	medium
	Clear vs Loud	-1.981	0.345	-5.748	<0.001	0.504	medium
	NM vs Masks	1.753	0.149	11.748	<0.001	0.446	small
	SM vs KN	1.050	0.097	10.862	<0.001	0.267	small
	Women vs. Men	1.435	0.336	4.271	<0.001	0.365	small
	OC vs PD:Habit vs Clear/Loud	-0.044	0.293	-0.150	0.882	0.011	ns
	OC vs PD:Clear vs Loud	-0.259	0.345	-0.751	0.459	0.066	ns
	OC vs PD:NM vs Masks	-0.210	0.149	-1.408	0.17	0.053	ns
	OC vs PD:SM vs KN	-0.089	0.097	-0.919	0.366	0.023	ns
	Habit vs Clear/Loud: NM vs Masks	0.034	0.128	0.265	0.791	0.009	ns
	Clear vs Loud:NM vs Masks	0.052	0.148	0.352	0.725	0.013	ns

Measure	Contrast	Estimate	Std. Error	t	p	Effect size parameter	Effect size
mid 1-3khz	Habit vs Clear/Loud: SM vs KN	-0.175	0.147	-1.186	0.236	0.044	ns
	Clear vs Loud:SM vs KN	0.306	0.170	1.801	0.072	0.078	ns
	OC vs PD:Habit vs Clear/Loud: NM vs Masks	0.002	0.128	0.014	0.989	0.000	ns
	OC vs PD:Clear vs Loud:NM vs Masks	-0.282	0.148	-1.913	0.056	0.072	ns
	OC vs PD:Habit vs Clear/Loud: SM vs KN	-0.179	0.147	-1.215	0.225	0.046	ns
	OC vs PD:Clear vs Loud:SM vs KN	-0.230	0.170	-1.357	0.175	0.059	ns
	(Intercept)	13.867	0.924	15.005	<0.001	1.680	large
	OC vs PD	2.608	0.870	2.998	0.006	0.316	small
	Habit vs Clear/Loud	-8.334	0.631	-13.214	<0.001	1.010	large
	Clear vs Loud	-5.600	0.942	-5.943	<0.001	0.678	medium
	NM vs Masks	1.259	0.168	7.504	<0.001	0.153	very small
	SM vs KN	0.360	0.152	2.367	0.025	0.044	very small
mid 1-3khz	Women vs. Men	-0.181	0.855	-0.212	0.834	0.022	ns
	OC vs PD:Habit vs Clear/Loud	-0.195	0.631	-0.310	0.759	0.024	ns
	OC vs PD:Clear vs Loud	-0.338	0.942	-0.359	0.722	0.041	ns
	OC vs PD:NM vs Masks	-0.075	0.168	-0.444	0.66	0.009	ns
	OC vs PD:SM vs KN	0.001	0.152	0.005	0.996	0.000	ns

Measure	Contrast	Estimate	Std. Error	t	p	Effect size parameter	Effect size
	Habit vs Clear/Loud: NM vs Masks	-0.122	0.151	-0.807	0.42	0.015	ns
	Clear vs Loud:NM vs Masks	-0.364	0.174	-2.088	0.037	0.044	very small
	Habit vs Clear/Loud: SM vs KN	-0.714	0.174	-4.107	<0.001	0.087	very small
	Clear vs Loud:SM vs KN	0.623	0.201	3.108	0.002	0.075	very small
	OC vs PD:Habit vs Clear/Loud: NM vs Masks	0.299	0.151	1.983	0.047	0.036	very small
	OC vs PD:Clear vs Loud:NM vs Masks	0.060	0.174	0.345	0.73	0.007	ns
	OC vs PD:Habit vs Clear/Loud: SM vs KN	-0.545	0.174	-3.135	0.002	0.066	very small
	OC vs PD:Clear vs Loud:SM vs KN	0.041	0.201	0.205	0.837	0.005	ns

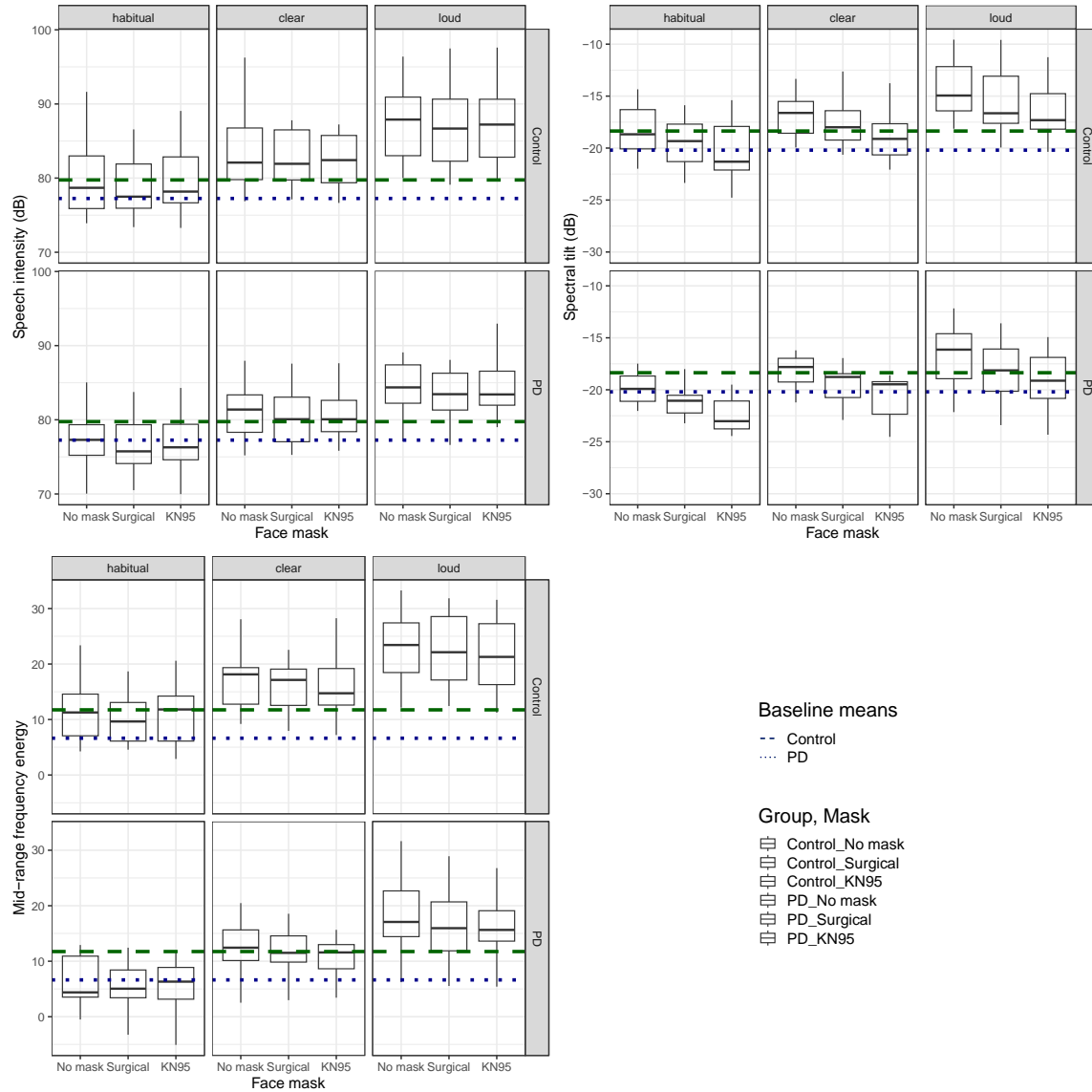


Figure 4 Effect of group, mask, and speech condition on all acoustic measures of interest. Horizontal lines reflect average speaker baselines in habitual speech without a mask: green dashed line reflects control speakers' baseline; blue dotted line reflects PD speakers' baseline. PD = Parkinson's disease.

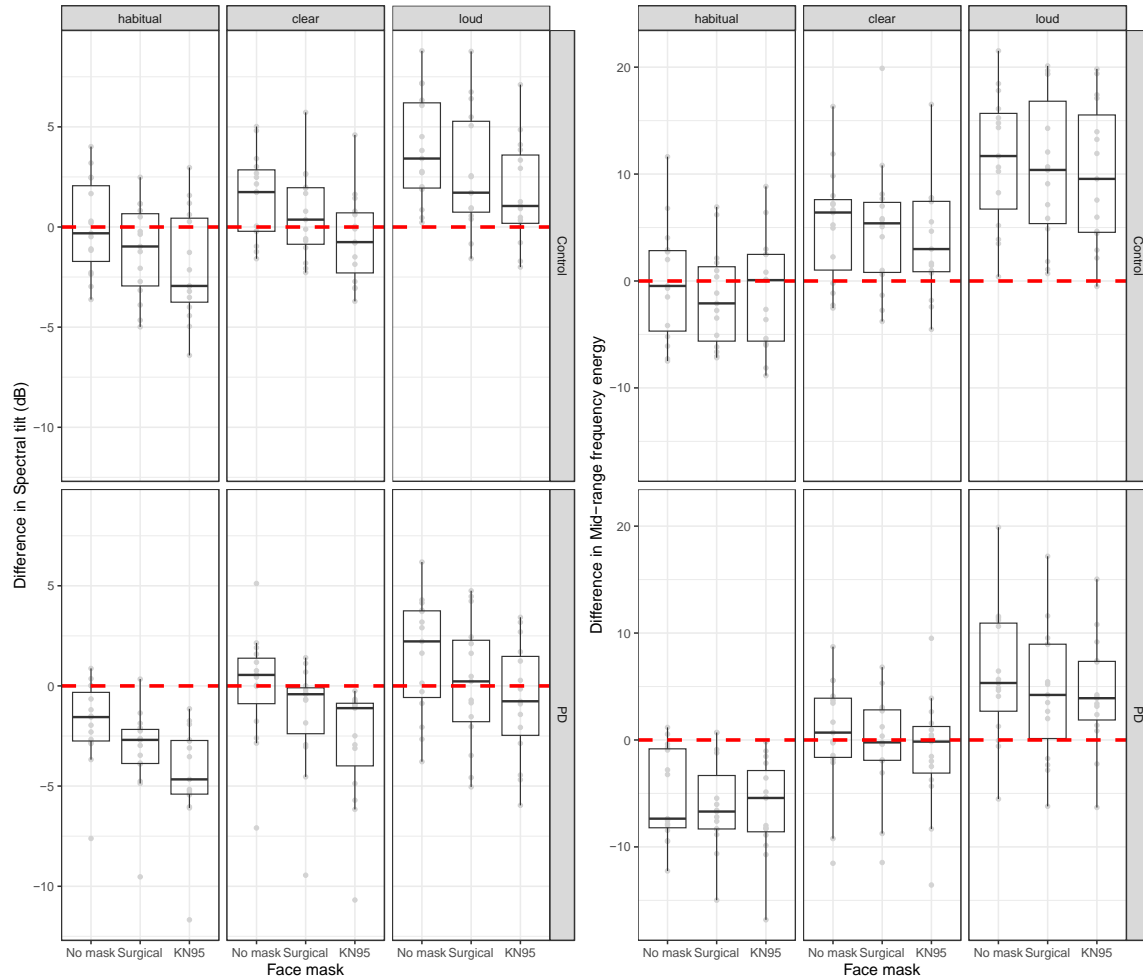


Figure 5 Effect of group, mask, and speech condition on the difference in spectral tilt and mid-frequency energy for each talker and group relative to the controls' baseline (habitual, unmasked speech). Horizontal dashed line reflects controls' baseline average ($y = 0$). Upright triangles reflect values that are greater than or equal to controls' baseline; downward triangles reflect values that fall below it. PD = Parkinson's disease.

Main effect of group

Controls produced greater concentrations of high frequency spectral energy, captured by small, statistically significant effects of group on spectral tilt ($\hat{\beta} = 1.095, p = 0.0041$) and mid-frequency energy ($\hat{\beta} = 2.608, p = 0.0061$). Significant main group differences were not observed for speech intensity.

Main effect of masks

When pooled across habitual, clear, and loud speech, the effects of wearing a mask patterned with those seen in habitual speech reported above when answering Research Question 1. That is, masks were associated with attenuation of higher frequency energy, and this effect was observed to be greater for the KN95 mask than for the surgical mask. Overall, effect sizes were small or very small. Positive estimates reflect higher outcomes without a mask compared to with the masks. Face masks were associated with overall lower speech intensity (very small effect, $\hat{\beta} = 0.392, p = 0.0011$), spectral tilt (small effect, $\hat{\beta} = 1.753, p < .001$), and mid-frequency energy (very small effect, $\hat{\beta} = 1.259, p < .001$). KN95 masks, compared to surgical masks, resulted in lower tilt (small effect, $\hat{\beta} = 1.05, p < .001$) and mid-frequency energy (very small effect; $\hat{\beta} = 0.36, p = 0.0251$), but *higher* speech intensity (very small effect, $\hat{\beta} = -0.302, p = 0.0061$).

Main effect of speech styles

When pooled across the three mask conditions and both speaker groups, the effects observed for loud and clear speech were consistent with the pattern observed on spectral moments in unmasked speech. Loud, followed by clear speech styles were associated with greater speech intensity and a greater concentration of higher frequency energy. When both clear and loud were pooled and compared to habitual speech (Habit vs Clear/Loud), large effects were observed for intensity ($\hat{\beta} = -5.679, p < .001$) and mid-frequency energy ($\hat{\beta} = -8.334, p < .001$), and a medium effect was observed for spectral tilt ($\hat{\beta} = -2.681, p < .001$).

Compared to clear speech, loud speech was produced with greater intensity (medium effect, $\hat{\beta} = -3.761, p < .001$), flatter tilt (medium effect, $\hat{\beta} = -1.981, p < .001$), and greater mid-frequency energy (medium effect; $\hat{\beta} = -5.6, p < .001$).

Interactions

Across all three speech styles and within the two groups, these patterns were largely maintained, as visible in Figure 4 and Figure 5. This is supported by an overall lack of two-way and three-way interactions between speech style and mask condition for most comparisons across the three outcome measures (see Table 5). Interactions between masks and speech styles were either non-significant or very small.

Discussion

The purpose of this study was to characterize the effects of masks and effortful speech styles on acoustic measures of spectral shape, tendency, and balance. Research Questions 1 and 2 sought to characterize the isolated effects of masks and speech styles, respectively, on spectral moments, while Research Question 3 characterized the combined effects of masks and speech styles on speech intensity, spectral tilt, and mid-range frequency energy.

Overall, face masks were found to attenuate higher frequency energy in the speech signal in similar ways for people with and without PD, across habitual, clear, and loud speech styles. Speaking loudly, followed by speaking clearly, was associated with increases in spectral balance measures (mid-range frequency and tilt), though these gains were limited when participants wore surgical or especially KN95 masks. As demonstrated in Figure 4, these increases resulting from clear or loud speech generally resulted in participants surpassing their spectral balance levels at baseline, i.e., habitual speech without a mask. In other words, for both older controls and people with PD, speaking more effortfully was generally successful in overcoming the damping effects of the masks with regards to the spectral outcomes addressed here. These findings are consistent with those found in the speech of young, typical controls in Knowles & Badh (2022). As can be seen in Figures 2 - 4, substantial individual variability exists. Our models accounted for this variability in the random effects structures, though we refer to effect sizes in the context of our reporting in order to compare the relative contributions of our variables of interest.

Effect of face masks on spectral moments in PD

Research Question 1 addressed the effects of surgical and KN95 face masks on the shape and tendency of the spectral density characteristics of speech of people with and without PD when using their habitual, everyday speaking style. Overall, results supported the generalization that masks attenuate higher frequency components of speech, and people with PD have lower concentrations of high frequency energy compared to controls at baseline.

In habitual speech, we observed statistically significant main effects of masks resulting in medium to large effect sizes for all spectral moments. Specifically, face masks, compared to no masks, were associated with medium to large decreases in first and second moments and medium to large increases in third and fourth moments when pooled across all speakers, and these patterns persisted in both speaker groups. This is consistent with previous literature documenting a low-pass filtering effect of masks on the speech signal (Corey et al., 2020; Maryn et al., 2021). Greater damping of higher frequency energy is associated with a lower overall spectral mean and subsequent variability (M1 and M2). Greater skewness (M3) is attributable to a right-sided tail of the LTAS, consistent with lower energy in higher frequencies. Greater kurtosis or peakedness (M4) captures a greater concentration of energy around the mean. Notably, these effects were consistent across both groups.

Of the two masks, the KN95 mask showed more attenuation than the surgical mask. The difference between the two masks were statistically significant for all spectral moments, though effect sizes were smaller than the difference between no masks and the two masks overall. This finding supports previous literature showing that attenuation amounts can vary by mask type and material (Corey et al., 2020; Maryn et al., 2021; Nguyen et al., 2021). However, previous studies have suggested that surgical masks and KN95 masks may have similar profiles of acoustic attenuation (Corey et al., 2020). For example, Corey et al. (2020) found that both surgical and KN95 masks demonstrated peak attenuation levels at 4 dB.

Statistically significant main effects of group for all moments but M2 provided evidence that the PD speakers have more low frequency energy compared to controls. However, these effects were small. The group differences are consistent with previous literature supporting lower concentrations of high frequency spectral energy in the speech of people with PD (Cushnie-Sparrow, 2021; Dromey, 2003; Smith & Goberman, 2014; though see Tjaden et al.,

2010). The pattern of the masks was consistent for both groups, as evidenced by the absence of any group by mask interactions.

Effect of effortful speech styles on spectral moments in PD

Our second research question sought to characterize the effect of speaking more clearly or more loudly on spectral moments in the absence of face masks. Our findings support previous literature that shows increased vocal effort is associated with not only greater speech intensity, but higher concentrations of energy in higher frequency components of the speech signal (Fant, 1960; Knowles & Badh, 2022; Krause & Braid, 2004; Smiljanic, 2021; Ternström et al., 2006). Compared to habitual speech, effortful speech in general was associated with a medium increase in M1 and small decreases in M3 and M4. No statistically significant difference was found for M2. Loud speech led to greater magnitudes of change for M1, M3, and M4 compared to clear speech, as evidenced by statistically significant comparisons between the speech styles. Effect sizes between clear and loud speech, however, were small. A small effect of group was found for M1, consistent with the results for Research Question 1 demonstrating overall lower spectral means for people with PD. The pattern of change in clear and loud speech across groups was very similar, evidenced again by an absence of group-by-speech style interactions for all spectral moments.

Combined effects of face masks and speech styles on spectral balance

Results from Research Questions 1 and 2 so far have affirmed the following: 1) people with PD have lower concentrations of high frequency spectral energy in their speech compared to controls, 2) face masks attenuate higher frequency spectral energy in speech, though do this to a similar extent for speech of people with and without PD, and 3) speaking more clearly and more loudly enhances spectral balance, and does so similarly for people with and without PD. Research Question 3 examined the combined effects of masks and speech styles on two measures of spectral balance as well as speech intensity. Overall, these results support those found in Research Questions 1 and 2. Small, statistically main significant effects of group were found for spectral tilt and mid-range frequency energy, but not speech intensity, which affirms that, even when all mask and speech styles were pooled together, participants with PD had poorer spectral balance compared to controls. The use of effortful speech, compared to habitual speech, resulted in medium to large increases for all three outcome measures. Loud speech resulted in medium-size increases compared to clear speech.

The masks were found to result in statistically significant decreases compared to no mask, but these differences were small compared to the change from clear and loud speech. Specifically, for both the No Mask versus Masks comparison and the surgical versus KN95 mask comparison, very small effect sizes were found for speech intensity and mid-range frequency energy, and small effect sizes were observed for spectral tilt.

Interactions between variables were either not statistically significant or were very small. In other words, the pattern of attenuation imposed by the masks was robust across speech styles, and the pattern of spectral enhancement resulting from effortful speech was robustly observed when people wore or did not wear face masks. This is consistent with the findings of our previous study which investigated young, typical speakers (Knowles & Badh, 2022). While distance from the talker was not a variable included in the present study, results from Knowles & Badh (2022) demonstrated an even larger effect of masks for spectral tilt at greater recording distances, i.e., at 2-meters. That is, a greater drop-off in spectral tilt was

observed when the talker was further away, such as during social distancing. This has potentially further repercussions for the impact on dysarthric speech signals.

Both speaker groups demonstrated similar degrees of change resulting from face masks and modifying their speech style. The effects of masks and more effortful speech result in the same *direction* and *magnitude* of effects for both groups, as evidenced by largely null interactions in Table 5. However, people with PD demonstrated lower baselines for most acoustic measures to begin with (main effect of group), and therefore do not benefit from clear and loud speech as much as they would without a mask. Masks have the effect of damping a poorer acoustic signal, and the effects of changing ones' speech to be clearer or louder are reduced for people who produce a poorer acoustic speech signal to begin with. This pattern is most readily observable for certain measures in Figure 4, in which the green dashed line reflects control baseline means, and the blue dotted line reflects PD baseline means. To even more clearly present this effect, we visualized mid-range frequency energy and spectral tilt as the difference compared to control participants' baselines in Figure 5. Take for example, spectral tilt, which reflects the difference in energy between 0 - 1 kHz and 1 - 10 kHz. In habitual speech, the PD group's tilt values fall below the control speakers (shown by the red dashed line), and even more so when wearing the masks. In clear speech *without* a mask, they approximate control speaker baseline and many participants exceed it (evidenced by upright triangle points), but still largely fall below the controls' baseline *with* a mask. Furthermore, when wearing a KN95 mask, all PD participants spectral tilt values fall below the controls' baseline, even in clear speech. In loud speech without a mask, most PD participants' tilt values surpass the control baseline, but once again many individuals still drop below when donning either of the masks, and especially the KN95 mask. This highlights the finding that the masks attenuate not only the acoustic signal, but also attenuate the gains made by producing clear and loud speech for people with PD. The present study provides an objective characterization of the acoustic consequences of face masks and the opposing consequences of effortful speech styles in the context of spectral balance deficits in parkinsonian speech. These acoustic characterizations are important in advancing an understanding of the nature of spectral balance in dysarthria and the mechanistic means of factors that may alter it.

When people with PD need to wear face masks, for example in medical settings, a tentative recommendation based on these findings would be to a) wear a less dense (but effective) mask such as a surgical mask and b) use loud speech techniques to compensate for the acoustic filtering effects and maximize speech transmission. Whether these changes in the acoustic signal have a functional impact on speech intelligibility or other perceptual outcomes is an open question that we are currently investigating. Given previous literature on the correspondence between spectral balance measures and perceived speech severity and intelligibility (Cushnie-Sparrow, 2021; Dromey, 2003; Tjaden et al., 2010), this very well may be the case. Cushnie-Sparrow (2021) found, for example, that spectral balance measures such as tilt were more predictive of perceptual estimations of loudness and intelligibility than overall intensity in the speech of people with and without PD. Furthermore, the authors found that synthetically amplifying mid and high frequency ranges of the speech signal while controlling for overall intensity were associated with increases in perceived loudness, suggesting that spectral balance plays an important role in multiple perceptual outcomes. Gutz et al. (2021) found that, for young, typical speakers, KN95 face masks did not result in reduced automatic speech recognition performance at baseline, but clear and loud speech were associated with improvements in word recognition when speakers wore masks. Smiljanic et al. (2021) found that native English listeners

recognized fewer words when speakers with foreign-accented speech wore masks, but this improved when talkers were instructed to speak more clearly.

Similar to challenges observed for listeners understanding foreign-accented speech with the additional degradation of a mask, people with atypical speech production, such as those with dysarthria who have reduced intelligibility at baseline, may be particularly difficult to understand. From a clinical perspective, if masks further impede the speech signal of people with dysarthria and standard behavioral interventions are limited by the impedance of the mask, different clinical approaches may be warranted. For example, it may be necessary for some individuals to use amplification devices in addition to using effortful speech styles when wearing a mask, whereas without a mask effortful speech alone may have been sufficient. A long term goal of this line of work is to identify acoustic predictors of perceptual challenges imposed by masks, including speech intelligibility and listener effort and to identify potential routes to better compensate as warranted. The present study provides an objective characterization of the acoustic consequences of face masks and the opposing consequences of effortful speech styles in the context of spectral balance deficits in parkinsonian speech. These acoustic characterizations are important in advancing an understanding of the nature of spectral balance in dysarthria and the mechanistic means of factors that may alter it. A long-term goal of this line of work is to identify acoustic predictors of perceptual challenges imposed by masks, including speech intelligibility and listener effort and to identify potential routes to better compensate as warranted.

These and future findings may also shed light on how filter characteristics due to vocal tract configurations or face coverings such as masks versus glottal source characteristics can impact spectral balance in speech. For example, one outcome of the present study is that while both face masks and effortful speech impact spectral balance in opposing ways, the masks have a relatively small effect compared to increased phonatory effort. Future work that examines kinematic or aerodynamic adjustments made when individuals wear masks could help to inform these contributions from a phonatory-articulatory standpoint. Integrating different mask types, such as more reverberant (e.g., plastic) or restrictive (e.g., N95s) masks is also an avenue to explore the resultant spectral outcomes. Findings from this line of work may provide guidance as to specific behavioral changes that individuals with or without speech disorders may be able to make in order to maximize the ability to compensate for the masks.

While face masks are less frequently worn now than at the height of the COVID-19 pandemic, anecdotal reports from health officials warn that increased masking may be warranted in future waves. Furthermore, masks and social distancing are commonly still mandated in many health care settings, which are places in which people with neurodegenerative diseases are more likely to access than the neurotypical population. Accurate communication in such settings is important, and thus an understanding of the impact of personal protective measures on speech of people with speech and voice disorders is of value from a clinical and public health perspective.

Limitations

While these results provide further evidence to established patterns in the literature of the filtering effects of masks and the acoustic benefits of effortful speech, several caveats are warranted. These findings reflect a relatively small speaker group with a large degree of heterogeneity in PD speech symptoms and severity. Future work would benefit from modelling the contributions of severity and prominent symptom clusters (i.e., hypophonia)

in spectral acoustic outcomes. Relatedly, while all participants were given identical instructions in how to modify their speech styles consistent with previous studies (Smiljanic, 2021; i.e., Tjaden & Wilding, 2004), it is likely that the precise implementations varied. Participants may have made different unconscious adjustments to their speech when wearing masks as well, though previous literature suggests these subconscious mask-related adjustments may be small (Gutz et al., 2021) and may be more likely to occur when speakers are intentionally trying to sound clearer (Cohn et al., 2021). While counterbalancing of the speech style conditions and randomization of the mask order was used to control for these potential differences here, more systematic manipulations of acoustic changes (i.e., through the use of synthesized or amplified speech) would help to tease these differences apart in future studies. Another caveat is that while the speech stimuli included phonetically balanced sentence lists, some utterances needed to be discarded due to production errors. While this affected a relatively small number of utterances overall (4%), this may have had an impact on the phonetic characteristics and subsequent acoustic outcomes. Face coverings have been shown to have differential effects on fricatives, for example, by lowering the spectral means of /f/ and /θ/, which have relatively flat spectral density characteristics, but not /s/ and /ʃ/, marked by concentrations of high frequency noise (Fecher & Watt, 2011). Further controlled accounts of the impact of masks on specific acoustic-phonetic characteristics is warranted.

Conclusion

In summary, findings from the present study can be summarized as follows. Face masks attenuate higher frequency energy in speech, and this is similarly true for people with and without PD, people with PD have lower concentrations of high frequency energy compared to. Speaking more clearly, followed by more loudly, is associated with increases in higher frequency energy, which has the effect of remediating the damping imposed by masks, and this is again true for people with and without PD. The effects of masks are preserved across habitual and effortful speech styles, though masks limit the gains achieved by making speech clearer and louder. These limitations may be especially detrimental for talkers with PD, who demonstrate lower levels of spectral balance to begin with. Future research should address the impact of masks on perceptual outcomes for people with dysarthria, as well as alternative or augmentative means of remediating such challenges.

Data availability statement

De-identified data and code will be publicly available following the publication of this manuscript.

References

- Adams, S. G., & Dykstra, A. D. (2009). Hypokinetic dysarthria. In M. R. McNeil (Ed.), *Clinical management of sensorimotor speech disorders*. Thieme Publishing Group.
- Alharbi, G. G., Cannito, M. P., Buder, E. H., & Awan, S. N. (2019). Spectral/Cepstral Analyses of Phonation in Parkinson's Disease before and after Voice Treatment: A Preliminary Study. *Folia Phoniatrica Et Logopaedica*, 71(5-6), 275–285.
<https://doi.org/10.1159/000495837>

- Atcherson, S. R., McDowell, B. R., & Howard, M. P. (2021). Acoustic effects of non-transparent and transparent face coverings. *The Journal of the Acoustical Society of America*, 149(4), 2249–2254. <https://doi.org/10.1121/10.0003962>
- Atcherson, S., Finley, E., McDowell, B., & Watson, C. (2020). *More Speech Degradations and Considerations in the Search for Transparent Face Coverings During the COVID-19 Pandemic*. <https://www.audiology.org/audiology-today-novemberdecember-2020/more-speech-degradations-and-considerations-search-transparent>
- Badh, G., & Knowles, T. (2022). *Acoustic and perceptual impact of face masks on speech: A scoping review*. <https://doi.org/10.31219/osf.io/fpsu2>
- Baylor, C., Yorkston, K., Eadie, T., Kim, J., Chung, H., & Amtmann, D. (2013). The Communicative Participation Item Bank (CPIB): Item bank calibration and development of a disorder-generic short form. *Journal of Speech, Language, and Hearing Research*.
- Boersma, P., & Weenink, D. (2021). *Praat: Doing phonetics by computer [Computer program]* (Version 6.1.35) [Computer software]. <http://www.praat.org/>
- Bottalico, P., Murgia, S., Puglisi, G. E., Astolfi, A., & Kirk, K. I. (2020). Effect of masks on speech intelligibility in auralized classrooms. *The Journal of the Acoustical Society of America*, 148(5), 2878.
- Brown, V. A., Van Engen, K. J., & Peelle, J. E. (2021). Face mask type affects audiovisual speech intelligibility and subjective listening effort in young and older adults. *Cognitive Research: Principles and Implications*, 6(49). <https://doi.org/gn22xd>
- Brysbaert, M., & Stevens, M. (2018). Power Analysis and Effect Size in Mixed Effects Models: A Tutorial. *Journal of Cognition*, 1(1, 1), 1–20. <https://doi.org/10.5334/joc.10>
- Cohen, J. (1962). The statistical power of abnormal-social psychological research: A review. *The Journal of Abnormal and Social Psychology*, 65(3), 145–153. <https://doi.org/10.1037/h0045186>
- Cohn, M., Pycha, A., & Zellou, G. (2021). Intelligibility of face-masked speech depends on speaking style: Comparing casual, clear, and emotional speech. *Cognition*, 210, 104570. <https://doi.org/10.1016/j.cognition.2020.104570>
- Corey, R. M., Jones, U., & Singer, A. C. (2020). Acoustic effects of medical, cloth, and transparent face masks on speech signals. *The Journal of the Acoustical Society of America*, 148(4), 2371. <https://doi.org/10.1121/10.0002279>
- Cushnie-Sparrow, D. A. (2021). *Modelling loudness: Acoustic and perceptual correlates in the context of hypophonia in Parkinson's disease*. Western University.

Dalrymple-Alford, J., MacAskill, M., Nakas, C., Livingston, L., Graham, C., Crucian, G., Melzer, T., Kirwan, J., Keenan, R., Wells, S., et al. (2010). The MoCA well-suited screen for cognitive impairment in Parkinson disease. *Neurology*, 75(19), 1717–1725.

Darley, F. L., Aronson, A. E., & Brown, J. R. (1969a). Differential diagnostic patterns of dysarthria. *Journal of Speech, Language, and Hearing Research*, 12(2), 246. <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, Language, and Hearing Research*, 12(3), 462–496. <https://doi.org/doi:10.1044/jshr.1203.462>

Dromey, C. (2003). Spectral measures and perceptual ratings of hypokinetic dysarthria. *Journal of Medical Speech-Language Pathology*, 11(2), 85–94.

Dromey, C., & Ramig, L. O. (1998). Intentional changes in sound pressure level and rate: Their impact on measures of respiration, phonation, and articulation. *Journal of Speech, Language, and Hearing Research*, 41(5), 1003–1018. <https://doi.org/10.1044/jslhr.4105.1003>

Duffy, J. R. (2019). *Motor speech disorders e-book: Substrates, differential diagnosis, and management*. Elsevier Health Sciences.

Fant, G. (1960). *Acoustic theory of speech production*. Mouton.

Fecher, N., & Watt, D. (2011). Speaking under Cover: The Effect of Face-concealing Garments on Spectral Properties of Fricatives. *ICPhs*. https://www.researchgate.net/profile/Dominic_Watt/publication/228766449_Speaking_under_cover_The_effect_of_face-concealing_garments_on_spectral_properties_of_fricatives/links/004635148963871e3f000000.pdf

Gilbert, R. C., Chandrasekaran, B., & Smiljanic, R. (2014). Recognition memory in noise for speech of varying intelligibility. *The Journal of the Acoustical Society of America*, 135(1), 389–399. <https://doi.org/10.1121/1.4838975>

Goldin, A., Weinstein, B., & Shiman, N. (2020). How Do Medical Masks Degrade Speech Reception? *Hearing Review*, 27(5), 8–9. <https://www.hearingreview.com/hearing-loss/health-wellness/how-do-medical-masks-degrade-speech-reception>

Gregory, N. D., Chandran, S., Lurie, D., & Sataloff, R. T. (2012). Voice Disorders in the Elderly. *Journal of Voice*, 26(2), 254–258. <https://doi.org/10.1016/j.jvoice.2010.10.024>

Guskaroska, A. (2021). *The role of face masks on speech intelligibility for native and nonnative speakers of English: Graduate students' experiences*. 117–133.

- Gutz, S., Rowe, H., & Green, J. (2021). Speaking with a KN95 Face Mask: ASR Performance and Speaker Compensation. *Proceedings of Interspeech 2021*, 4798–4802. <https://doi.org/10.21437/Interspeech.2021-99>
- Hazan, V., & Baker, R. (2011). Acoustic-phonetic characteristics of speech produced with communicative intent to counter adverse listening conditions. *The Journal of the Acoustical Society of America*, 130(4), 2139–2152.
- Hazan, V., & Markham, D. (2004). Acoustic-phonetic correlates of talker intelligibility for adults and children. *The Journal of the Acoustical Society of America*, 116(5), 3108–3118. <https://doi.org/10.1121/1.1806826>
- Hazan, V., Tuomainen, O., Kim, J., Davis, C., Sheffield, B., & Brungart, D. (2018). Clear speech adaptations in spontaneous speech produced by young and older adults. *The Journal of the Acoustical Society of America*, 144(3), 1331–1346. <https://doi.org/10.1121/1.5053218>
- Hillenbrand, J., Cleveland, R. A., & Erickson, R. L. (1994). Acoustic Correlates of Breathy Vocal Quality. *Journal of Speech, Language, and Hearing Research*, 37(4), 769–778. <https://doi.org/10.1044/jshr.3704.769>
- Hillenbrand, J., & Houde, R. A. (1996). Acoustic Correlates of Breathy Vocal Quality: Dysphonic Voices and Continuous Speech. *Journal of Speech, Language, and Hearing Research*, 39(2), 311–321. <https://doi.org/10.1044/jshr.3902.311>
- IEEE. (1969). IEEE Subcommittee on Subjective Measurements IEEE Recommended Practices for Speech Quality Measurements. *IEEE Transactions on Audio and Electroacoustics*, 17, 225–246. <https://doi.org/10.1109/ieeestd.1969.7405210>
- 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
- Kendall, K. (2007). Presbyphonia: A review. *Current Opinion in Otolaryngology & Head and Neck Surgery*, 15(3), 137–140. <https://doi.org/10.1097/MOO.0b013e328166794f>
- Knowles, T., Adams, S. G., & Jog, M. (2021). Variation in Speech Intelligibility Ratings as a Function of Speech Rate Modification in Parkinson's Disease. *Journal of Speech, Language, and Hearing Research*, 64(6), 1773–1793. https://doi.org/10.1044/2021_JSLHR-20-00593
- Knowles, T., & Badh, G. (2022). The impact of face masks on spectral acoustics of speech: Effect of clear and loud speech styles. *J. Acoust. Soc. Am.*, 11. <https://doi.org/10.1121/10.0011400>

- Krause, J. C., & Braida, L. D. (2004). Acoustic properties of naturally produced clear speech at normal speaking rates. *The Journal of the Acoustical Society of America*, 115(1), 362–378. <https://doi.org/10.1121/1.1635842>
- Krause, J. C., & Braida, L. D. (2009). Evaluating the role of spectral and envelope characteristics in the intelligibility advantage of clear speech. *The Journal of the Acoustical Society of America*, 125(5), 3346–3357. <https://doi.org/10.1121/1.3097491>
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, 82(13), 1–26. <https://doi.org/10.18637/jss.v082.i13>
- Logemann, J. A., Fisher, H. B., Boshes, B., & Blonsky, E. R. (1978). Frequency and cooccurrence of vocal tract dysfunctions in the speech of a large sample of Parkinson patients. *The Journal of Speech and Hearing Disorders*, 43(1), 47–57. <https://doi.org/10.1044/jshd.4301.47>
- Lu, Y., & Cooke, M. (2009). The contribution of changes in F0 and spectral tilt to increased intelligibility of speech produced in noise. *Speech Communication*, 51(12), 1253–1262. <https://doi.org/10.1016/j.specom.2009.07.002>
- Lüdecke, D. (2021). *sjPlot: Data Visualization for Statistics in Social Science*. <https://CRAN.R-project.org/package=sjPlot>
- Magee, M., Lewis, C., Noffs, G., Reece, H., Chan, J. C., Zaga, C. J., Paynter, C., Birchall, O., Azocar, S. R., Ediriweera, A., et al. (2020). Effects of face masks on acoustic analysis and speech perception: Implications for peri-pandemic protocols. *bioRxiv*.
- Martel, -S. V., Breton, M., Chabot, A., & Langlois, M. (2021). The Impact of Clear Speech on the Perceptual and Acoustic Properties of Fricative–Vowel Sequences in Speakers With Dysarthria. *American Journal of Speech-Language Pathology*, 30, 1410–1428. https://doi.org/10.1044/2021_AJSLP-20-00157
- Martins, R. H. G., Gonçalves, T. M., Pessin, A. B. B., & Branco, A. (2014). Aging voice: presbyphonia. *Aging Clinical and Experimental Research*, 26(1), 1–5. <https://doi.org/10.1007/s40520-013-0143-5>
- Maryn, Y., Wuyts, F. L., & Zarowski, A. (2021). Are Acoustic Markers of Voice and Speech Signals Affected by Nose-and-Mouth-Covering Respiratory Protective Masks? *Journal of Voice*, S0892199721000370. <https://doi.org/10.1016/j.jvoice.2021.01.013>
- McKenna, V. S., Kendall, C. L., Patel, T. H., Howell, R. J., & Gustin, R. L. (2022). Impact of Face Masks on Speech Acoustics and Vocal Effort in Healthcare Professionals. *The Laryngoscope*, 132(2), 391–397. <https://doi.org/10.1002/lary.29763>
- Mefferd, A. S. (2017). Tongue- and Jaw-Specific Contributions to Acoustic Vowel Contrast Changes in the Diphthong /ai/ in Response to Slow, Loud, and Clear

Speech. *Journal of Speech, Language, and Hearing Research*, 60(11), 3144–3158.
https://doi.org/10.1044/2017_JSLHR-S-17-0114

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-MS18-18-0192

Mendel, L. L., Gardino, J. A., & Atcherson, S. R. (2008). Speech understanding using surgical masks: A problem in health care? *Journal of the American Academy of Audiology*, 19(9), 686–695. <https://doi.org/10.3766/jaaa.19.9.4>

Miller, N., Allcock, L., Jones, D., Noble, E., Hildreth, A. J., & Burn, D. J. (2007). Prevalence and pattern of perceived intelligibility changes in Parkinson's disease. *Journal of Neurology, Neurosurgery, and Psychiatry*, 78(11), 1188–1190. <https://doi.org/10.1136/jnnp.2006.110171>

Müller, J., Wenning, G. K., Verny, M., McKee, A., Chaudhuri, K. R., Jellinger, K., Poewe, W., & Litvan, I. (2001). Progression of dysarthria and dysphagia in postmortem-confirmed parkinsonian disorders. *Archives of Neurology*, 58(2), 259–264. <https://doi.org/10.1001/archneur.58.2.259>

Mutch, W. J., Strudwick, A., Roy, S. K., & Downie, A. W. (1986). Parkinson's disease: Disability, review, and management. *British Medical Journal*, 293(6548), 675–677. <https://doi.org/10.1136/bmj.293.6548.675>

Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J. L., & Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, 53(4), 695–699.

Neel, A. (2009). Effects of loud and amplified speech on sentence and word intelligibility in Parkinson disease. *Journal of Speech, Language, and Hearing Research*.

Nguyen, D. D., McCabe, P., Thomas, D., Purcell, A., Doble, M., Novakovic, D., Chacon, A., & Madill, C. (2021). Acoustic voice characteristics with and without wearing a facemask. *Scientific Reports*, 11(1), 5651. <https://doi.org/10.1038/s41598-021-85130-8>

Oosthuizen, I., Saunders, G. H., Manchaiah, V., & Swanepoel, D. W. (2022). Impact of SARS-CoV-2 Virus (COVID-19) Preventative Measures on Communication: A Scoping Review. *Frontiers in Public Health*, 10. <https://doi.org/10.3389/fpubh.2022.815259>

Ostwald, P. F. (1960). Visual Denotation of Human Sounds: Preliminary Report of an Acoustic Method | JAMA Psychiatry | JAMA Network. *Archives of General Psychiatry*, 3(2), 117–121. <https://doi.org/10.1001/archpsyc.1960.01710020001001>

- Palmiero, A. J., Symons, D., Morgan, J. W. 3rd., & Shaffer, R. E. (2016). Speech intelligibility assessment of protective facemasks and air-purifying respirators. *Journal of Occupational and Environmental Hygiene*, 13(12), 960–968. <https://doi.org/10.1080/15459624.2016.1200723>
- Park, S., Theodoros, D., Finch, E., & Cardell, E. (2016). Be Clear: A New Intensive Speech Treatment for Adults With Nonprogressive Dysarthria. *American Journal of Speech-Language Pathology*, 25(1), 97–110. https://doi.org/10.1044/2015_AJSLP-14-0113
- Pörschmann, C., Lübeck, T., & Arend, J. M. (2020). Impact of face masks on voice radiation. *The Journal of the Acoustical Society of America*, 148(6), 3663–3760. <https://doi.org/10.1121/10.0002853>
- R Core Team. (2020). *R: A language and environment for statistical computing* [Manual]. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Ramig, L. O., Fox, C., & Sapir, S. (2004). Parkinson's disease: Speech and voice disorders and their treatment with the Lee Silverman Voice Treatment. *Seminars in Speech and Language*, 25, 169–180.
- Schulman, R. (1989). Articulatory dynamics of loud and normal speech. *The Journal of the Acoustical Society of America*, 85(1), 295–312. <https://doi.org/10.1121/1.397737>
- Shin, H., Shivabasappa, P., & Koul, R. (2022). Effect of clear speech intervention program on speech intelligibility in persons with idiopathic Parkinson's disease: A pilot study. *International Journal of Speech-Language Pathology*, 24(1), 33–41. <https://doi.org/10.1080/17549507.2021.1943522>
- Smiljanic, R. (2021). Clear Speech Perception. In *The Handbook of Speech Perception* (pp. 177–205). John Wiley & Sons, Ltd. <https://doi.org/10.1002/9781119184096.ch7>
- Smiljanic, R., Keerstock, S., Meemann, K., & Ransom, S. M. (2021). Face masks and speaking style affect audio-visual word recognition and memory of native and non-native speech. *The Journal of the Acoustical Society of America*, 149(6), 4013–4023. <https://doi.org/10.1121/10.0005191>
- Smith, L. K., & Goberman, A. M. (2014). Long-time average spectrum in individuals with Parkinson Disease1. *NeuroRehabilitation*, 35(1), 77–88. <https://doi.org/10.3233/NRE-141102>
- Švec, J. G., & Granqvist, S. (2018). Tutorial and guidelines on measurement of sound pressure level in voice and speech. *Journal of Speech, Language, and Hearing Research*, 61(3), 441–461.
- Tanner, K., Roy, N., Ash, A., & Buder, E. (2005). Spectral Moments of the Long-term Average Spectrum: Sensitive Indices of Voice Change After Therapy? *Journal of Voice*

: *Official Journal of the Voice Foundation*, 19, 211–222.
<https://doi.org/10.1016/j.jvoice.2004.02.005>

Ternström, S., Bohman, M., & Södersten, M. (2006). Loud speech over noise: Some spectral attributes, with gender differences. *The Journal of the Acoustical Society of America*, 119(3), 1648–1665. <https://doi.org/10.1121/1.2161435>

Thomas, F., Allen, C., Butts, W., Rhoades, C., Brandon, C., & Handrahan, D. L. (2011). Does wearing a surgical facemask or N95-respirator impair radio communication? *Air Medical Journal*, 30(2), 97–102. <https://doi.org/10.1016/j.amj.2010.12.007>

Tjaden, K. (2008). Speech and swallowing in Parkinson's disease. *Topics in Geriatric Rehabilitation*, 24(2), 115–126.
<https://doi.org/10.1097/01.TGR.0000318899.87690.44>

Tjaden, K., Richards, E., Kuo, C., Wilding, G., & Sussman, J. (2013). Acoustic and perceptual consequences of clear and loud speech. *Folia Phoniatrica Et Logopaedica*, 65(4), 214–220. <https://doi.org/10.1159/000355867>

Tjaden, K., Sussman, J. E., Liu, G., & Wilding, G. (2010). Long-Term Average Spectral (LTAS) Measures of Dysarthria and Their Relationship to Perceived Severity. *Journal of Medical Speech-Language Pathology*, 18(4), 125–132.
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5800529/>

Tjaden, K., Sussman, J. E., & Wilding, G. E. (2014). Impact of Clear, Loud, and Slow Speech on Scaled Intelligibility and Speech Severity in Parkinson's Disease and Multiple Sclerosis. *Journal of Speech, Language, and Hearing Research : JSLHR*, 57(3), 779–792. https://doi.org/10.1044/2014_JSLHR-S-12-0372

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\(2004/058\)](https://doi.org/10.1044/1092-4388(2004/058))

Toscano, J. C., & Toscano, C. M. (2021). Effects of face masks on speech recognition in multi-talker babble noise. *PloS One*, 16(2), e0246842.

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>

Westfall, J., Kenny, D. A., & Judd, C. M. (2014). Statistical power and optimal design in experiments in which samples of participants respond to samples of stimuli. *Journal of Experimental Psychology: General*, 143(5), 2020–2045.
<https://doi.org/10.1037/xge0000014>

Whelan, B.-M., Theodoros, D., Cahill, L., Vaezipour, A., Vogel, A. P., Finch, E., Farrell, A., & Cardell, E. (2022). Feasibility of a Telerehabilitation Adaptation of the Be Clear Speech Treatment Program for Non-Progressive Dysarthria. *Brain Sciences*, 12(2, 2), 197. <https://doi.org/10.3390/brainsci12020197>

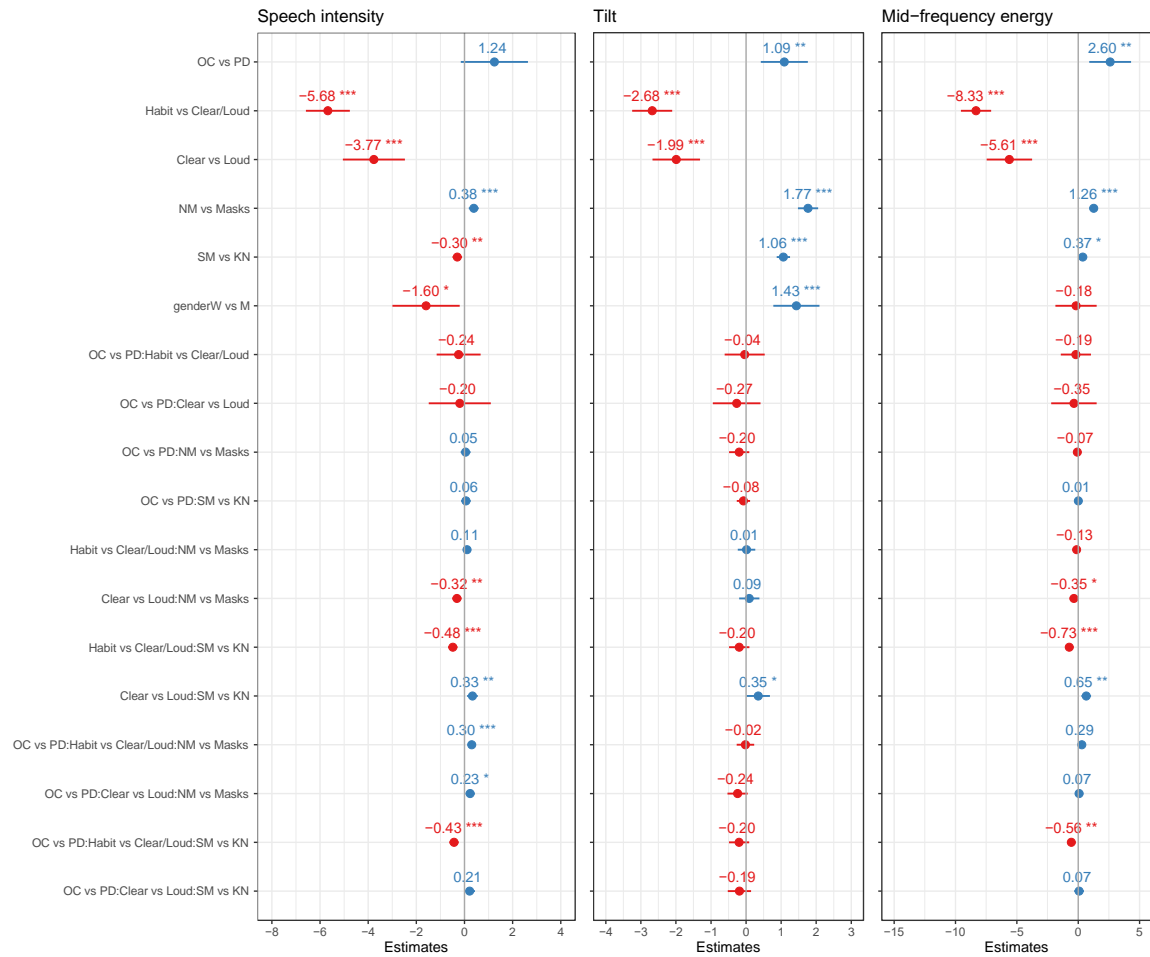
Wong, M. N., Kuruvilla-Dugdale, M., & Ng, M. L. (2017). Differential effects of increased loudness on tongue kinematics in individuals with PD? Analyses of two cases. *Speech, Language and Hearing, 20*(1), 49–56.
<https://doi.org/10.1080/2050571X.2016.1206694>

Yorkston, K. M., Hakel, M., Beukelman, D. R., & Fager, S. (2007). Evidence for effectiveness of treatment of loudness, rate, or prosody in dysarthria: A systematic review. *Journal of Medical Speech-Language Pathology, 15*(2).

Appendix

Appendix Table 1: Mean values (averaged over the individual participants) for speech outcomes across all groups and conditions.

Speech style	Group	Face mask	M1	M2	M3	M4	Intensity	Tilt	1 - 3 kHz energy
habitual	Control	No mask	646.4602	970.3593	5.578245	50.53543	79.74911	-18.36070	11.741967
		Surgical	523.0858	675.3347	6.433631	74.51831	78.61748	-19.57154	10.191759
		KN95	491.4189	563.4630	7.180618	105.60692	79.47042	-20.30952	10.655609
	PD	No mask	521.4364	931.6915	6.701912	77.03431	77.25870	-20.19686	6.630286
		Surgical	414.5218	632.7531	7.974372	120.62346	76.79416	-21.59756	5.648980
		KN95	372.2391	484.6455	9.605673	217.67661	77.24077	-22.79627	5.408419
clear	Control	No mask	732.2689	971.8249	5.094572	43.05932	83.43687	-16.88721	17.072927
		Surgical	634.8242	730.5256	5.556000	55.74045	83.26155	-17.78595	16.557224
		KN95	568.6012	608.5575	6.048198	74.75660	82.96386	-18.87605	15.509653
	PD	No mask	617.7304	926.7529	5.889434	58.99163	80.82368	-18.36184	12.204930
		Surgical	503.7631	666.9715	6.544623	86.94354	80.60143	-19.77856	11.361491
		KN95	446.4003	493.7634	7.490336	136.17956	80.85511	-21.19114	10.619552
loud	Control	No mask	875.0511	949.1581	4.409286	33.87972	87.41414	-14.48151	23.212247
		Surgical	772.2631	725.4552	4.757359	45.55467	86.98615	-15.60453	22.173954
		KN95	697.2946	607.7710	5.084900	56.59800	87.17517	-16.64336	21.711761
	PD	No mask	701.7811	889.9908	5.346931	50.47876	84.78722	-16.85509	17.803399
		Surgical	599.6839	632.4288	5.576339	65.66757	83.89483	-18.21742	16.152552
		KN95	545.5409	532.0934	5.907378	79.17944	84.30322	-19.07512	16.036000



Appendix Figure A1. Forest plot of model outcomes. PD = Parkinson's disease; OC = older control; NM = no masks; SM = surgical mask; KN = KN95