Effect of face masks and speech style on speech intelligibility and listener effort in Parkinson's disease

Nathaniel Cline[1], Thea Knowles[2], Gursharan Badh[1]

[1]Department of Communicative Disorders and Sciences, University at Buffalo, Buffalo, NY;
[2]Communicative Sciences and Disorders, Michigan State University, East Lancing, MI

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 The aim of this study was to quantify the combined effects of face masks and effortful speech styles on listener intelligibility and perceived listener effort in talkers with and without Parkinson's disease (PD).

Method Ten people with Parkinson's disease and 10 healthy, older controls read aloud sentences in two face mask and three speech style conditions. Masks included no mask and KN95 masks. Speech styles included habitual, clear, and loud. Listener participants were tasked with listening to each sentence mixed with background noise and then transcribing what they heard and rating how effortful it was to understand. Listener accuracy and effort were each modeled as a function of speaker group, face mask, and speech style using mixed effects regression models.

Results Listeners were less accurate and reported greater listening effort for the PD group and for the mask condition. Listeners were more accurate and reported less effort when listening to clear and loud compared to habitual speech. Listener accuracy and listener effort were strongly negatively correlated across all conditions. Face masks were also associated with a steeper decline in speech intelligibility and an increase in listener effort for talkers with PD.

Discussion Face masks resulted in steeper speech intelligibility decline for talkers with PD compared to controls. Speaking more loudly *or* more clearly when wearing a face mask improved intelligibility for talkers with PD compared to habitual speech, and both speech styles resulted in speech intelligibility levels that approximated talkers' baseline intelligibility levels without a mask.

Introduction

Effect of face masks on speech

In response to the surge of the COVID-19 pandemic, the Centers for Disease Control and Prevention (CDC) recommended that the population at large begin wearing face masks to prevent the spread of the disease. Unfortunately, face masks have been shown to negatively impact intelligibility of masked speakers by degrading the speech signal (Cohn et al., 2021; Toscano & Toscano, 2021). In quiet environments, most face masks have a negligible effect on intelligibility, but in the presence of competing background noise, masks have a pronounced effect on intelligibility (Brown et al., 2021; Carraturo et al., 2021; Kim & Thompson, 2022; Rahne et al., 2021). Not only does the impedance of the acoustic signal impact intelligibility, but the obscurance of the mouth plays a role as well; although transparent face masks severely impact the acoustic signal, the availability of visual cues may partially make up for it (Brown et al., 2021; Yi et al., 2021). The degree of this impact varies by mask type and likely depends on the material and construction of the mask; for instance, surgical masks seem to have a minimal effect on speech acoustics due to their thin material, while transparent masks show the greatest attenuation as they are made from acoustically reflective material, and most other masks (such as KN95s and cloth masks) tend to fall somewhere in the middle (Corey et al., 2020). KN95s are similar to N95s in construction and that they are a type of disposable respirator effective in filtering out small particles, but they differ in that they are not approved by the National Institute for Occupational Safety and Health (CDC, 2022). Not only do masks reduce intelligibility, but they also increase the effort required for listeners to comprehend masked talkers' speech (Carraturo et al., 2021). Populations who demonstrated reduced speech intelligibility to begin with, such as those with speech disorders, may be at a further risk of communication difficulties when wearing masks.

There are a number of underlying acoustic mechanisms at play here. Possibly the most prevalent of these is that masks have been shown to act as a low-pass filter, attenuating frequencies above approximately 1 kHz (Corey et al., 2020; Maryn et al., 2021). Use of face masks is also associated with steeper spectral tilt, less energy in midrange frequencies, and a small reduction in speech intensity (Knowles & Badh, 2022, 2023; Nguyen et al., 2021). Knowles and Badh (2022, 2023) also measured four spectral moments (center of gravity, standard deviation of center of gravity, skewness, and kurtosis), each of which was significantly altered in the presence of the KN95 mask compared to the surgical mask, supporting an overall attenuation of high-frequency energy in the signal. Others, however, report different findings, as Maryn et al. (2021) found no significant effects of masks on spectral moments of prerecorded vowel prolongations.

Speech and speech intelligibility in Parkinson's disease

The majority of people with idiopathic Parkinson's disease (PD) develop hypokinetic dysarthria (J. A. Logemann et al., 1978; Müller et al., 2001; Mutch et al., 1986), a speech disorder associated with a reduction in the mobility of muscle movements used in speech, particularly shown in abnormally smaller and less forceful speech movements (Duffy, 2019). Perceptual symptoms of hypokinetic dysarthria include reduced loudness (hypophonia; Adams & Dykstra, 2009; Ludlow & Bassich, 1984) and breathy or hoarse voice quality (J. Logemann et al., 1973). Acoustically, these can also be described in terms of reduced vocal intensity (Adams et al., 2005; Fox & Ramig, 1997; Ho et al., 1999) and spectral attenuation (Cushnie-Sparrow, 2021). Hypokinetic dysarthria is also associated with deficits related to imprecise articulation, such as distortion of stops, affricates, and fricatives, apparently resulting from an inadequate narrowing of the vocal tract (J. A. Logemann & Fisher, 1981), voicing during the typically voiceless closure interval of voiceless stops (Weismer, 1984), and longer voice onset times (Forrest et al., 1989).

All of these factors contribute to reduced speech intelligibility, which may be further compounded by the addition of face masks. Hypophonia is involved in reduced intelligibility, communicative effectiveness, and communicative participation (Adams & Dykstra, 2009; Tjaden, 2008), with loudness and spectral balance being critical components of the cluster of symptoms (Cushnie-Sparrow, 2021). This, paired with symptoms such as abnormal voice quality, monopitch, monoloudness, reduced stress and intonation patterns, and overall lack of articulatory precision, leads to reduced speech intelligibility. Considering the attenuation that masks impose, the addition of face masks to people with PD results in further reduction of speech intelligibility.

Clear and loud speech strategies in PD:

For talkers with and without PD, altering one's speech style to be clearer or louder can improve speech intelligibility (e.g., Tjaden et al., 2013, 2014). Recent evidence has shown this to also be the case for neurologically healthy talkers when wearing face masks (Knowles & Badh, 2022; Yi et al., 2021). Clear and loud speech styles, compared to habitual speech, have been shown to result in similar spectral changes, reflecting increased vocal effort (Rosenthal et al., 2014), such as being produced with greater speech intensity compared to habitual speech (with this increase being greater in loud speech; Tjaden et al., 2013) and having increased energy in higher frequency ranges, leading to flatter spectral slopes (which have been attributed to greater relative energy in the first formant range; Fant, 1960; Ternström et al., 2006). When talkers are instructed to speak clearly, speech is associated with an increase in energy in midrange frequencies (i.e., 1–3kHz; Gilbert et al., 2014; Hazan & Baker, 2011; Hazan et al., 2018; Krause & Braida, 2004, 2009; Smiljanic, 2021), although different instructions may have different results (Lam et al., 2012; Stipancic et al., 2022). Overall, in neurologically healthy talkers, both clear and loud speaking styles have also been shown to yield better intelligibility in the context of face masks (Cohn et al., 2021; Gutz et al., 2021; Smiljanic, 2021; Yi et al., 2021). However, there is reason to believe that speakers may be subconsciously altering their speech style to accommodate while wearing masks, instinctively speaking louder and clearer to overcome the obstacle that masks have on conversation (Cohn et al., 2021).

For people with PD, clear and loud speaking styles have been shown to be effective in improving speech intelligibility (Neel, 2009; Tjaden et al., 2013, 2014). For people with PD who need to wear face masks, such as in hospital settings where masks are still often required, clear or loud speech might reduce the compounded detriment on their speech intelligibility.

Acoustic outcomes of effortful speech styles and face masks in PD

This study builds on our previous report of the acoustic effects of face masks and effortful speech styles in the same group of talkers with and without PD (Knowles & Badh, 2023) as well as in younger talkers (Knowles & Badh, 2022). Specifically, this study utilizes the same speech production data as Knowles and Badh (2023), providing perceptual data while the previous study focused on acoustic data. Our previous work demonstrated that clear and loud speech styles lead to increases in relative high-frequency energy in the speech spectrum, demonstrated by changes in spectral moments, spectral tilt, increases in energy in the 1- to 3-kHz range, and speech intensity. In particular, loud compared to clear speech was produced with greater intensity, flatter tilt, and greater midfrequency energy. Face masks were shown to have an opposing effect, leading to decreases in all of these measures, although the relative effect sizes were smaller for masks compared to speech styles. These findings were consistent across speakers with and without PD (Knowles & Badh, 2022; 2023). This work extends these study questions to changes in auditory–perceptual speech outcomes. To the authors' knowledge, there are no reports of the effects of face masks on clinical speaker populations. Given the continued use of face masks in health care settings (at the time of writing) that these individuals must access, a holistic understanding of potential communication challenges and remediation strategies is vital to ensure optimal clinical care. Furthermore, an understanding of the effects of acoustic filtration such as that imposed by face masks on the speech of individuals with speech disorders offers a window into how changes to the speech signal contribute to perceptual outcomes.

Summary and Purpose

In summary, masks have been shown to impair speech intelligibility in large part due to attenuating frequencies above approximately 1 kHz. This impairment on speech intelligibility is worse for people who suffer from dysarthria as a result of neurological disorders such as PD, as they already experience reduced speech intelligibility compared to controls. Fortunately, clear and loud speech styles have been shown to be effective in combating both the impairment imposed by masks (Gutz et al., 2021, 2022) and the impairment imposed by hypokinetic dysarthria (Tjaden et al., 2013, 2014).

Our purpose with this study is twofold: first, to quantify the effects of face masks on speech intelligibility and perceived listener effort in talkers with and without PD and, second, to examine the effects of clear and loud speech on these perceptual outcomes in the context of face masks.

Based on the previous literature regarding face masks, as well as our previous work on speech acoustics of masks and speech styles in PD, we predict the following. Predictions 1–3 pertain to whether speech intelligibility and listener effort are affected by our conditions, while Prediction 4 integrates our previous acoustic findings to describe the underlying nature of any observed changes.

1. People with PD will be overall less intelligible and more effortful to understand than controls. This difference may or may not widen when face masks are worn.

2. Habitual speech will be less intelligible and more effortful to understand than clear and loud speech. Based on previous literature (e.g., Tjaden et al., 2013; 2014), clear speech is predicted to be more intelligible and easier to understand than loud and habitual speech in both participant groups. This difference may or may not change when face masks are worn.
3. Speech produced with KN95 masks will be overall less intelligible and more effortful to understand than speech without a mask. This difference may or may not change when speakers speak clearly and/or loudly.
4. Acoustic measures found to be sensitive to the effects of masks and speech styles in younger talkers (Knowles & Badh, 2022) and older talkers with and without PD (Knowles & Badh, 2023) will be associated with changes in speech intelligibility. Specifically, we predict that higher speech intelligibility will be associated with increases in speech intensity, energy in the 1- to 3-kHz range, and spectral tilt.

Methods

Speech corpus

All study procedures were approved by the institutional review board at the University at Buffalo. Speaker participants and stimuli represent a subset of those reported in Knowles and Badh (2023) and are reported in Table 1. Speaker participants comprised two groups: one group of 10 people with PD (six male, four female; mean age = 68 years), and one group of 10 older, healthy adults with no history of speech, language, hearing, or neurological concerns (six male, four female; mean age = 65 years).

Each participant read aloud six phonetically balanced lists of 10 sentences selected from the first 18 lists of the Harvard Sentence Corpus (IEEE, 1969). One sentence list was used for each combination of two mask and three speech conditions. Mask conditions for the perceptual study included a no-mask condition and the use of a KN95 face mask[1], and speech conditions included habitual, clear, and loud speech. Participants always began with the habitual speech condition. The order of the clear and loud speech conditions was counterbalanced across talkers, and the order of the face masks was randomized within each speech condition for each participant. Instructions for the clear speech condition were "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," and instructions for loud speech were "Speak at a volume that feels two times louder than your normal speaking voice." Participants had the

[1] The original study also included a surgical mask which was excluded in the perceptual study. Previous findings have demonstrated that surgical masks impose relatively minimal challenges to speech intelligibility (Toscano & Toscano, 2021) and were found to be less acoustically detrimental to the speech signal compared to the KN95 mask (Knowles & Badh, 2023). For simplicity, only the no-mask and-KN95 mask conditions were included here.

opportunity to practice each of the modified speech styles until they were ready to proceed with the experiment. They were not provided with feedback on their productions. Talkers were recorded at a 6-foot distance from a Shure SM58 microphone in order to mimic social distancing protocols. Before the experiment, a 1000-Hz tone of a fixed intensity was played in order to calibrate speech intensity off-line. A small loudspeaker was positioned under the speaker participant's chin and used to play the tone. This tone was played and recorded 3 times. Real-time intensity was recorded using a sound-level meter (Galaxy Audio CM-170) positioned adjacent to the microphone. The average intensity of this tone was used to adjust the intensity of the speech audio files for each speaker participant before including the audio in the listening experiment. Further details are described in the work of Knowles and Badh (2023).

Table 1. Participant demographics.

ID	Age	Gender	MoCA (/30)	Years since PD diagnosis	CPIB (/30)	Hypophonia severity	Deviant perceptual characteristics
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 consonsants
PD06	69	w	29	7	2	moderate	breathy/hoarse, strained
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
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		-	-	-
OC16	67	m	24		-	-	-
OC17	74	m	26		-	-	-

Note: Dashes indicate data not relevant for neurotypical controls. MoCA = Montreal Cognitive Assessment; CPIB = Communicative Participation Item Bank; PD = Parkinson's disease; M = men; W = women.

Auditory-perceptual experiment

Stimulus preparation

Audio files were then prepared for the listening portion of the experiment, which was hosted online via the Prolific crowdsourcing platform (https://www.prolific.co). Sentence boundaries were manually identified and automatically extracted in Praat (Boersma & Weenink, 2021). The actual utterance intensity was then adjusted using the calibration factor described above in order to reflect the true speech intensity at which it was uttered. Calibrated sentence audio files were mixed with +5 dB-signal-to-noise ratio (SNR) of multitalker background noise (a level chosen based on pilot testing) in order to limit ceiling and floor effects and mimic real-world speaking environments. In the original production experiment, speakers had the chance to practice each speech style with a novel set of sentences before the start of the new condition. This helped minimize speaker errors driven by the novelty of the speech style. Recorded sentences with major hesitations (such as long pauses or restarts), reading errors, or nonspeech disruptions were discarded before the perceptual experiment (n = 29). Each speaker contributed approximately 60 audio files: 10 sentences per condition combination across the two mask and three speech conditions. Speech stimuli were presented via a custom experiment written using jsPsych (de Leeuw, 2015) and hosted on Pavlovia (https://pavlovia.org/). Audio files were presented in a fully randomized order to each listener.

Listener participants

In total, 233 listener participants were recruited and compensated for their time. Listeners were excluded if they (a) reported a history of speech, language, or hearing concerns; (b) were not native North American English speakers; or (c) were not wearing headphones, all per self-report (n = 16). Additional listeners were excluded if they did not complete the experiment (n = 12) or if they fell below 2 SDs of the mean group accuracy in the perceptual outcomes (n = 13). A total of 192 listener participants were included in the final analyses, with an age range of 18–35 years.

Each listener heard randomized audio files from a single talker (n = ~60 stimulus items), and each talker was heard by a minimum of nine listeners (maximum: 11) after exclusions were made, which reflects the industry standard for remote data collection (Byun et al., 2015). Listeners additionally heard a random 20% of items repeated for intrarater reliability calculations. The experiment took approximately 20–30 min for listeners to complete.

Before the experiment, listeners were informed that they would be listening to speakers with and without neurological movement disorders in a variety of communication situations and were then presented with a practice period of five novel sentences from a novel control speaker (speaking habitually with no mask) in background noise. During this practice period, listeners were asked to adjust their volume to a comfortable level and were instructed to not change it during the experiment but were otherwise given no additional environmental instructions. After this practice period, listeners were presented with the stimuli sentences and asked to (a) "type exactly what [they] heard the speaker say[2]" and then (b) "rate how much effort it took to understand the speech" using a visual analog scale with anchors *very easy to understand* and *very difficult to understand*. Listeners were permitted to replay each audio recording one time, and there was no delay between the presentation of each sentence. Outcome measures included the proportion of keywords correctly transcribed and the effort rating expressed as a proportion along the visual scale. Transcriptions were assessed by the inclusion of five target words from each sentence, selected for their grammatical prominence in the sentence. Each sentence was given a score from 0% to 100% in increments of 20% based on how many of the five target words were included. Prevalent homophones and common homophonic typographical errors were considered correct.

Statistical analyses

Intrarater and interrater reliability for accuracy and effort were measured using intraclass correlation coefficients (ICCs; Koo & Li, 2016). Intrarater reliability was assessed using a two-way mixed-model (ICC 3, k), and interrater reliability was assessed using a two-way random model (ICC 2, k).

To evaluate the relationship between the two perceptual measures, repeated-measures correlation coefficients were calculated between listener accuracy and listener effort within each of the mask conditions using the `rmcorr` package (Bakdash & Marusich, 2022). For this, the data were averaged across listeners, resulting in one data point per condition per speaker.

We built two linear mixed-effects models, one for each outcome, in order to quantify the effects of face masks and speech styles in accordance with our research questions. These models were built in the following way. Listener accuracy and effort were each logit-transformed and modeled as a function of speaker group, face mask, and speech style using linear mixed-effects regression with the *lme4* package (Bates et al., 2015) in R (R Core Team, 2022). Post hoc pairwise comparisons were calculated using estimated marginal means (Lenth, 2020), and p values were adjusted using the Tukey method.

Two linear mixed-effects models were constructed, one for each outcome measure. Both included identical fixed effects structures. Fixed effects included speaker group, mask condition, speech condition, and all possible interactions. Random intercepts were included for listener participants, speaker participants, and sentence identifiers. Random by-speaker slopes for both speech condition contrast levels were also included. Random slope terms for listener participants led to model nonconvergence for intelligibility and were excluded but were included in the effort model.

Group and mask fixed effects terms were sum coded (OC = +1, PD = −1; NM = +1, KN95 = −1). Speech condition was coded using reverse Helmert contrasts, resulting in two contrast

---

[2] In addition to this instruction, further guidance included "Type exactly what you think you heard, even if you are unsure. If you have no idea, you can type NA."

levels (habitual vs. clear/loud: habitual = +2/3, clear = −1/3, loud = −1/3; clear vs. loud: clear = +1/2, loud = −1/2). The first contrast level permits the comparison of habitual speech to the average of clear and loud speech outcomes combined. The second contrast level compares clear to loud speech. In all cases, contrasts are coded such that each baseline (denoted on the left) is assigned a positive value. A positive model estimate, thus, indicates the baseline for that contrast resulted in a higher outcome than the comparison. For example, a positive estimate for OC versus PD in the intelligibility model would indicate higher intelligibility ratings for OC. A negative estimate for the habitual versus clear/loud contrast would indicate lower intelligibility outcomes for habitual speech compared to clear and loud speech.

As a secondary analysis to address Research Question 4, we investigated the additive effect of three acoustic measures of spectral balance on speech intelligibility. This analysis was driven by the results of Knowles and Badh (2023), which quantified the effects of face masks and effortful speech styles on acoustic outcomes in this sample. Details of the acoustic analysis are provided in Knowles and Badh (2023) but are briefly summarized here. These included mean speech intensity, the mean energy in the 1- to 3-kHz spectral range, and the difference in energy between 0–1 and 1–10 kHz (i.e., a measure of spectral tilt). In order to quantify their potential contributions to speech intelligibility, we added each of these measures as fixed effects to the intelligibility models one at a time. For the addition of each acoustic measure, we performed likelihood ratio tests to determine whether any of the three acoustic measures significantly improved the model fit compared to the original model (i.e., without the acoustic variable), using a more conservative threshold of $p < .01$. Acoustic measures that improved model fit were then iteratively added to a single model. Collinearity was assessed by examining variance inflation factors to ensure correlations between predictors remained acceptably low (i.e., < 5; Akinwande et al., 2015). For simplicity's sake and given the exploratory nature of this approach, we limited this secondary analysis to the intelligibility outcomes only (i.e., not listener effort).

Results

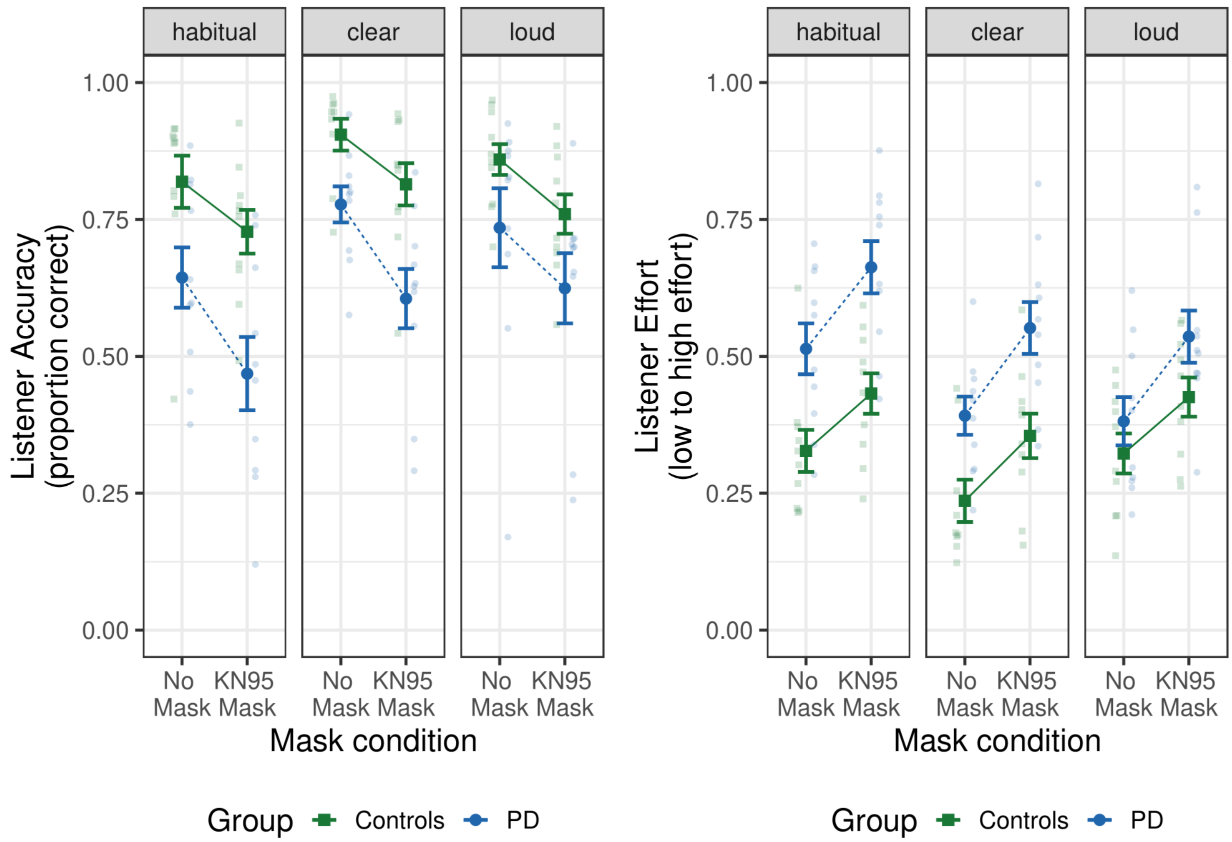

Figure 1. *Observed group means for each mask and speech style condition. Data have been averaged over listeners. Error bars represent standard error. Points represent individual data points, averaged over listeners. PD = Parkinson's disease.*

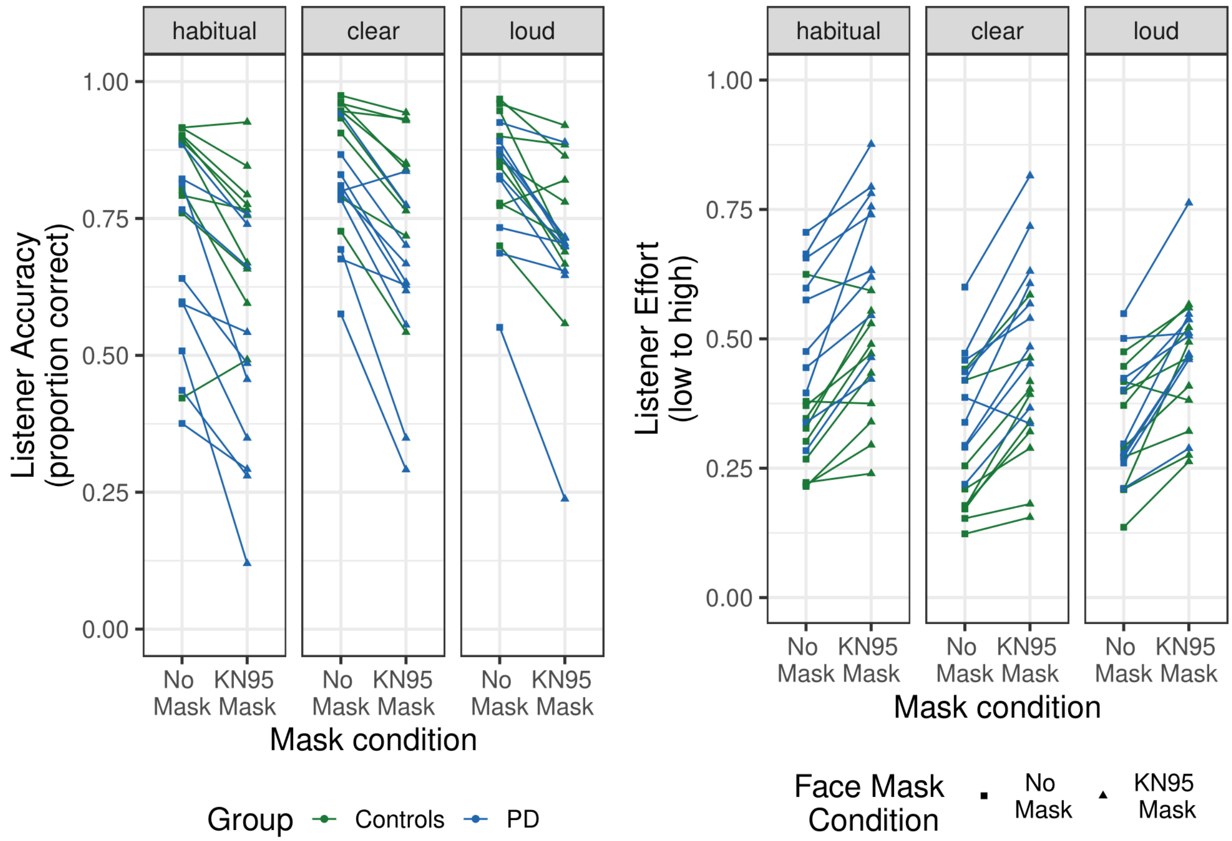

Figure 2. *Observed individual speakers means in each mask and speech style condition. Each line represents a given speaker. Means have been averaged over listeners and sentence lists. PD = Parkinson's disease.*

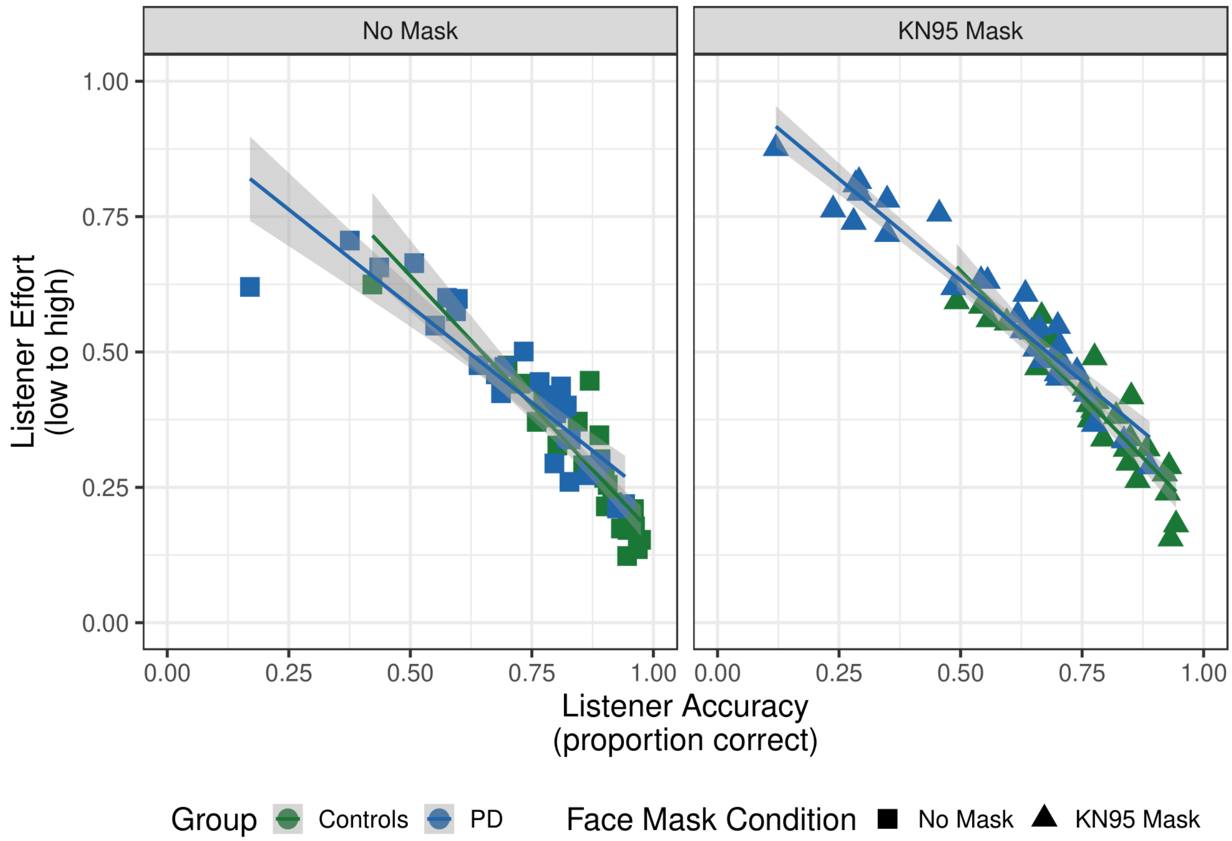

Figure 3. *Relationship between listener accuracy and listener effort for each mask condition, collapsed across speaker group and speech condition. PD = Parkinson's disease.*

Table 2. Model results for speech intelligibility and listener effort.

	Listener Accuracy			**Listener Effort**		
Predictors	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*
Intercept	1.53	1.07 – 1.99	**<0.001**	-0.42	-0.76 – -0.08	**0.016**
Group: OC vs. PD	0.62	0.17 – 1.06	**0.009**	-0.48	-0.80 – -0.15	**0.006**
Mask: NM vs. KN	0.38	0.34 – 0.41	**<0.001**	-0.36	-0.39 – -0.32	**<0.001**
Speech: H vs. C&L	-0.64	-0.92 – -0.36	**<0.001**	0.50	0.29 – 0.70	**<0.001**
Speech: C vs. L	0.20	-0.09 – 0.49	0.167	-0.20	-0.34 – -0.06	**0.008**
[Group: OC vs. PD] x [Mask: NM vs. KN]	-0.10	-0.13 – -0.07	**<0.001**	0.08	0.05 – 0.10	**<0.001**
[Group: OC vs. PD] x [Speech: H vs. C&L]	0.26	-0.01 – 0.54	0.060	-0.28	-0.48 – -0.08	**0.009**
[Group: OC vs. PD] x [Speech: C vs. L]	0.07	-0.22 – 0.36	0.611	-0.23	-0.37 – -0.09	**0.003**
[Mask: NM vs. KN] x [Speech: H vs. C&L]	-0.12	-0.20 – -0.04	**0.004**	0.13	0.06 – 0.19	**<0.001**
[Mask: NM vs. KN] x [Speech: C vs. L]	0.24	0.14 – 0.33	**<0.001**	-0.13	-0.22 – -0.05	**0.001**
[Group: OC vs. PD] x [Mask: NM vs. KN] x [Speech: H vs. C&L]	-0.08	-0.15 – -0.01	**0.031**	0.02	-0.04 – 0.08	0.532
[Group: OC vs. PD] x [Mask: NM vs. KN] x [Speech: C vs. L]	-0.15	-0.23 – -0.07	**<0.001**	-0.01	-0.08 – 0.06	0.832

Random Effects

σ^2	2.77	1.99
τ_{00}	0.52 $_{\text{participant}}$	0.56 $_{\text{participant}}$
	0.60 $_{\text{id}}$	0.44 $_{\text{id}}$
	0.84 $_{\text{speaker}}$	0.42 $_{\text{speaker}}$
τ_{11}	0.32 $_{\text{speaker.HvCL}}$	0.10 $_{\text{participant.HvCL}}$
	0.34 $_{\text{speaker.CvL}}$	0.02 $_{\text{participant.CvL}}$
		0.15 $_{\text{speaker.HvCL}}$
		0.06 $_{\text{speaker.CvL}}$
ρ_{01}	0.15 $_{\text{speaker.HvCL}}$	-0.26 $_{\text{participant.HvCL}}$
	-0.60 $_{\text{speaker.CvL}}$	0.31 $_{\text{participant.CvL}}$
		0.00 $_{\text{speaker.HvCL}}$
		-0.15 $_{\text{speaker.CvL}}$
ICC	0.41	0.41
N	194 $_{\text{participant}}$	194 $_{\text{participant}}$
	20 $_{\text{speaker}}$	20 $_{\text{speaker}}$
	179 $_{\text{id}}$	179 $_{\text{id}}$
Observations	11259	11259
Marginal R^2 / Conditional R^2	0.123 / 0.486	0.117 / 0.483

Note: Outcome variables have been logit-transformed. Bolded *p* values reflect *p* < .05. CI = confidence interval; OC = older controls; PD = Parkinson's disease; NM = no mask; KN = KN95; H = habitual; C = clear; L = loud; C&L = clear and loud; ICC = intraclass correlation coefficient.

Results are reported in Table 2 and Figures 1–3, with Figure 2 being included as supplementary to Figure 1 by showing individual variation. Across all listeners, intrarater reliability indicated *moderate* agreement for both accuracy (mean ICC = .76) and effort (mean ICC = .716; Koo & Li, 2016). Across all speaker playlists, interrater reliability was found to be good for both outcomes (mean ICC, accuracy: .866; mean ICC, effort: .857).

While we predicted a similar pattern of effects across listener accuracy and effort, we first sought to establish their relationship to contextualize any differences. Listener accuracy and effort were found to be strongly correlated with each other across all conditions (repeated-measures correlation coefficient r_{rm}: –.917, CI [–0.943, –0.879]; $p < .001$), as illustrated by Figure 3. This relationship was stronger within the KN95 mask condition (repeated-measures correlation coefficient r_{rm}: –.944, CI [–0.970, –0.897]; $p < .001$) compared to within the no-mask condition (repeated-measures correlation coefficient $r = -.819$, CI [–0.901, –0.681]; $p < .001$).

Main effects of group, mask, and speech style

Overall, listeners were less accurate and reported greater listening effort for the talkers with PD compared to controls (Prediction 1: main effect of group for accuracy: $\hat{\beta} = 0.617$, $p = .0091$; effort: $\hat{\beta} = -0.476$, $p = .0061$), for habitual compared to effortful speech (Prediction 2: main effect of speech style; accuracy: $\hat{\beta} = -0.639$, $p < .001$; effort: $\hat{\beta} = 0.499$, $p < .001$), and when talkers spoke with a mask on compared to without (Prediction 3: main effect of mask for accuracy: $\hat{\beta} = 0.377$, $p < .001$; effort: $\hat{\beta} = -0.356$, $p < .001$). Between the two effortful styles, clear speech was less effortful than loud speech for listeners to understand ($\hat{\beta} = -0.201$, $p = .0081$), but there was no statistically significant difference in accuracy between the two altered speech styles ($\hat{\beta} = 0.2$, $p = .1671$).

Interactions with speaker group and mask

Each of the first three predictions sought to identify not only overall effects of group, speech styles, and masks, but also how these effects are mediated across conditions. Regarding Prediction 1, talkers with PD were even more disadvantaged than controls when wearing masks, which was found as two-way interactions between speaker group and mask for both accuracy and effort. That is, listeners were even less accurate and reported greater effort when listening to talkers with PD in masks (accuracy: $\hat{\beta} = -0.099$, $p < .001$; effort: $\hat{\beta} = 0.076$, $p < .001$). Furthermore, talkers with PD did not gain as much in clear speech (compared to loud speech) when wearing masks, a finding captured by the three-way interaction for accuracy between group, mask, and speech style for habitual versus clear and loud speech ($\hat{\beta} = -0.08$, $p = .0311$) and Clear versus Loud speech ($\hat{\beta} = -0.152$, $p < .001$). No three-way interactions were found for listener effort, however, indicating that this asymmetry in the ability to understand talkers with PD may not have been perceived as more difficult, even if listeners were less successful overall. In other words, the benefit of effortful speech, especially for speakers with PD, was greater without a mask (Predictions 2 and 3), but this was a nuance captured by objective intelligibility and not by subjective listener effort.

To determine whether the KN95 mask led to consistently poorer perceptual outcomes, post-hoc pairwise comparisons tested the difference in listener accuracy between the mask conditions for each group and speech style. In all cases, the KN95 masks were associated with significantly lower speech intelligibility and higher listener effort across the board (p < .001 for most comparisons). Greater declines from the KN95 mask were observed for the PD group compared to controls, as evidenced by higher odds ratios. These differences can be observed in Figure 1 and Table 3. Due to the strong relationship between listener accuracy and effort and in order to limit the number of post hoc tests, listener effort was not included in these pairwise comparisons.

In order to determine whether the intelligibility advantages of using effortful speech persisted when speakers wore masks, a second set of post hoc comparisons was run, reported in Table 4. These comparisons tested the difference between each speech style for each speaker group and mask condition. P values were adjusted using the Bonferroni method for three tests. In general, both groups were most intelligible when using clear speech without a mask (PD: 78%; OC: 90%) and least intelligible when using habitual speech with a KN95 mask (PD: 47%; OC: 73%). For the control group, the only statistically significant comparison was clear versus habitual speech without a mask (odds ratio = 0.472, p = .001). Continued in Table 4, no other comparisons reached significance for the controls, suggesting limited differences in listener accuracy regardless of speech style with or without a mask for control speakers. For the PD group, however, nearly all comparisons were significantly different, demonstrating that even when wearing a face mask, clear and loud speech styles resulted in higher listener accuracy than with a habitual speech style. Without a mask, clear speech also offered an advantage over loud speech (odds ratio = 1.677, p = .039) but with a mask, this difference disappeared (odds ratio = 0.771, p = .61).

These findings suggest that although wearing a face mask reduces the effectiveness of effortful speech styles for people with PD, using clear or loud speech with a mask on still results in better speech intelligibility than not making any adjustments at all. These patterns are visible in Figure 1, in which talkers with PD show a steeper decline in intelligibility as a result of masks when using habitual and clear speaking styles. This pattern is visibly similar but not as extreme for listener effort.

This finding was also supported by a two-way interaction between mask and speech style, pictured in Figure 1, which demonstrated that the clear speech benefit was attenuated by masks for both accuracy (habitual vs. clear/loud: $\hat{\beta}$ = –0.118, p = .0041; clear vs. loud: $\hat{\beta}$ = –0.118, p = .0041) and effort (habitual vs. clear/loud: $\hat{\beta}$ = 0.126, p < .001; clear vs. loud: $\hat{\beta}$ = –0.135, p = .0011). In summary, clear speech was most easily and more accurately understood without a mask, but loud and clear were similarly understood when speakers wore a mask.

Table 3. Pairwise comparisons for differences in listener accuracy between no mask and KN95 masks for each speaker group and speech condition.

Group	Contrast	Speech style	Odds Ratio	SE	z.ratio	p.value
OC	No Mask / KN95	habitual	1.339	0.128	3.056	**0.002**
	No Mask / KN95	clear	2.166	0.200	8.391	**0.000**
	No Mask / KN95	loud	1.827	0.153	7.179	**0.000**
PD	No Mask / KN95	habitual	2.466	0.220	10.102	**0.000**
	No Mask / KN95	clear	3.921	0.341	15.690	**0.000**
	No Mask / KN95	loud	1.803	0.162	6.554	**0.000**

Note: Bolded p values reflect $p < .05$. SE = standard error; OC = older controls; PD = Parkinson's disease.

Table 4. Pairwise comparisons for differences in listener accuracy between speech styles for each speaker group and face mask condition.

Group	Contrast	Mask	Odds Ratio	SE	z.ratio	p.value
OC	habitual / clear	No Mask	0.472	0.101	-3.497	**0.001**
	habitual / loud	No Mask	0.674	0.152	-1.750	0.24
	clear / loud	No Mask	1.428	0.294	1.732	0.25
	habitual / clear	KN95	0.763	0.162	-1.271	0.611
	habitual / loud	KN95	0.919	0.208	-0.372	1
	clear / loud	KN95	1.205	0.248	0.902	1
PD	habitual / clear	No Mask	0.302	0.065	-5.542	**<.001**
	habitual / loud	No Mask	0.506	0.113	-3.044	**0.007**
	clear / loud	No Mask	1.677	0.349	2.484	**0.039**
	habitual / clear	KN95	0.480	0.102	-3.457	**0.002**
	habitual / loud	KN95	0.370	0.085	-4.352	**<.001**
	clear / loud	KN95	0.771	0.158	-1.272	0.61

Note: Bolded *p* values reflect $p < .05$. *SE* = standard error; OC = older controls; PD = Parkinson's disease.

Potential acoustic contributors to intelligibility in masks

In our previous article involving these same speaker participants, we investigated the combined effects of face masks and effortful speech from this same data set on three acoustic outcomes of spectral balance: mean speech intensity, the mean energy in the 1- to 3-kHz spectral range, and the difference in energy between 0–1 and 1–10 kHz (Knowles & Badh, 2023). Each of these was added to the intelligibility model in the current study one at a time and evaluated using likelihood ratio tests. Two of the three measures were found to improve model fit compared to the original model at $p < .01$, with spectral tilt yielding the lowest Bayesian information criterion, followed by intensity (tilt: χ^2 = 35.955, p = < .001; intensity: χ^2 = 15.439, p = < .001). Midrange frequency energy was not found to improve model fit at $p < .01$ (χ^2 = 5.166, p = .0231).

The model including spectral tilt was then rerun with intensity as a predictor. Intensity again was found to improve model fit (χ^2 = 17.978, p = < .001). Variance inflation factors for all predictors were below 2, suggesting acceptably low collinearity.

The final model suggested that more positive spectral tilt was associated with higher speech intelligibility ($\hat{\beta}$ = 0.068, p = < .001), while with all other predictors controlled for, overall higher speech intensity was associated with *lower* speech intelligibility ($\hat{\beta}$ = –0.055, p = < .001). These results are illustrated in Figure 4. Perceptual results are interpreted in the context of these acoustic findings in the Discussion section.

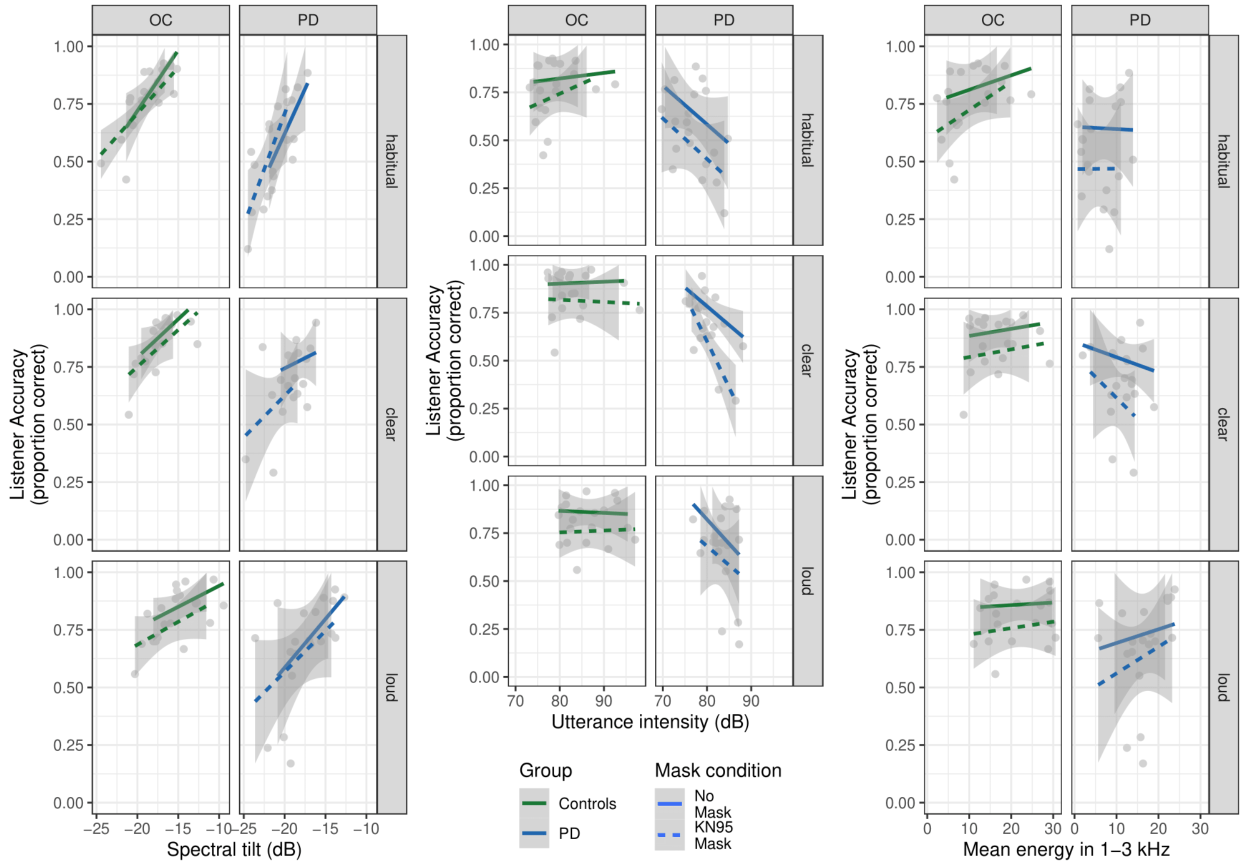

Figure 4. *Relationship between listener accuracy and each of the three acoustic measures from Knowles and Badh (2023): spectral tilt, speech intensity, and mean energy in the 1- to 3-kHz range. PD = Parkinson's disease; OC = older controls.*

Discussion

This study provides support for a growing body of evidence that clear and loud speaking styles are useful in accommodating for the low-pass filtering effect of face masks. Consistent with previous literature, speakers in this study were less intelligible when wearing face masks (Cohn et al., 2021; Magee et al., 2020; Toscano & Toscano, 2021). The use of effortful speech styles enhanced speech intelligibility compared to habitual speech, partially compensating for the effect of the masks, as has been previously demonstrated (Gutz et al., 2021, 2022; Smiljanic et al., 2021). This study demonstrated that this was true for individuals with PD as well as age- and gender-matched controls. An important secondary finding was that the clear speech advantage over loud speech was not present when speakers wore masks. That is, without masks, listeners were most accurate in understanding clear speech, followed by loud speech, and this difference was found to be statistically significant. *With* face masks, though, intelligibility of clear and loud did not significantly differ, although both were more intelligible than habitual speech. Functionally, this suggests that while effortful speech styles are effective in, at least in part, compensating for the intelligibility loss from face masks, the effectiveness of clear speech is reduced with a mask on. The acoustic correlates and clinical implications of this are discussed below. Here, we summarize our main findings in the context of our original predictions:

1. People with PD were overall less intelligible than controls. This difference was greater when speakers wore a KN95 face mask, as captured by the Group × Mask interaction.
2. Habitual speech was less intelligible and more effortful to understand compared to clear and loud speech (main effect of speech style). For the speakers with PD, clear speech was associated with greater gains than loud speech without a mask. This advantage of clear over loud speech for intelligibility (but not listener effort) was lost when the KN95 mask was worn.
3. The KN95 mask was associated with overall poorer perceptual outcomes (main effect of mask). This relationship was preserved across speech styles, although the magnitude of the effect lessened with effortful speech (as captured by the two- and three-way interactions with mask).
4. Higher speech intelligibility was found to be associated with higher spectral tilt (as predicted) and lower speech intensity (the opposite of our prediction), and was not associated with midrange frequency energy (inconsistent with our prediction).

These findings are, for the most part, consistent with our previous work, which demonstrated that the acoustic attenuation of KN95 masks is parallel across habitual, clear, and loud speech styles for these same speakers (Knowles & Badh, 2023) as well as with previous literature (Gutz et al., 2021, 2022; Knowles & Badh, 2022). In our previous article of talkers with and without PD, we demonstrated that loud followed by clear speech provided a greater increase in speech intensity, spectral tilt, and midrange frequency energy compared to habitual speech, and this finding held with and without the presence of face masks (Knowles & Badh, 2023). While that study found a hierarchy of loud > clear > habitual speech with regard to increases in spectral balance measures, this study found a hierarchy of clear > loud > habitual speech for intelligibility without a mask and [clear = loud] > habitual with a mask.

The acoustic underpinnings of intelligibility in PD may be altered when the talker wears a mask, given the way the mask alters the properties of the speech signal and therefore the information available to the listener. Previous research on speech intelligibility in PD suggests that, for many speakers, clear, followed by loud speech styles yield higher speech intelligibility (Tjaden et al., 2014). The results of this study support this finding without a mask, but this pattern disappeared with a mask. This discrepancy could be related to the acoustic differences in clear and loud speech and the acoustic damping imposed by the mask. Clear speech is typically characterized by slower, louder, hyperarticulated speech that leads to changes such as increased vowel and consonant distinctiveness (e.g., as reviewed in Smiljanic et al., 2021). Acoustically, this results in, among other features, increased spectral energy in canonical vowel formant frequency bands and strengthening of higher frequency information responsible for distinguishing certain consonants. Previous literature on face coverings indicate this increased acoustic salience in these critical frequency regions is likely damped by the attenuation profiles of masks (e.g., Corey et al., 2020; Maryn et al., 2021). Therefore, it is plausible that the clear speech benefit driven by increased energy in this range fails to be maintained when a mask is worn. This would be consistent with findings from Neel (2009), who found that while one third to half of improvements in intelligibility in loud speech could be attributed to increases in SNR, changes in spectral tilt or fundamental frequency were likely also contributors. Indeed, we found that increased spectral tilt was a predictor of speech intelligibility in our current study, but mid-range frequency energy (1–3 kHz) was not. Despite the fact that both metrics were shown to increase as predicted in clear and loud speech in Knowles and Badh (2023), the fact that midrange frequency energy did not contribute to intelligibility in this dataset suggests that the damping of the masks in this range may attenuate the clear speech gains that would normally be observed. The nuances of this are further discussed below.

As described earlier, we entered the three primary acoustic metrics from Knowles and Badh (2023) into our present model of speech intelligibility and found that, consistent with our expectations, flatter spectral tilt was associated with improvements in speech intelligibility. That is, a higher relative concentration of high- to low-frequency energy (0–1 vs. 1–10 kHz) across the speech spectrum was associated with speech that was more accurately understood by listeners. The relative size of this effect, however, compared to the effect size of masks and effortful speech styles, was quite small ($\hat{\beta}$ = 0.068) compared to, for example, the main effect of group ($\hat{\beta}$ = 0.615), mask ($\hat{\beta}$ = 0.316), or effortful speech (habitual vs. clear and loud: $\hat{\beta}$ = –0.757). Furthermore, the overall explanatory power of the model (conditional r^2) was low (48.6%), with the fixed effects only accounting for 12.3% of the change in intelligibility (marginal r^2). This suggests additional factors not presently included may have been stronger predictors of speech intelligibility and likely also of listener effort.

The outcome of the acoustic additions to the model yielded two *unexpected* findings as well. First, given previous literature on acoustic predictors of the clear speech benefit (Gilbert et al., 2014; Hazan & Baker, 2011; Hazan et al., 2018; Krause & Braida, 2004, 2009; Smiljanic, 2021), we expected the inclusion of acoustic energy in the 1- to 3-kHz range to improve our model's predictive power. It did not; only spectral tilt and speech intensity improved the model fit. It is possible that attenuation in this same range imposed by the KN95 mask reduces the effectiveness of this component of the signal and that higher frequency information (which here was captured by spectral tilt) is more important to the listener in recuperating the overall speech signal. While we did not look at interactions with these acoustic measures, it is evident from Figure 4 that there was considerable variability in the relationship between speech intelligibility and mean energy between 1 and 3 kHz across speaker groups and speech conditions, underscoring a need for more research in this area. Gutz et al. (2022) found that KN95 masks provided the greatest amount of attenuation to the acoustic signal above 2.5 and 4 kHz. A measure of spectral tilt that captured a wider low-frequency band contribution, such as 0–4 kHz instead of the 0–1 kHz as used here, may be more sensitive to changes in intelligibility due to masks.

Another surprising finding was that while including speech intensity improved the model's predictive power, the direction of this effect was in the opposite direction as expected. Higher speech intensity was actually associated with *poorer* speech intelligibility when spectral tilt was controlled, although again the relative size of this effect was small ($\hat{\beta}$ = –0.055). That is, while better spectral tilt (i.e., balance), which is often a consequence of louder, more effortful speech (Fant, 1960; Ternström et al., 2006), benefited speakers' ability to be understood, simply increasing their speech intensity actually made understanding more difficult. Trends visible in Figure 4 suggest that this pattern for intensity was most evident for talkers with PD rather than controls; in fact, controls showed little to no relationship between intensity and intelligibility. On the other hand, the trend for intelligibility to increase with spectral tilt was consistent across both speaker groups.

Further research is required to understand the nuanced relationships between speech intensity, measures of spectral balance, and the filtering effects of the barriers imposed by face masks and how these impact a listener's ability to retrieve the signal. It is possible that when a person increases their speech volume, additional distortions that affect prosody, voice quality, or articulation are introduced. While these may not pose substantial difficulty when a speaker is maskless, these distortions could be compounded by the high-frequency attenuation of a mask, leading to greater difficulty for the listener. For example, Neel (2009) found that while increased intensity *was* associated with increased intelligibility in talkers with PD in loud speech, other phonatory and articulatory changes also contributed.

In this study, an asymmetry was observed between speech intelligibility and listener effort for the three-way interaction involving group, mask, and speech style. Namely, the negative impact of the masks on the clear versus loud speech benefit impacted intelligibility to a greater extent than listener effort. This suggests that listeners were sometimes less accurate when transcribing masked clear speech despite reporting similar levels of difficulty. This is consistent with Kim and Thompson (2022), who found that listeners rated clear speech spoken through a mask (for young, healthy talkers) as intelligible as habitual speech without a mask, despite a decline in transcription accuracy. The authors found that lower transcription scores were driven by confusion of consonants reliant on high-frequency acoustic information (such as fricatives). A future acoustic–phonetic analysis of perceptual errors in Parkinsonian speech with a mask would be warranted to understand the specific contributors responsible for these patterns. Collectively, this study's findings suggest that the relative contribution of acoustic factors in perceptual outcomes for talkers with PD may change when face masks are worn, although the exact nature of these contributions is unknown. Furthermore, how masks impact the relationship between perceptual outcomes is an open question.

It is worth noting that our findings of speech intelligibility and listener effort were in the context of background noise. While we did not test intelligibility in quiet, previous work has demonstrated that many face masks typically pose little, if any, burden to intelligibility in optimal, quiet listening conditions (Brown et al., 2021; Carraturo et al., 2021). Whether this is the case for speakers with speech disorders is not known.

Clinical implications

The results of this study suggest that speaking clearly or loudly can help overcome intelligibility challenges imposed by face masks in people with and without PD. Clinically, this is relevant as masks continue to be required in many healthcare settings, and our results also demonstrate that speech in people with PD is disproportionately affected with a mask on. For talkers with PD, speaking effortfully with a mask was better than using habitual speech but brought their intelligibility down to similar levels to their habitual, unmasked speech. That is, the face masks effectively eliminated the intelligibility benefits they derived from clear and loud speech strategies. What this implies from a clinical perspective is that behavioral speech strategies may be ineffective alone in overcoming challenges posed by masks, especially for people with lower baseline intelligibility. These individuals will likely require additional measures in place to maximize their ability to communicate in these settings. Examples of this might include environmental modifications, such as taking efforts to reduce background noise, listener modifications, ensuring the listener is attending to the person speaking, or augmentative options, such as exploring the use of a speech amplification device.

Limitations

Trying to emulate naturalistic speech in a controlled environment carries trade-offs. Most notably, while the Harvard sentence corpus is phonetically balanced, it is not naturalistic to read sentences off of a paper. In addition, this study does not control for compensatory behavior that accompanies wearing face masks. This study also only considers one type of mask on account of economy, and only an adverse listening environment is considered. Finally, while crowdsourcing offers many participants, it also comes with the potential for things like bots and inattentive participants, although it is our hope that filtering out participants based on reliability and other factors combated that possibility.

Conclusion

In the end, we found that listeners were less accurate and reported greater listening effort for the PD group and for the mask condition. However, face masks were associated with a steeper decline in speech intelligibility and an increase in listener effort for talkers with PD. In addition, listeners were more accurate and reported less effort when listening to clear and loud speech compared to habitual speech, and both clear and loud speech styles were associated with improvements in perceptual outcomes both with and without masks.

Data Availability Statement

Deidentified data and reproducible code for the current study are available at https://github.com/thealk/cline-et-al_2023_ajslp_masks-intelligibility.

Acknowledgments

Acknowledgments: The authors thank the participants for their time and contributions to this work.

References

Adams, S. G., & Dykstra, A. D. (2009). Hypokinetic dysarthria. In M. R. McNeil (Ed.), *Clinical management of sensorimotor speech disorders*. Thieme Publishing Group.

Adams, S., Haralabous, O., Dykstra, A., Abrams, K., & Jog, M. (2005). Effects of multi-talker background noise on the intensity of spoken sentences in Parkinson's disease. *Canadian Acoustics*, *33*(3), 94–95.

Akinwande, M. O., Dikko, H. G., & Samson, A. (2015). Variance Inflation Factor: As a Condition for the Inclusion of Suppressor Variable(s) in Regression Analysis. *Open Journal of Statistics*, *5*(7), 754–767. https://doi.org/10.4236/ojs.2015.57075

Bakdash, J. Z., & Marusich, L. R. (2022). rmcorr: Repeated Measures Correlation. https://CRAN.R-project.org/package=rmcorr

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

Boersma, P., & Weenink, D. (2021). *Praat: Doing phonetics by computer [Computer program, Version 6.2]*.

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*(1), 49. https://doi.org/10.1186/s41235-021-00314-0

Byun, T. M., Halpin, P. F., & Szeredi, D. (2015). Online crowdsourcing for efficient rating of speech: A validation study. *Journal of Communication Disorders*, *53*, 70–83.

Carraturo, S., Brown, V. A., Van Engen, K. J., & Peelle, J. (2021). Face masks, speech intelligibility, and listening effort. *The Journal of the Acoustical Society of America*, *150*(4), A275–A275. https://doi.org/gpg7v5

CDC. (2022). Masks and Respirators. In *Centers for Disease Control and Prevention*. https://www.cdc.gov/coronavirus/2019-ncov/prevent-getting-sick/types-of-masks.html.

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* [PhD thesis]. Western University.

de Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a Web browser. *Behavior Research Methods*, *47*(1), 1–12. https://doi.org/10.3758/s13428-014-0458-y

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.

Forrest, K., Weismer, G., & Turner, G. S. (1989). Kinematic, acoustic, and perceptual analyses of connected speech produced by parkinsonian and normal geriatric adults. *The Journal of the Acoustical Society of America*, *85*(6), 2608–2622.

Fox, C. M., & Ramig, L. O. (1997). Vocal sound pressure level and self-perception of speech and voice in men and women with idiopathic Parkinson's disease. *American Journal of Speech-Language Pathology*, *6*(2), 85–94. https://doi.org/10.1044/1058-0360.0602.85

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

Gutz, S. E., Rowe, H. P., Tilton-Bolowsky, V. E., & Green, J. R. (2022). Speaking with a KN95 face mask: A within-subjects study on speaker adaptation and strategies to improve intelligibility. *Cognitive Research: Principles and Implications*, *7*(1), 73. https://doi.org/10.1186/s41235-022-00423-4

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., 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

Ho, A. K., Bradshaw, J. L., Iansek, R., & Alfredson, R. (1999). Speech volume regulation in Parkinson's disease: Effects of implicit cues and explicit instructions. *Neuropsychologia*, *37*, 1453–1460.

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

Kim, Y., & Thompson, A. (2022). An Acoustic-Phonetic Approach to Effects of Face Masks on Speech Intelligibility. *Journal of Speech, Language, and Hearing Research, 65*(12), 4679-4689. 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.*, *151*(5), 3359–3368. https://doi.org/10.1121/10.0011400

Knowles, T., & Badh, G. (2023). Impact of face masks on speech in Parkinson's disease: Effect of clear and loud speech styles. *Journal of Speech, Language, and Hearing Research*, 1-24. https://doi.org/10.1044/2022_JSLHR-22-00291

Koo, T. K., & Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine*, *15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

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

Lam, J., Tjaden, K., & Wilding, G. (2012). Acoustics of clear speech: Effect of instruction. *Journal of Speech, Language, and Hearing Research*, *55*(6), 1807–1821. https://doi.org/10.1044/1092-4388(2012/11-0154)

Lenth, R. (2020). *Emmeans: Estimated marginal means, aka least-squares means* [Manual].

Logemann, J. A., & Fisher, H. B. (1981). Vocal tract control in Parkinson's disease. *Journal of Speech and Hearing Disorders*, *46*(4), 348–352. https://doi.org/10.1044/jshd.4604.348

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

Logemann, J., Boshes, B., Fisher, H., & Siegfried, J. (1973). The steps in the degeneration of speech and voice control in Parkinson's disease. *Parkinson's Diseases: Rigidity, Akinesia, Behavior*, 101–112.

Ludlow, C. L., & Bassich, C. J. (1984). Relationships between perceptual ratings and acoustic measures of hypokinetic speech. In McNcil, M. R., J. C. Rosenbek, & A. E. Aronson (Eds.), *The dysarthrias: Physiology, acoustics, perception, management* (pp. 163–195). College-Hill Press.

Magee, M., Lewis, C., Noffs, G., Reece, H., Chan, J. C. S., Zaga, C. J., Paynter, C., Birchall, O., Rojas Azocar, S., Ediriweera, A., Kenyon, K., Caverlé, M. W., Schultz, B. G., & Vogel, A. P. (2020). Effects of face masks on acoustic analysis and speech perception: Implications for peri-pandemic protocols. *The Journal of the Acoustical Society of America*, *148*(6), 3562–3568. https://doi.org/10.1121/10.0002873

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

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

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

R Core Team. (2022). *R: A language and environment for statistical computing* [Manual]. R Foundation for Statistical Computing.

Rahne, T., Fröhlich, L., Plontke, S., & Wagner, L. (2021). Influence of surgical and N95 face masks on speech perception and listening effort in noise. *PLOS ONE*, *16*(7), e0253874. https://doi.org/10.1371/journal.pone.0253874

Rosenthal, A. L., Lowell, S. Y., & Colton, R. H. (2014). Aerodynamic and Acoustic Features of Vocal Effort. *Journal of Voice*, *28*(2), 144–153. https://doi.org/10.1016/j.jvoice.2013.09.007

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

Stipancic, K. L., van Brenk, F., Kain, A., Wilding, G., & Tjaden, K. (2022). Clear Speech Variants: An Investigation of Intelligibility and Speaker Effort in Speakers With Parkinson's Disease. *American Journal of Speech-Language Pathology, 31*(6), 2789-2805.

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

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., & 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

Toscano, J. C., & Toscano, C. M. (2021). Effects of face masks on speech recognition in multi-talker babble noise. *PloS One*, *16*(2), e0246842. https://doi.org/10.1371/journal.pone.0246842

Weismer, G. (1984). Articulatory characteristics of Parkinsonian dysarthria: Segmental and phrase-level timing, spirantization, and glottal-supraglottal coordination. In M. McNeil (Ed.), *The Dysarthrias: Physiology, Acoustics, Perception, Management* (pp. 101–130). College Hill Press San Diego.

Yi, H., Pingsterhaus, A., & Song, W. (2021). Effects of Wearing Face Masks While Using Different Speaking Styles in Noise on Speech Intelligibility During the COVID-19 Pandemic. *Frontiers in Psychology*, *12*, 682677. https://doi.org/10.3389/fpsyg.2021.682677