ASHA American Speech-Language-Hearing Association

JSLHR

Research Article

Cognitive Predictors of Perception and Adaptation to Dysarthric Speech in Young Adult Listeners

Kaitlin L. Lansford,[a] Tyson S. Barrett,[b] and Stephanie A. Borrie[c]

[a] School of Communication Science & Disorders, Florida State University, Tallahassee [b] Department of Psychology, Utah State University, Logan
[c] Department of Communicative Disorders and Deaf Education, Utah State University, Logan

ARTICLE INFO

Article History:
Received July 1, 2022
Revision received August 9, 2022
Accepted September 2, 2022

Editor-in-Chief: Cara E. Stepp
Editor: Nancy Pearl Solomon

https://doi.org/10.1044/2022_JSLHR-22-00391

ABSTRACT

Purpose: Although recruitment of cognitive-linguistic resources to support dysarthric speech perception and adaptation is presumed by theoretical accounts of effortful listening and supported by cross-disciplinary empirical findings, prospective relationships have received limited attention in the disordered speech literature. This study aimed to examine the predictive relationships between cognitive-linguistic parameters and intelligibility outcomes associated with familiarization with dysarthric speech in young adult listeners.
Method: A cohort of 156 listener participants between the ages of 18 and 50 years completed a three-phase perceptual training protocol (pretest, training, and posttest) with one of three speakers with dysarthria. Additionally, listeners completed the National Institutes of Health Toolbox Cognition Battery to obtain measures of the following cognitive-linguistic constructs: working memory, inhibitory control of attention, cognitive flexibility, processing speed, and vocabulary knowledge.
Results: Elastic net regression models revealed that select cognitive-linguistic measures and their two-way interactions predicted both initial intelligibility and intelligibility improvement of dysarthric speech. While some consistency across models was shown, unique constellations of select cognitive factors and their interactions predicted initial intelligibility and intelligibility improvement of the three different speakers with dysarthria.
Conclusions: Current findings extend empirical support for theoretical models of speech perception in adverse listening conditions to dysarthric speech signals. Although predictive relationships were complex, vocabulary knowledge, working memory, and cognitive flexibility often emerged as important variables across the models.

In optimal listening conditions, speech perception unfolds seemingly effortlessly; listeners parse the connected speech stream through word recognition, mapping word-sized acoustic units onto lexical items stored in memory (Mattys et al., 2005). However, deciphering speech becomes increasingly more effortful in adverse listening conditions (e.g., reduced audibility due to background noise and phonemic uncertainty due to hearing impairment or a disordered speech signal). Hierarchical frameworks of speech perception suggest that as listening conditions become more adverse and speech deviates from typical norms (i.e., noncanonical speech), listeners make use of lower level segmental (e.g., acoustic–phonetic) and, subsequently, suprasegmental (e.g., lexical stress) acoustic cues to aid in word recognition and speech segmentation (Mattys et al., 2005). Theoretical accounts posit that the unique demands of the listening situation modulate the level and type of acoustic information extracted and processed by the listener (Mattys et al., 2012; see also the studies of Borrie, Baese-Berk, et al., 2017, for empirical support). Furthermore, listener success with accurately perceiving degraded or, otherwise, noncanonical speech is predicted by their ability to readily and flexibly adapt their perceptual strategies to exploit the most salient acoustic information to support speech understanding (Borrie, Baese-Berk, et al., 2017).

Likewise, perceptual adaptation or learning (i.e., experience-evoked adjustments to the speech perception

Correspondence to Kaitlin L. Lansford: klansford@fsu.edu. ***Disclosure:*** *The authors have declared that no competing financial or nonfinancial interests existed at the time of publication.*

system) of noncanonical speech also relies on the flexibility of the listener's perceptual system. Structured familiarization paradigms, in which listeners are exposed to noncanonical speech signals, are often used to examine the perceptual learning phenomenon. Prior experience with noncanonical speech allows the listener to map the atypical or degraded acoustic cues onto linguistic categories stored in memory, resulting in improved perception of that speech in ensuing encounters (Samuel & Kraljic, 2009). As the ideal adaptor framework outlines, the distributional regularities of the noncanonical speech signal drive this cue-to-category mapping process. During familiarization, listeners exploit acoustic regularities available in the speech signal, subsequently retuning their linguistic categories (i.e., "beliefs") to account for the relevant degraded (but still useful) acoustic–phonetic information (Kleinschmidt & Jaeger, 2015). This learning, or belief-updating, improves lexical access, thereby supporting word recognition and speech segmentation. Thus, the extent to which a listener benefits from prior experience depends largely on their ability to identify and extract salient acoustic information from the impoverished speech signal to support cue-to-category mapping for improved perception of and adaptation to noncanonical speech.

Dysarthria, a motor speech disorder arising from neurological damage or disease, results in speech production patterns that deviate from typical norms. Consequently, perception of dysarthric speech can be effortful. The production of crucial lower level segmental (e.g., acoustic–phonetic) and suprasegmental (e.g., lexical stress, rate, and rhythm) acoustic cues is often impacted by the underlying movement disorder in dysarthria, thereby challenging the listener's ability to exploit salient acoustic information to segment dysarthric speech signals accurately and recognize spoken words (Liss et al., 1998, 2000). To the extent that the aberrant production patterns occur systematically, perception of dysarthric speech can improve with experience (e.g., Borrie, Lansford, et al., 2017; Lansford et al., 2019). Indeed, across a series of studies, researchers have rigorously demonstrated improved perception of dysarthric speech for listeners who completed a structured familiarization experience, in which lexical feedback is provided to facilitate mapping of degraded acoustic cues onto linguistic targets (see the study of Borrie & Lansford, 2021, for a recent review). The clinical implications of this body of work are substantial and point to a listener-targeted intervention option for intelligibility deficits in dysarthria.

Despite the well-documented demonstration of robust and clinically significant intelligibility improvements following familiarization with dysarthric speech, considerable variability in intelligibility outcomes, both before and after training, is observed across listeners (e.g., Borrie, Lansford, et al., 2017; Hirsch et al., 2021). Such individual differences in intelligibility performance suggest that some listeners are simply better able to extract salient acoustic–phonetic information from the dysarthric speech signal to support perception and adaptation. Although recent work has implicated rhythm perception abilities as a potential source of individual variability in intelligibility outcomes (Borrie, Lansford, et al., 2017; Borrie et al., 2018), more research is needed to determine listener-related parameters that drive intelligibility improvement following familiarization with dysarthric speech. Given that successful speech perception and adaptation demand flexibility of the listener's perceptual system, it follows that select cognitive-linguistic domains might account for some of the individual variability in intelligibility outcomes.

Theoretical models of effortful listening (e.g., Ease of Language Understanding and the Framework for Understanding Effortful Listening) provide additional theoretical support for examining cognitive-linguistic domains as potential sources of individual variability in perceptual outcomes associated with dysarthric speech (e.g., Pichora-Fuller et al., 2016; Rönnberg, 2003). Briefly, models of effortful listening posit cognitive resources are recruited and allocated to support speech perception when there is a mismatch between the incoming acoustic signal and the linguistic categories stored in memory. Cognitive resource allocation permits simultaneous processing of multimodal information (e.g., audiovisual and contextual information) associated with the incoming degraded speech stream to facilitate lexical access (see the study of Rönnberg et al., 2019, for a detailed discussion). Model assumptions indicate, however, that if the mismatch between the acoustic input and stored linguistic categories is extreme, as is often the case with dysarthric speech, additional recruitment of cognitive faculties might be insufficient for resolving the incoming speech signal (Ohlenforst et al., 2017; Rönnberg et al., 2019).

Although examining relationships between cognitive-linguistic parameters and dysarthric speech perception and adaptation is theoretically motivated, these relationships have received limited attention in the disordered speech literature. Certainly, we can draw upon empirical support from cross-disciplinary findings, which collectively show that cognitive-linguistic parameters support perception and adaptation to noncanonical speech, including speech in noise (e.g., Rudner & Rönnberg, 2008; Rudner et al., 2009), time-compressed speech (Kennedy-Higgins et al., 2020), noise-vocoded speech (O'Neill et al., 2019), and accented speech (e.g., Adank & Janse, 2010; Bent et al., 2016; Heffner & Myers, 2021; Ingvalson et al., 2017a). However, such relationships with noncanonical speech might not be comparable to those with dysarthric speech, which, due to the underlying movement disorders, can result in inconsistent and sometimes unpredictable speech degradation. Thus, it is warranted to evaluate intelligibility

outcomes associated with dysarthric speech perception and adaptation relative to specific cognitive-linguistic factors implicated in supporting effortful speech perception, including working memory, attention, cognitive flexibility, processing speed, and vocabulary knowledge. In the following section, we review these specific cognitive-linguistic factors, delineating the theoretical and empirical support for evaluating each relative to perception and adaptation to dysarthric speech.

Cognitive-Linguistic Parameters Linked to Effortful Speech Perception

Working memory, defined broadly as a limited capacity memory system that temporarily stores and manipulates information involved in complex tasks, is theorized to support speech perception when there is a mismatch between the acoustic signal and linguistic categories stored in long-term memory that restricts lexical access (Rönnberg, 2003). Empirical evidence from related disciplines largely supports these theoretical postulations: Increased working memory capacity has been linked to better perception of speech in noise (Ellis et al., 2015; Gordon-Salant & Fitzgibbons, 1997; Souza & Arehart, 2015) and accented speech (Ingvalson et al., 2017a). The relationship between working memory and dysarthric speech perception, however, remains largely unclear due to limited research and contradictory findings across studies (e.g., Ingvalson et al., 2017b; Lee et al., 2014; McAuliffe et al., 2013). Reliance on working memory is hypothesized to diminish, however, following structured familiarization with the noncanonical speech (Lunner, 2003). Familiarization is theorized to promote mapping the noncanonical acoustic cues onto linguistic categories stored in memory, thereby improving lexical access and reducing the need to recruit additional cognitive resources to resolve the speech signal. Results from several studies support these theoretical assumptions; dependence on working memory is reduced for listeners following familiarization with compressed speech signals (Lunner, 2003; Lunner & Sundewall-Thorén, 2007; Rudner et al., 2008; Rudner et al., 2009). The relationship between working memory and perceptual adaptation to dysarthric speech has not yet been explored.

Inhibitory control of attention, alternatively referred to as selective attention, is an umbrella term defined as the ability to suppress or inhibit goal-irrelevant stimuli to complete a given task (Diamond, 2013; Tiego et al., 2018). To successfully understand speech in adverse listening conditions, listeners appear to suppress or reduce their attention to distracting or irrelevant acoustic information and attend to the salient acoustic information that facilitates lexical segmentation and word recognition (Pichora-Fuller et al., 2016). Thus, it follows that individuals with better inhibitory control of attention should be better equipped to attend to useful acoustic–phonetic information that supports both perception of and adaptation to the noncanonical speech signal. Data supporting this presupposition are mixed. Bent et al. (2016) found no predictive relationship between a measure of inhibitory control and perception of accented or dysarthric speech, and this finding was corroborated, in large part, by a pair of investigations also with accented and dysarthric speech (Ingvalson et al., 2017a, 2017b). However, the listener's age might mediate the influence of inhibitory control on the perception of noncanonical speech (e.g., Bent et al., 2016; Dey & Sommers, 2015; Ingvalson et al., 2017a). Importantly, increased inhibitory control has been linked to greater and quicker adaptation to unfamiliar accents (Banks et al., 2015; Janse & Adank, 2012). To our knowledge, inhibitory control has not yet been explored relative to adaptation to dysarthric speech.

Cognitive flexibility, also referred to as task- or attention-switching, is an executive function that describes the ability to quickly and flexibly adapt to a changing environment (Cools, 2015). Undeniably, listening to speech in adverse conditions presents a situation where successful speech perception is likely supported by elevated cognitive flexibility. To accurately perceive speech under conditions of phonemic uncertainty, listeners are thought to readily adapt the perceptual strategies used to parse the incoming speech signal into the correct word-sized units (e.g., relying on lower level segmental and suprasegmental information to identify word boundaries) and match those units onto the correct linguistic forms stored in memory (Mattys et al., 2012). It is therefore postulated that those with greater cognitive flexibility capacity might be better able to flexibly adapt their perceptual strategies to decipher the noncanonical acoustic–phonetic input (Bent et al., 2016). Cognitive flexibility is also assumed to support perceptual adaptation to noncanonical speech signals. To derive benefit from the familiarization experience, listeners appear to adeptly navigate the incoming speech signal, extracting salient acoustic cues to facilitate mapping to linguistic categories stored in memory (Samuel & Kraljic, 2009). There is emerging empirical support for these theoretical assumptions. Adank and Janse (2010) found cognitive flexibility in older adults predicted perception of accented speech. Likewise, increased cognitive flexibility capacity was found to support older adults' perception of accented and dysarthric speech (Ingvalson et al., 2017a, 2017b). However, the relationship between cognitive flexibility and adaptation to noncanonical speech signals remains underexplored, and the limited findings are largely inconsistent (e.g., Colby et al., 2018; Heffner & Myers, 2021; Janse & Adank, 2012; Scharenborg et al., 2015).

Processing speed, defined generally as the time required to process a specific quantity of information (DeLuca & Kalmar, 2013), is assumed to support the listener's completion of time-constrained tasks in adverse conditions. Timely and accurate lexical segmentation and word recognition rely on the listener's ability to disambiguate and decode the degraded acoustic–phonetic information quickly enough to match the pace of the unfolding speech signal (Pichora-Fuller, 2003; Rönnberg, 2003). Furthermore, increased processing speed capabilities are hypothesized to support the listener's ability to rapidly extract acoustic regularities in the unfolding speech signal that drive adaptation (Neger et al., 2014). Empirical evidence lends some support to theoretical assumptions. Improved perception of time-compressed speech was revealed when pauses were inserted at syntactically appropriate places, leading the authors to speculate that the pauses allowed speech processing to catch up with the unfolding, time-compressed signal (Wingfield et al., 1999). However, accented and dysarthric speech perception studies have revealed inconsistent findings (Adank & Janse, 2010; Ingvalson et al., 2017a, 2017b). Evidence linking processing speed to perceptual adaptation, however, is scant and mixed. Although processing speed did not predict learning of nonnative phonetic contrasts (Heffner & Myers, 2021), faster processing speed has been linked to faster adaptation to noise-vocoded speech (Neger et al., 2014).

Advanced *vocabulary knowledge,* that is, knowledge of words and their meanings, might result from increased experience with a greater variety of exemplars for each lexical item. This experience might aid in word recognition under adverse listening conditions (Bent et al., 2016). It is assumed that if sufficient acoustic–phonetic information can be realized from an impoverished signal, those with increased vocabulary knowledge might be better able to leverage this information to reconstruct and ultimately decipher the degraded input (Cooke, 2006). This assumption appears to be supported by research findings that link advanced vocabulary knowledge to better perception of accented (Bent et al., 2016; Janse & Adank, 2012) and dysarthric speech signals (Bent et al., 2016; Borrie, Lansford, et al., 2017; McAuliffe et al., 2012; although cf. Ingvalson et al., 2017a). The relationship between vocabulary knowledge and perceptual adaptation, however, is less clear. It stands to reason that increased vocabulary knowledge should support the cue-to-category mapping process that underlies perceptual adaptation; however, the empirical support for this hypothesis is mixed. Although vocabulary knowledge has been revealed to support lexically guided perceptual learning (Colby et al., 2018), learning a novel accent (Banks et al., 2015; Janse & Adank, 2012), and learning nonnative phonetic contrasts (Heffner & Myers, 2021), it did not predict intelligibility improvement following familiarization with a moderately severe speaker with dysarthria (Borrie, Lansford, et al., 2017).

This Study

Recruitment of cognitive resources to support dysarthric speech perception and adaptation is presumed by theoretical accounts of effortful listening (Pichora-Fuller et al., 2016; Rönnberg, 2003). However, models also posit that when the mismatch between the acoustic input and linguistic categories stored in memory is extreme, as is often the case with severe dysarthria, additional recruitment of cognitive supports might not be sufficient to achieve accurate speech understanding (Rönnberg, 2003). This postulation suggests that the relationships between cognitive functions and perceptual outcomes might depend on the nature and severity of the speech degradations. Although select cognitive factors have been linked to perception of dysarthric speech (Bent et al., 2016; Ingvalson et al., 2017b; McAuliffe et al., 2014), the extent of their roles in perceptual learning of dysarthric speech is currently unknown.

In this study, a large sample of 156 young adult participants completed speaker-specific perceptual training, yielding two primary perceptual outcomes: *initial intelligibility* and *intelligibility improvement.* Initial intelligibility is presumed to reflect the extent to which a listener accurately maps the degraded acoustic cues stored to linguistic units stored in memory (i.e., perception), whereas *intelligibility improvement* reflects the listener's ability to identify and acquire knowledge of the speaker's acoustic cue distributions during the lexically guided familiarization experience (i.e., adaptation). The listeners also completed five subtests from the National Institutes of Health (NIH) Toolbox Cognition Battery. The resulting cognitive-perceptual data were analyzed to address the following research questions: (a) Are select cognitive-linguistic factors and their two-way interactions predictive of initial intelligibility and intelligibility improvement associated with structured familiarization with dysarthric speech? (b) Do predictive models yield different constellations of important cognitive-linguistic factors for three speakers with dysarthria who differ in terms of perceptual characteristics and level of intelligibility impairment?

To address these questions, we build predictive cognitive models of initial intelligibility and intelligibility improvement for three speakers with dysarthria, who vary in terms of perceptual features present and level of intelligibility impairment. We anticipate that cognitive-linguistic models will yield complex predictive relationships of initial intelligibility and intelligibility improvement of dysarthric speech. We also expect that the constellations of important cognitive-linguistic factors and their interactions will vary across the speaker conditions due to the inherent differences

in both levels of intelligibility impairment and perceptual characteristics (i.e., segmental and suprasegmental degradation) of the speech disorder. Furthermore, given that the intelligibility measures (i.e., initial intelligibility and intelligibility improvement) are considered to capture related, yet distinct, perceptual information, we anticipate some, but not complete, overlap across the predictive models.

Method

This investigation was completed as part of a larger, ongoing study investigating speaker and listener parameters associated with perceptual learning of dysarthric speech. Only procedures relating to the current investigation of cognitive parameters and intelligibility outcomes in younger adults with intact hearing are reported here. This study was approved by the Florida State University (FSU) and Utah State University (USU) institutional review boards.

Participants

A total of 158 listener participants were recruited from the FSU and USU communities and surrounding areas for this prospective study and were randomly assigned to one of three speaker-training conditions. All listeners were native speakers of American English and were between the ages of 18 and 50 years. As per self-report, listeners had no history of speech, language, or cognitive disorders. Listeners with hearing thresholds > 10 dB, obtained from the NIH Toolbox words-in-noise test (Zecker et al., 2013), were excluded from the analysis. Thus, data from the remaining 156 adult listeners with normal hearing were included in the analyses. Relevant listener demographic information is presented in Table 1. Listeners provided informed consent and received a gift card or course credit for their participation.

Speakers and Stimuli

The speakers and speech stimuli used for the current investigation were selected from a database of speakers with dysarthria, collected in the Motor Speech Disorders Lab at Arizona State University (see the study of Liss et al., 2009, for a description of recording procedures). This study utilized speech stimuli, audio-recorded productions of a set of 80 phrases, and a reading passage, produced by three male speakers of American English with moderate-to-severe dysarthria due to Parkinson's disease (PD), amyotrophic lateral sclerosis (ALS), or cerebellar disease (ataxia). Table 2 provides a detailed description of the three speakers with dysarthria.

The audio-recorded speech stimuli were used to create a three-phase perceptual training paradigm (pretest, familiarization, and posttest), resulting in three speaker-specific perceptual training conditions. The set of 80 syntactically plausible but semantically anomalous phrases was divided into two smaller subsets and used as the speech stimuli for the pretest (20 phrases) and posttest (60 phrases) phase of the paradigm. These low-predictability phrases, developed by Liss and colleagues for studies of dysarthric speech perception, restrict the listener's use of higher level cognitive-linguistic information to resolve the speech signal (e.g., *frame her seed to answer;* Liss et al., 1998, 2000). Since familiarization is thought to support the cue-to-category mapping process (i.e., bottom-up processing), the use of such stimuli is well justified. These phrases alternate in metrical stress and range from three to five words in length. The pretest and posttest stimuli were presented in randomized order. However, all listeners were exposed to the same sets of pretest and posttest stimuli. The pretest and posttest stimuli sets were carefully balanced for the number of words and metrical stress. Furthermore, intelligibility levels of testing sets were balanced, according to listener transcription data collected during pilot phases of this work.

The audio recordings of an adapted version of the Grandfather Passage were paired with lexical feedback in the form of an orthographic transcription of the intended targets and used as speech stimuli for the familiarization phase of the paradigm. The Grandfather Passage, written by Darley et al. (1975) for the assessment of dysarthric speech production, provides adequate lexical and structural complexity for a brief connected speech sample

Table 1. Listener demographics per speaker condition.

Characteristic	Ataxia *n* = 54: *M* (*SD*)	ALS *n* = 49: *M* (*SD*)	PD *n* = 53: *M* (*SD*)	*p* value
Age in years	22.9 (6.3)	22.3 (7.3)	21.5 (2.9)	.13
Working memory	104 (10)	104 (12)	102 (13)	.90
Inhibitory control of attention	94 (15)	93 (14)	90 (17)	.60
Cognitive flexibility	107 (15)	104 (15)	104 (17)	.60
Vocabulary knowledge	108 (12)	105 (14)	110 (12)	.20
Processing speed	115 (19)	111 (21)	114 (19)	.60

Note. PD = Parkinson's disease; ALS = amyotrophic lateral sclerosis. Moreover, *p* values are based on Kruskal–Wallis Rank Sum Tests.

Table 2. Speaker characteristics.

Speaker/etiology	Age	Dysarthria type	Severity	Perceptual characteristics
Ataxia	73	Ataxic	Severe	Slow rate, imprecise consonants, irregular articulatory breakdown, equal and even stress, monotone, monoloudness
PD	80	Hypokinetic	Moderate	Variable rate, short rushes of speech, irregular pauses, imprecise consonants, monotone, monoloudness
ALS	56	Mixed	Moderate	Slow rate, imprecise consonants, equal and even stress, monotone, monoloudness, strained-strangled vocal quality, hypernasality

Note. PD = Parkinson's disease; ALS = amyotrophic lateral sclerosis.

(Powell, 2006). The passage comprises 35 phrases, ranging in length from three to 12 words, with three to 14 syllables per phrase.

Experimental Protocol

The perceptual and behavioral data were collected in a single session (approximately 2 hr) at the FSU or USU laboratories. All participants completed the perceptual training protocol followed by the NIH Toolbox Cognition Battery (Weintraub et al., 2013). Data were collected while adhering to social distancing and other COVID-19 mitigation guidelines (e.g., maintaining a 6 ft distance between the research assistant and the participant, wearing masks, temperature checks, and conducting a COVID-19 exposure questionnaire). Before the perceptual training task, participants were presented with a short audio clip and asked to adjust the volume to a comfortable listening level. The volume remained at this level for the duration of the perceptual task.

Perceptual Training Paradigm

A three-phase, lexically guided perceptual training paradigm (pretest, familiarization, and posttest) was used to collect initial intelligibility and intelligibility improvement data from the listeners. The listener participants were randomly assigned to one of the speaker-specific perceptual training conditions, receiving training with either the ataxic, PD, or ALS speaker while wearing headphones and seated at a computer workstation. All task-related instructions were provided on the computer screen throughout perceptual training. During the pretest phase, listeners were informed that they would be listening to a speaker with a speech disorder and that it might be difficult to understand what was being said. Listeners were instructed to listen carefully, as they would hear each phrase only once, and to transcribe what they heard. They were encouraged to guess if they were unsure. Listeners transcribed 20 phrases during the pretest phase. Immediately following the pretest, participants completed a lexically guided familiarization task. Briefly, participants were instructed to listen to the same speaker's production of each phrase of the Grandfather Passage while simultaneously following along with the orthographic transcription presented on the screen (i.e., lexical feedback). The passage phrases were presented one at a time, and participants were instructed to advance to the next phrase when ready. After familiarization, listeners completed the posttest phase, in which they were asked to listen to and transcribe 60 novel phrases produced by the same speaker heard in the prior two phases. The same task instructions provided for the pretest were reiterated for the posttest. All phases of the perceptual training paradigm were self-guided and untimed, but on average, most listeners completed the full task in 30 min.

NIH Toolbox Cognition Battery

The NIH Toolbox Cognition Battery, administered via an iPad Pro, was used to obtain measures of the following cognitive-linguistic constructs: working memory, inhibitory control of attention, cognitive flexibility, processing speed, and vocabulary knowledge. The specific cognitive tasks used to quantify each construct and their brief descriptions can be found in Table 3. All tasks were administered according to the standardized procedures outlined in the NIH Toolbox Administrator's manual, with minimal adjustments to adhere to social distancing guidelines. All analyses used the obtained age-corrected standard scores.

Data Analysis

Transcript Analysis

Pretest and posttest listener transcripts were scored for words correct using Autoscore, an open-source computer-based tool for automated intelligibility scoring (http://autoscore.usu.edu; Borrie, Barrett, & Yoho, 2019). We selected rules in Autoscore to score words as correct if they match the intended target exactly or differed only by tense or plurality. Homophones and obvious spelling errors were scored as correct using a preprogrammed "default" list of common misspellings. A percent word correct (PWC) score was tabulated for the pretest and posttest experimental phases, resulting in a pretest PWC score and a posttest PWC score for each listener. These PWC scores were used to quantify initial intelligibility

Table 3. Tasks used by the National Institutes of Health Toolbox Cognition Battery to quantify each of the cognitive-linguistic constructs.

Construct	Measurement task	Description of the task
Working memory	List Sorting Working Memory Test	A series of pictures from two different categories (e.g., food and animals) are presented one at a time during a trial. At the end of the trial, the participant is instructed to recall the names of the pictures from one category first, in size order (smallest to largest), and then complete the same task with the other category.
Inhibitory control of attention	Flanker Inhibitory Control and Attention Test	A row of arrows is presented. The participant is instructed to select the direction that the middle arrow points (right or left). Distractor arrows sometimes point in the same or opposite direction.
Cognitive flexibility	Dimensional Change Card Sort Test	Two test pictures that differ along two dimensions relative to a target picture are presented. The participant is instructed to select the test picture that matches the target picture based on one dimension. After several trials, the participant sorts the images based on the other dimension.
Processing speed	Pattern Comparison Processing Speed Test	Two pictures are presented, and the participant is instructed to determine if the pictures are the same or different.
Vocabulary knowledge	Picture Vocabulary test	Four pictures are presented, and the participant hears a word spoken aloud. The participant is instructed to select the picture that matches the spoken word.

(pretest PWC) and intelligibility improvement (posttest PWC after accounting for pretest PWC).

Statistical Analysis

Initial analyses were used to investigate distributions of the variables, assess for baseline differences, test for changes from pretest to posttest, and calculate correlations between variables of interest. Descriptive statistics (means and standard deviations) were calculated for all variables of interest (e.g., cognitive measures and intelligibility), and baseline differences between conditions were tested using Kruskal–Wallis Rank Sum tests, as these data did not always meet statistical assumptions of parametric testing (e.g., normality). Paired-samples t tests determined average intelligibility improvement from pretest to posttest with effect sizes estimated with Cohen's d for each speaker condition. Lastly, Pearson correlations were estimated between each cognitive measure and their relationship with intelligibility (both initial and improvement) by speaker condition.

Predictive modeling—using elastic net regression models—was used to assess the ability of the cognitive measures to predict both initial intelligibility and intelligibility improvement. Elastic net is a popular linear modeling approach built on linear regression that handles high multicollinearity naturally and is commonly used in human interaction literature (e.g., Borrie, Barrett, Willi, & Berisha, 2019; Borrie et al., 2020). Elastic net was selected for this study for three primary reasons: (a) the cognitive measures are generally correlated, and elastic net handles multicollinearity well; (b) prior information suggests the relationships between the cognitive variables and the intelligibility measures are likely complex, and elastic net permits the use of a complicated specification to accommodate modeling of complex data; and (c) elastic net is built on regression, facilitating ease of interpretation. Unlike bivariate correlations (and related methods), which can miss the more complex underpinnings of relationships between variables, elastic net can better capture the complexity by allowing for far more complex specifications of the model (e.g., include all interactions). Model-specific hyperparameters were selected based on tenfold cross-validation. Predictive models were run separately for each speaker condition. For each speaker, two model specifications were used (a) main effects only and (b) main effects and all two-way interactions.

Results were derived from the model predictions and variable importance metrics, as well as post hoc assessment of predictions to understand the nature of the important interactions. Predictive accuracy included the root-mean-squared error (RMSE) and R^2 values for the cross-validated predictions. This approach pushes the model to find the variation that is predictive of data it has not learned from. In other words, the model is incentivized to find the relationships that generalize to unlearned data and are not specific to the peculiarities of the sample. Variable importance was extracted from each model using a permutation method wherein each variable is randomly permuted, and the difference in predictive accuracy between the observed variable and its randomly permuted counterpart was derived. Relative importance was calculated based on these permutations, ranging from 1 (*most important*) to 0 (*least important*). As such, relative importance reflects the degree that model performance suffers when that variable is removed, relative to the other variables that remain in the model. Interactions selected as important were descriptively assessed by graphical means, such that we assessed the prediction stratified by the variables in the interaction. All analyses were completed in the R statistical environment version 4.2.0 (R Core Team,

2022) with the caret, tidyverse, janitor, and gtsummary packages (Firke et al., 2021; Kuhn, 2022; Sjoberg et al., 2021; Wickham et al., 2019).

Results

Descriptive Results

First, descriptive statistics associated with listeners in each speaker-training condition, including listener demographics and cognitive measures averages, are reported in Table 1. As shown, Kruskal–Wallis Rank Sum tests indicated that there were no significant differences between the listeners in the three speaker conditions on any of the measures. Figure 1 shows the change from pretest to posttest for each listener by speaker condition. Paired-samples *t* tests showed there were significant intelligibility improvements for each speaker from pretest to posttest (all $p < .001$). Average intelligibility for the ataxic speaker increased from 35.3% (*SD* = 7.7%) at pretest to 50% (*SD* = 7.1%) at posttest. For the speaker with PD, intelligibility increased from 52.9% (*SD* = 7.1%) at pretest to 57.8% (*SD* = 7.8%) at posttest, and for the speaker with ALS, intelligibility increased from 77.8% (*SD* = 5.4%) at pretest to 84.4% (*SD* = 3.6%) at posttest. All differences had large effect sizes (*d*s = 2.42, 0.77, 1.47, for the speakers with ataxia, PD, and ALS, respectively).

Pearson correlations between all cognitive measures, initial intelligibility, and intelligibility improvement, stratified by speaker condition, are shown in Figure 2. All correlations larger than approximately ± 0.29 are significant at alpha = .05 across the conditions. For the speaker with ataxia, working memory and vocabulary knowledge were the largest correlates with initial intelligibility and intelligibility improvement. For the speaker with PD, only vocabulary knowledge significantly correlated with initial intelligibility (and none correlated with intelligibility improvement). For the speaker with ALS, none of the cognitive measures significantly correlated with either initial intelligibility or intelligibility improvement, although cognitive flexibility was close.

Figure 1. Pretest to posttest intelligibility for each listener by speaker condition. The point and lines are individual pretest to posttest changes. The box-and-whisker plots show the distribution at each time point, with the median (line in the box middle of the box), the range for the middle 50% of data (the box), and the minimum and maximum up to 1.5 times the interquartile range (the whiskers). ALS = amyotrophic lateral sclerosis; PD = Parkinson's disease.

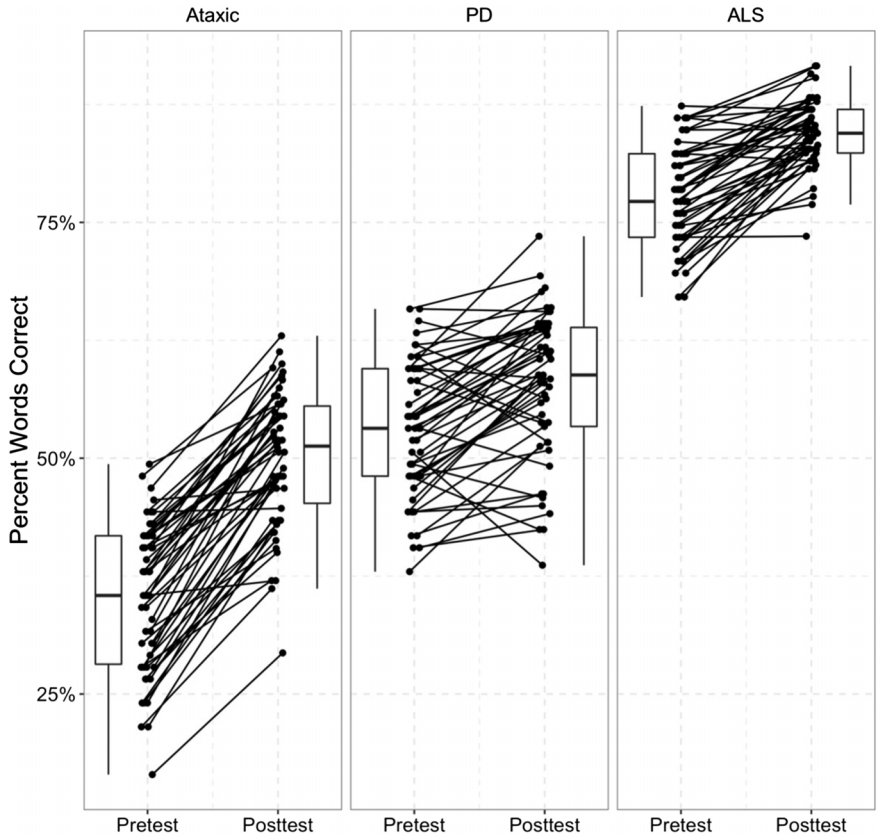
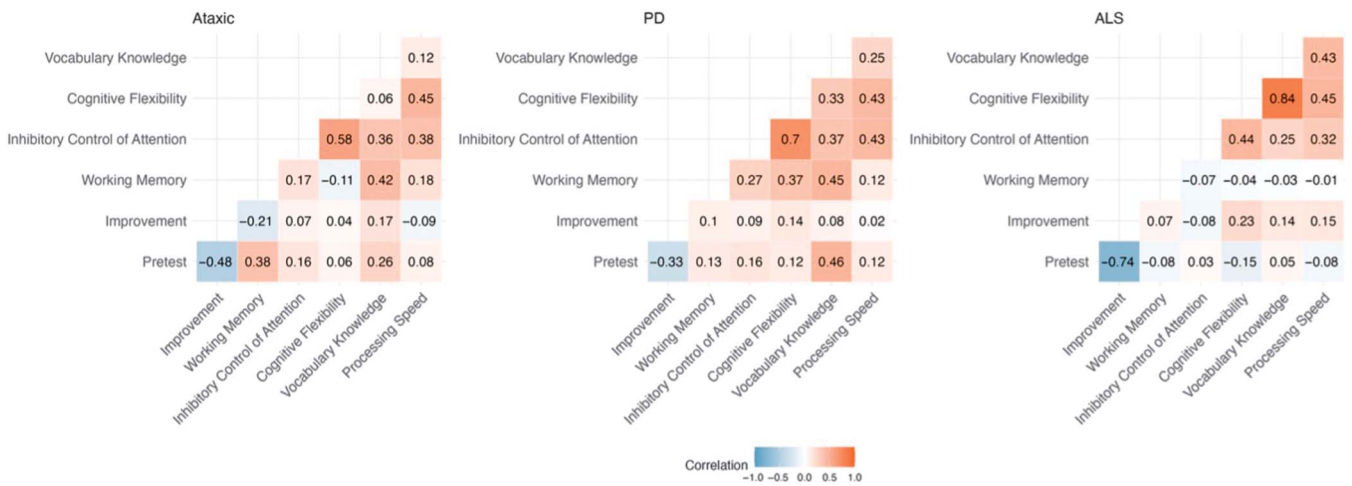


Figure 2. Correlograms showing Pearson correlations between each cognitive measure and intelligibility by speaker. ALS = amyotrophic lateral sclerosis; PD = Parkinson's disease



Initial Intelligibility

We used six elastic net models to model the complex interplay of cognitive measures on initial intelligibility, two for each speaker condition. The predictive accuracies (cross-validated R^2 and RMSE) of these models are presented in Table 4. Based on R^2 values, the best model for the speaker with ataxia was the main effects model (i.e., no interactions). For both the speakers with PD and ALS, the best models were the interactions specification (i.e., interactions and main effects both specified). These results show that cognitive measures explained between 36% and 44% of the cross-validated variation in initial intelligibility, depending on speaker condition. Within each speaker condition, the RMSE values did not differ much between the specifications.

Table 4. Cross-validated predictive accuracy (R^2 and RMSE) for each elastic net model by initial intelligibility and intelligibility improvement by speaker condition.

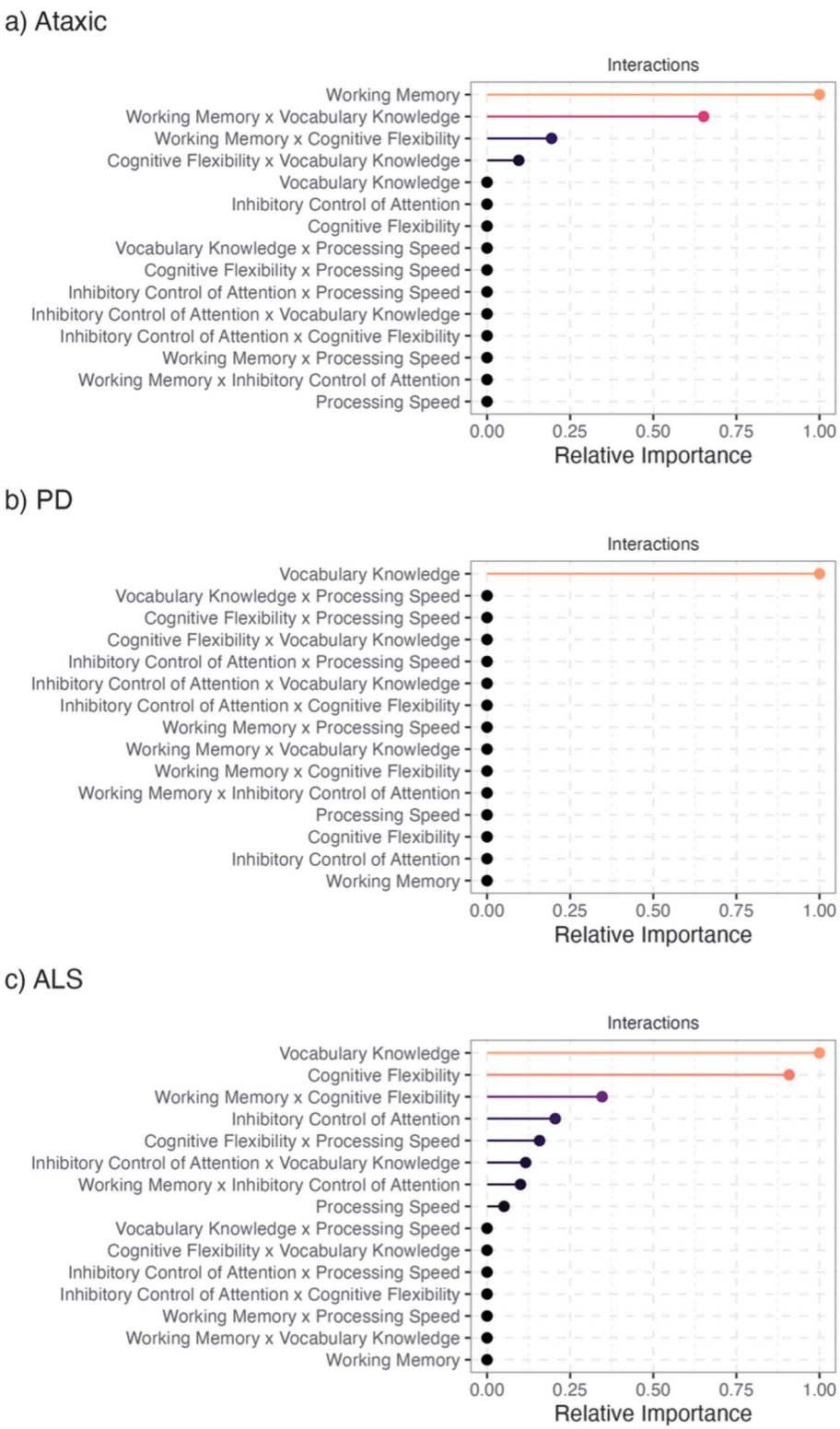
Condition	Model specification	R^2	RMSE
Initial intelligibility			
Ataxia	Main effects	.396	0.075
Ataxia	Interactions	.244	0.078
PD	Main effects	.350	0.063
PD	Interactions	.442	0.064
ALS	Main effects	.159	0.054
ALS	Interactions	.367	0.055
Intelligibility improvement			
Ataxia	Main effects	.232	0.054
Ataxia	Interactions	.406	0.054
PDM	Main effects	.367	0.060
PDM	Interactions	.422	0.062
ALS	Main effects	.199	0.039
ALS	Interactions	.409	0.036

Note. RMSE = root-mean-squared error; PD = speaker with Parkinson's disease; ALS = speaker with amyotrophic lateral sclerosis.

The relative variable importance for initial intelligibility is shown in Figure 3, with all variables and their interactions included. Important variables varied by speaker condition. For the speaker with ataxia (Panel a), the most important variable for prediction was working memory, followed by three interactions, suggesting that some predictive power came from including those interactions specifically. These interactions included working memory and vocabulary knowledge, working memory and cognitive flexibility, and cognitive flexibility and vocabulary knowledge. As such, working memory, vocabulary knowledge, and cognitive flexibility (and their interactions) were most important for the speaker with ataxia. It is important to note that interactions are described here even though the model without interactions performed better for the speaker with ataxia. Model performance for the interaction specification was negatively impacted by the inclusion of several interactions that did not aid in prediction. That is, the additional, unimportant interactions created more noise in the model. A follow-up study will be necessary to assess the predictive accuracy of the selected variables alone (including interactions). Such a model would be expected to yield better predictive accuracy than the main effects model. We did not do this here to avoid overfitting the data.

For the speaker with PD (Panel b), a single variable was important: vocabulary knowledge. All other variables had zero variable importance for initial intelligibility for the speaker with PD. Several variables and interactions were important for the speaker with ALS (Panel c). The first two were similar in importance—vocabulary knowledge and cognitive flexibility—followed by the interaction

Figure 3. Variable importance for predicting initial intelligibility for each elastic net model by speaker condition. ALS = amyotrophic lateral sclerosis; PD = Parkinson's disease



between working memory and cognitive flexibility. The last five were not particularly important but were all more than zero; these included main effects and interactions of attention, cognitive flexibility, processing speed, and vocabulary knowledge.

Intelligibility Improvement

We used six elastic net models to model the complex interplay of cognitive measures on intelligibility improvement, two for each speaker condition. The predictive accuracies (cross-validated R^2 and RMSE) of these models are also shown in Table 4. Based on R^2 values, the best model specification for the three speaker conditions was the interaction specification (i.e., interactions and main effects both specified). These results show that cognitive measures explained between 41% and 42% of the cross-validated variation in intelligibility improvement, depending on speaker condition. Within each speaker condition, the RMSE values did not differ much between the specifications but mostly showed better performance for the interaction specification.

The relative variable importance for intelligibility improvement is shown in Figure 4. For the speaker with ataxia (Panel a), the most important variable for prediction was vocabulary knowledge, followed by working memory and the interaction between working memory and processing speed. The last two that had any relative importance were also interactions. For the PD speaker (Panel b), three interactions between three variables were most important: cognitive flexibility, vocabulary knowledge, and working memory. As such, these three variables and their combinations were most important for predicting improvement for the speaker with PD. The last that had any relative importance was vocabulary knowledge on its own. For the speaker with ALS (Panel c), the most important variable was attention. This was followed by three interactions with cognitive flexibility—processing speed, working memory, and vocabulary knowledge. The last variable with any importance was cognitive flexibility on its own. As such, inhibitory control of attention on its own and cognitive flexibility (with its combined effect with other variables) were most important for the speaker with ALS.

Discussion

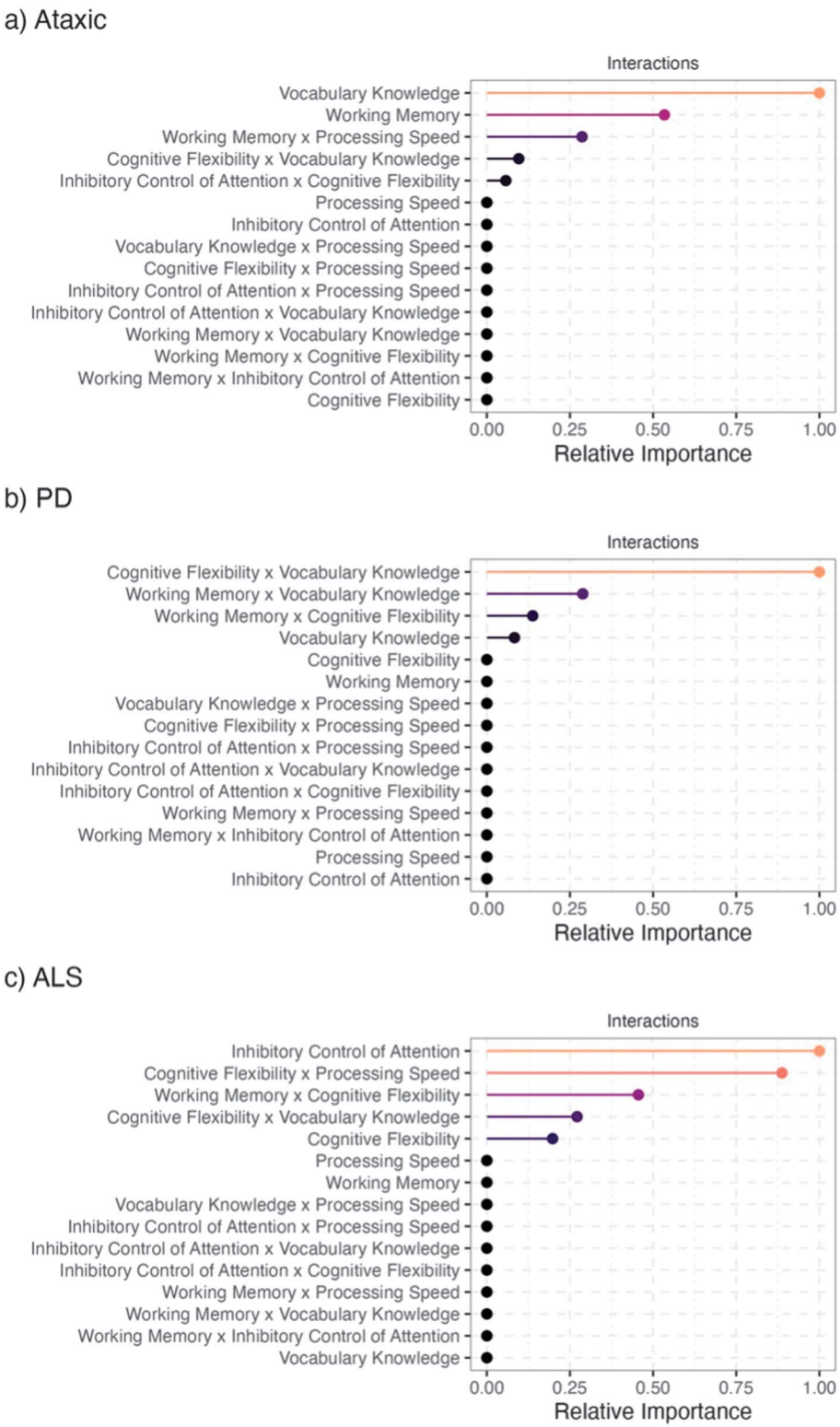
Motivated by theoretical frameworks that posit recruitment and allocation of additional cognitive-linguistic resources to support speech perception in effortful listening conditions (Pichora-Fuller et al., 2016; Rönnberg, 2003), this study examined the role of cognitive-linguistic parameters in the perception of, and adaptation to, neurologically degraded speech. While complex, the results collectively provide empirical support for theoretical models of effortful listening. A priori, we expected that the relationships between the cognitive-linguistic measures, intelligibility, and intelligibility improvement of dysarthric speech would likely be complex and that any single measure would be unlikely to represent variation in the intelligibility outcomes to any meaningful degree. The results of this work support this hypothesis. While some consistency across models was revealed (e.g., vocabulary knowledge was often positively linked to outcomes), unique constellations of select cognitive factors and their interactions predicted initial intelligibility and intelligibility improvement of the three different speakers with dysarthria. Recall that the speakers with dysarthria differed in terms of their overall intelligibility levels and perceptual characteristics (see Table 2); thus, it is unsurprising that the cognitive-linguistic prediction models of intelligibility outcomes yielded distinctive constellations for each speaker.

Speech Perception

Across the three speakers with dysarthria, prediction models that included only the main effects accounted for 16%–30% of the variance in initial intelligibility scores. When all two-way interactions were included, the prediction models accounted for 24%–44% of the variance. Aside from the main effects model for the speaker with ataxia, the prediction models that included the interactions accounted for more variance in initial intelligibility scores. Given the models performed best with the interactions included, the discussion is focused on interpreting those results. Broadly, the results of the prediction models highlighted the importance of vocabulary knowledge, either directly (speakers with PD and ALS) or indirectly (speaker with ataxia), for predicting initial intelligibility scores. For the speaker with PD (average baseline intelligibility of 53%), vocabulary knowledge was the only important variable indicated by both models. The relationship between vocabulary knowledge and initial intelligibility was positive for all three speakers, suggesting that those with advanced vocabulary knowledge demonstrated elevated perception of dysarthric speech before familiarization. This finding is consistent with prior work in this area (e.g., Bent et al., 2016; Borrie, Lansford, et al., 2017; McAuliffe et al., 2012) and supports theoretical assumptions that listeners leverage vocabulary knowledge to help reconstruct and decode degraded acoustic–phonetic information.

For the most severe speaker with ataxia (average baseline intelligibility of 35%), the most important predictor variable was working memory. The main effect of working memory was positive, suggesting those with advanced working memory abilities were better able to understand the severely degraded speech before

Figure 4. Variable importance for predicting intelligibility improvement for each elastic net model by speaker condition. ALS = amyotrophic lateral sclerosis; PD = Parkinson's disease



familiarization. Furthermore, the interaction between working memory and vocabulary knowledge was also important for predicting initial intelligibility. Reduced working memory abilities in listeners were offset by increased vocabulary knowledge. Overall, these findings are aligned with theoretical postulations; working memory is leveraged by listeners to temporarily store the unfolding speech signal while simultaneously processing multimodal input to facilitate lexical access.

Interestingly, working memory was not implicated by predictive models of initial intelligibility for the other two speakers with moderate intelligibility deficits (speakers with PD and ALS). Earlier work in this area revealed inconsistent findings relative to working memory and perception of dysarthric speech. While the work of McAuliffe et al. (2013) failed to demonstrate a significant relationship between working memory and dysarthric speech perception, results from another study by Lee et al. (2014) demonstrated a perceptual advantage for listeners with high working memory abilities compared to those with low working memory abilities. Another study revealed a predictive relationship between working memory and dysarthric speech perception for older but not younger listeners (Ingvalson et al., 2017b). The divergent findings across dysarthric perception studies might be due to differences in sample size or other methodological differences related to measurement (see Füllgrabe & Rosen, 2016, for a related meta-analysis that explores inconsistent findings across studies of working memory and speech-in-noise perception for listeners with normal hearing thresholds). However, in this study, we examined the relationship between working memory and initial intelligibility of dysarthric speech in three large samples of young listeners, using identical task and measurement procedures, and revealed different findings across the three speaker conditions. It is plausible, then, that discrepancies across studies might also track to speaker characteristics. Of interest, previous examinations of working memory and perception of dysarthric speech were restricted to mildly to moderately impaired speakers. In this study, elevated working memory was linked to better perception of the least intelligible speaker (with ataxia) exclusively. Thus, from these findings, we speculate that additional working memory resources might not be required when speakers have less degraded acoustics (i.e., moderate impairment). This is an empirical question deserving of future attention.

In addition to vocabulary knowledge, cognitive flexibility and, to a lesser degree, its interactions with working memory and processing speed were revealed as important factors by predictive models of initial intelligibility for the speaker with ALS, who was also the most intelligible (average baseline intelligibility of 78%). The main effect of cognitive flexibility was negative, suggesting that better perceptual outcomes were associated with reduced cognitive flexibility. This finding was not expected, as those with greater cognitive flexibility capacity were assumed to be better able to adapt their perceptual strategies to decipher the degraded acoustic–phonetic input (Bent et al., 2016). However, relevant important interactions indicate that increased cognitive flexibility might offset reduced working memory and processing speed abilities, resulting in better perception of the speaker with ALS. Cognitive flexibility and how it relates to intelligibility improvement are discussed below, but the current findings suggest that the relationship between this construct and intelligibility is complicated, likely involving other domains. Other, less important two-way interactions between inhibitory control of attention and vocabulary and working memory provide further evidence of the complexity of the relationships between cognitive-linguistic domains and intelligibility of dysarthric speech.

Perceptual Adaptation

Listeners demonstrated significant intelligibility improvement, on average and across speakers, following a structured familiarization experience. This study, to our knowledge, is the first to consider whether cognitive-linguistic domains, other than vocabulary knowledge (see the study of Borrie, Lansford, et al., 2017), predict intelligibility improvement following familiarization with dysarthric speech. Based on theoretical accounts (e.g., Kleinschmidt & Jaeger, 2015) and cross-disciplinary empirical support (e.g., Banks et al., 2015; Janse & Adank, 2012), predictive relationships between cognitive-linguistic domains and intelligibility improvement were hypothesized for listeners familiarized with dysarthric speech. Crucially, the current findings support this hypothesis; prediction models that included only the main effects accounted for 20%–37% of the variance in intelligibility improvement scores, depending on the speaker condition. Models that included all two-way interactions and the main effects accounted for more variance, at 41%–42% for all speakers. As aforementioned, given that the models performed best with the interactions included, the discussion is focused on those results.

While distinctive constellations and relative importance of cognitive-linguistic predictors of intelligibility improvement are evident across the three speaker conditions, there is an appreciable overlap that is worth noting. For example, greater perceptual adaptation following familiarization is predicted by greater cognitive flexibility and vocabulary knowledge, both directly and indirectly. Furthermore, these relationships were additive for the speakers with ataxia and PD; listeners with elevated cognitive flexibility and vocabulary knowledge experienced greater intelligibility improvements following familiarization than those with elevated levels in a single domain.

Additionally, and across the speaker conditions, cognitive flexibility and vocabulary knowledge interacted with other cognitive domains, including working memory and processing speed, to predict intelligibility improvement. These interactive relationships were largely positive and additive, such that listeners with increased ability in the cognitive-linguistic domains appeared to have greater intelligibility improvement following familiarization with dysarthric speech. Thus, together, the current findings provide additional evidentiary support linking cognitive-linguistic ability to perceptual adaptation to noncanonical speech.

Despite these commonalities, some interesting speaker-dependent findings emerged in the data. First, inhibitory control of attention was identified as the most important factor for predicting intelligibility improvement for the speaker with ALS. The positive relationship indicated that listeners who were better able to inhibit task-irrelevant information gleaned more from the familiarization experience, resulting in greater intelligibility improvement. Although inhibitory control of attention was not directly linked to intelligibility improvement for the other less intelligible speakers with dysarthria, the current findings add to the growing literature linking this executive function to perceptual adaptation to noncanonical speech (Colby et al., 2018; Scharenborg et al., 2015). Second, working memory was identified as an important factor for predicting intelligibility improvement for listeners of the ataxic speaker. The negative relationship suggested that those with reduced working memory learned more from the familiarization experience. This relationship was magnified for listeners with elevated levels of cognitive flexibility. Unsurprisingly, working memory was not positively related to intelligibility improvement, as structured familiarization has been demonstrated to minimize the need to recruit working memory to help resolve the degraded speech signal (Lunner, 2003; Lunner & Sundewall-Thorén, 2007; Rudner et al., 2008; 2009). However, working memory's overall negative relationship and its interaction with cognitive flexibility are more difficult to interpret. Interestingly, and consistent with theoretical assumptions and previous findings, working memory was not directly linked to intelligibility improvement for the other two speakers in this study. Thus, the relationship between working memory and adaptation to dysarthric speech is likely complex, mediated by interdependent variables, and should be the focus of future work.

Limitations and Future Work

Collectively, the cognitive-linguistic prediction models of intelligibility and intelligibility improvement align with theoretical models, adding to the emerging body of literature linking cognition to perception and adaptation of noncanonical speech. While we anticipated that the predictive models would vary across the speaker conditions due to the unique characteristics of each speaker, which varied in terms of not only intelligibility but also deviant segmental and suprasegmental features, it would be premature to definitively link the current findings to specific speaker characteristics from this study alone. Indeed, it is plausible to assume that the overall intelligibility of the speaker modulates the strength of the various predictive models, and the current results provide some support for this hypothesis. However, model differences cannot be explained exclusively by the level of intelligibility impairment. Thus, it will be important for future work to examine whether constellations of deviant speech features that characterize different speakers with dysarthria have a mediating role in the predictive relationships between cognitive-linguistic factors and perceptual outcomes. For example, one could imagine that processing dysarthric speech characterized by an abnormally slow speaking rate might demand greater working memory resources. This line of inquiry warrants systematic investigation.

It is important to interpret the current findings with caution. First, given that a between-subjects design was used to address the research questions, it is possible that differences in cognitive predictors for each speaker could also reflect differences in the participant groups. Second, although the NIH Toolbox offers a convenient and validated tool for quantifying the select cognitive-linguistic variables examined in this study (Heaton et al., 2014; Weintraub et al., 2013), we expected and observed multicollinearity among many of the cognitive-linguistic variables. Unexpected, however, were the differences in the strength and direction of the relationships between the variables across the speaker conditions (see Figure 2), despite no group-level differences in any of the individual cognitive parameters (see Table 1). In theory, the relationships between the cognitive-linguistic variables should be relatively consistent across listener groups, given that the measurements obtained were not dependent on the speaker condition. Thus, these findings suggest that even validated measures of cognitive behavior might be vulnerable to measurement error and, as such, specific relationships between cognition and intelligibility outcomes should be considered within such context.

The current results demonstrate that even in a young, healthy sample of listeners who likely have relatively minimal variability in cognitive abilities, cognitive-linguistic domains support intelligibility outcomes associated with understanding dysarthric speech. However, given that age-related cognitive changes have been well documented in older adults, even those free from neurological disease (e.g., Salthouse, 2004), the specific relationships might differ for older listeners. Indeed, preliminary work in this area supports this hypothesis. For example, relationships between working memory and perception of

dysarthric speech were stronger for older listeners than for their younger counterparts (Ingvalson et al., 2017b). Additionally, increased cognitive flexibility capacity was found to support older, but not younger, adults' perception of accented and dysarthric speech signals (Bent et al., 2016; Ingvalson et al., 2017a, 2017b).

Interestingly, in this study, inhibitory control of attention was not a key predictor of perceptual outcomes, linked only to adaptation to the speaker with ALS. Evidence from cross-disciplinary studies suggests that the listener's age might mediate the influence of inhibitory control on perception of noncanonical speech (e.g., Bent et al., 2016; Dey & Sommers, 2015; Ingvalson et al., 2017a). Thus, it might be the case that in older listeners who presumably display more variability in cognitive ability, relationships with inhibitory control and other domains might emerge. Indeed, older adults are frequent communication partners of people with dysarthria and have been demonstrated to derive perceptual benefits from perceptual training with dysarthric speech (Lansford et al., 2018). Thus, prediction models of perception and adaptation to speakers with dysarthria in older adult listeners are well justified.

Clinical Implications

While this study was focused on cognitive-linguistic predictors of understanding and adapting to dysarthric speech, we provide further evidence of the value of perceptual training, in which listeners are better able to understand a speaker with dysarthria following a structured familiarization experience. Familiarization paradigms offer a promising platform for listener-targeted remediation in dysarthria, in which the burden of behavioral change is offset from the speaker onto the listener (e.g., caregiver, spouse, friend, or practitioner). While the listeners, as a group, achieved significant intelligibility improvements from pretest to posttest, Figure 1 highlights considerable variability in intelligibility outcomes with the three speakers with dysarthria examined in this study. Indeed, most listeners achieved clinically significant gains in intelligibility; however, some appeared to receive no benefit from the familiarization experience. This finding highlights the importance of identifying the sources of such individual variability, particularly factors that might distinguish listeners who likely will benefit from training from those who likely will not. This study implicates several cognitive-linguistic domains as sources of variability, adding to earlier work linking expertise in rhythm perception to intelligibility outcomes (Borrie, Lansford, et al., 2017; Borrie et al., 2018). Further efforts in delineating the combination of factors that have utility in predicting whether an individual will benefit from perceptual training could be used to inform candidacy for this promising clinical intervention.

Conclusions

This study revealed cognitive-linguistic prediction models of perception (initial intelligibility) and adaptation following familiarization (intelligibility improvement) with speakers with dysarthria for young adult listeners. Thus, the findings extend empirical support for theoretical models of speech perception in adverse listening conditions to the neurologically degraded speech signal. Predictive relationships were complex; however, vocabulary knowledge, working memory, and cognitive flexibility emerged as important variables across many models. Model-based predictions provided insights on the high degree of interdependence of the cognitive domains in how they relate to intelligibility outcomes. Findings also suggest that the relative importance of each cognitive domain in predicting intelligibility outcomes depends on the baseline intelligibility of the speaker with dysarthria. Further attention to the role of speaker intelligibility and extension to older adult listeners are important directions for this inquiry.

Author Contributions

Kaitlin L. Lansford: Conceptualization (Lead), Formal analysis (Supporting), Funding acquisition (Lead), Methodology (Lead), Project administration (Lead), Resources (Lead), Writing – original draft (Lead), Writing – review & editing (Lead). **Tyson S. Barrett:** Conceptualization (Supporting), Data curation (Lead), Formal analysis (Lead), Funding acquisition (Supporting), Methodology (Lead), Software (Lead), Validation (Lead), Visualization (Lead) Writing – original draft (Supporting), Writing – review & editing (Supporting). **Stephanie A. Borrie:** Conceptualization (Lead), Formal analysis (Supporting), Funding acquisition (Lead), Methodology (Lead), Project administration (Lead), Resources (Lead), Writing – original draft (Supporting), Writing – review & editing (Supporting).

Data Availability Statement

The written stimuli targets, data, analysis code, and output are provided publicly at https://osf.io/y32qc/.

Acknowledgments

This research was supported by the National Institute on Deafness and Other Communication Disorders Grants R21DC018867. A portion of this work was conducted as part of an undergraduate honors thesis completed by Courtney Wilkinson at Florida State University. The authors would like to thank Julie Liss at Arizona State

University for sharing her speech sample database and extend our gratitude to Audrey Hendrix in the Motor Speech Disorders Lab at Florida State University and Taylor Hepworth in the Human Interaction Lab at Utah State University for assistance with data collection. Finally, the first author extends appreciation to the writing group, "The Write Stuff," for providing the structure and encouragement to prepare this article.

References

Adank, P., & Janse, E. (2010). Comprehension of a novel accent by young and older listeners. *Psychology and Aging, 25*(3), 736–740. https://doi.org/10.1037/a0020054

Banks, B., Gowen, E., Munro, K. J., & Adank, P. (2015). Cognitive predictors of perceptual adaptation to accented speech. *The Journal of the Acoustical Society of America, 137*(4), 2015–2024. https://doi.org/10.1121/1.4916265

Bent, T., Baese-Berk, M., Borrie, S. A., & McKee, M. (2016). Individual differences in the perception of regional, nonnative, and disordered speech varieties. *The Journal of the Acoustical Society of America, 140*(5), 3775–3786. https://doi.org/10.1121/1.4966677

Borrie, S. A., Baese-Berk, M., Van Engen, K., & Bent, T. (2017). A relationship between processing speech in noise and dysarthric speech. *The Journal of the Acoustical Society of America, 141*(6), 4660–4667. https://doi.org/10.1121/1.4986746

Borrie, S. A., Barrett, T. S., Liss, J. M., & Berisha, V. (2020). Sync pending: Characterizing conversational entrainment in dysarthria using a multidimensional, clinically informed approach. *Journal of Speech, Language, and Hearing Research, 63*(1), 83–94. https://doi.org/10.1044/2019_JSLHR-19-00194

Borrie, S. A., Barrett, T. S., Willi, M. M., & Berisha, V. (2019). Syncing up for a good conversation: A clinically meaningful methodology for capturing conversational entrainment in the speech domain. *Journal of Speech, Language, and Hearing Research, 62*(2), 283–296. https://doi.org/10.1044/2018_JSLHR-S-18-0210

Borrie, S. A., Barrett, T. S., & Yoho, S. E. (2019). Autoscore: An open-source automated tool for scoring listener perception of speech. *The Journal of Acoustical Society of America, 145*(1), 392–399. https://doi.org/10.1121/1.5087276

Borrie, S. A., & Lansford, K. L. (2021). A perceptual learning approach for dysarthria remediation: An updated review. *Journal of Speech, Language, and Hearing Research, 64*(8), 3060–3073. https://doi.org/10.1044/2021_JSLHR-21-00012

Borrie, S. A., Lansford, K. L., & Barrett, T. S. (2017). Rhythm perception and its role in perception and learning of dysrhythmic speech. *Journal of Speech, Language, and Hearing Research, 60*(3), 561–570. https://doi.org/10.1044/2016_JSLHR-S-16-0094

Borrie, S. A., Lansford, K. L., & Barrett, T. S. (2018). Understanding dysrhythmic speech: When rhythm does not matter, and learning does not happen. *The Journal of the Acoustical Society of America, 143*(5), EL379–EL385. https://doi.org/10.1121/1.5037620

Colby, S., Clayards, M., & Baum, S. (2018). The role of lexical status and individual differences for perceptual learning in younger and older adults. *Journal of Speech, Language, and Hearing Research, 61*(8), 1855–1874. https://doi.org/10.1044/2018_JSLHR-S-17-0392

Cooke, M. (2006). A glimpsing model of speech perception in noise. *The Journal of the Acoustical Society of America, 119*(3), 1562–1573. https://doi.org/10.1121/1.2166600

Cools, R. (2015). Neuropsychopharmacology of cognitive flexibility. In A. W. Toga (Ed.), *Brain mapping: A comprehensive reference, three volume set: Vol. 3: Social cognitive neuroscience, cognitive neuroscience, clinical brain mapping* (pp. 349–353). Academic Press. https://doi.org/10.1016/B978-0-12-397025-1.00253-0

Darley, F. L., Aronson, A. E., & Brown, J. R. (1975). *Motor speech disorders*. Saunders.

DeLuca, J., & Kalmar, J. H. (2013). *Information processing speed in clinical populations*. Psychology Press. https://doi.org/10.4324/9780203783054

Dey, A., & Sommers, M. S. (2015). Age-related differences in inhibitory control predict audiovisual speech perception. *Psychology and Aging, 30*(3), 634–646. https://doi.org/10.1037/pag0000033

Diamond, A. (2013). Executive functions. *Annual Review of Psychology, 64*(1), 135–168. https://doi.org/10.1146/annurev-psych-113011-143750

Ellis, R., Molander, P., Rönnberg, J., & Lyxell, B. (2015). *Predicting speech-in-noise perception using the trail-making task: Results from a large-scale internet study* [Paper presentation]. Conference on Cognitive Hearing Science for Communication, Linköping, Sweden. https://www.diva-portal.org/smash/record.jsf?pid=diva2%3A878445&dswid=-1708

Firke, S., Denney, B., Haid, C., Knight, R., Grosser, M., & Zadra, J. (2021). *Janitor: Simple tools for examining and cleaning dirty data* (R package Version 2.1.0) [Computer software]. The Comprehensive R Archive Network. https://CRAN.R-project.org/package=janitor

Füllgrabe, C., & Rosen, S. (2016). On the (un)importance of working memory in speech-in-noise processing for listeners with normal hearing thresholds. *Frontiers in Psychology, 7*, 1268. https://doi.org/10.3389/fpsyg.2016.01268

Gordon-Salant, S., & Fitzgibbons, P. J. (1997). Selected cognitive factors and speech recognition performance among young and elderly listeners. *Journal of Speech, Language, and Hearing Research, 40*(2), 423–431. https://doi.org/10.1044/jslhr.4002.423

Heaton, R. K., Akshoomoff, N., Tulsky, D., Mungas, D., Weintraub, S., Dikmen, S., Beaumont, J., Casaletto, K. B., Conway, K., Slotkin, J., & Gershon, R. (2014). Reliability and validity of composite scores from the NIH toolbox cognition battery in adults. *Journal of the International Neuropsychological Society, 20*(6), 588–598. https://doi.org/10.1017/S1355617714000241

Heffner, C. C., & Myers, E. B. (2021). Individual differences in phonetic plasticity across native and nonnative contexts. *Journal of Speech, Language, and Hearing Research, 64*(10), 3720–3733. https://doi.org/10.1044/2021_JSLHR-21-00004

Hirsch, M. E., Lansford, K. L., Barrett, T. S., & Borrie, S. A. (2021). Generalized learning of dysarthric speech between male and female talkers. *Journal of Speech, Language, and Hearing Research, 64*(2), 444–451. https://doi.org/10.1044/2020_jslhr-20-00313

Ingvalson, E. M., Lansford, K. L., Fedorova, V., & Fernandez, G. (2017a). Cognitive factors as predictors of accented speech perception for younger and older adults. *The Journal of the Acoustical Society of America, 141*(6), 4652–4659. https://doi.org/10.1121/1.4986930

Ingvalson, E. M., Lansford, K. L., Fedorova, V., & Fernandez, G. (2017b). Receptive vocabulary, cognitive flexibility, and inhibitory control differentially predict older and younger adults'

success perceiving speech by talkers with dysarthria. *Journal of Speech, Language, and Hearing Research, 60*(12), 3632–3641. https://doi.org/10.1044/2017_JSLHR-H-17-0119

Janse, E., & Adank, P. (2012). Predicting foreign-accent adaptation in older adults. *Quarterly Journal of Experimental Psychology, 65*(8), 1563–1585. https://doi.org/10.1080/17470218.2012.658822

Kennedy-Higgins, D., Devlin, J. T., & Adank, P. (2020). Cognitive mechanisms underpinning successful perception of different speech distortions. *The Journal of the Acoustical Society of America, 147*(4), 2728–2740. https://doi.org/10.1121/10.0001160

Kleinschmidt, D. F., & Jaeger, T. F. (2015). Robust speech perception: Recognize the familiar, generalize to the similar, and adapt to the novel. *Psychological Review, 122*(2), 148–203. https://doi.org/10.1037/a0038695

Kuhn, M. (2022). *caret: Classification and regression training*. R package version 6.0–92, https://CRAN.R-project.org/package=caret

Lansford, K. L., Borrie, S. A., & Barrett, T. S. (2019). Regularity matters: Unpredictable speech degradation inhibits adaptation to dysarthric speech. *Journal of Speech, Language, and Hearing Research, 62*(12), 4282–4290. https://doi.org/10.1044/2019_JSLHR-19-00055

Lansford, K. L., Luhrsen, S., Ingvalson, E. M., & Borrie, S. A. (2018). Effects of familiarization on intelligibility of dysarthric speech in older adults with and without hearing loss. *American Journal of Speech-Language Pathology, 27*(1), 91–98. https://doi.org/10.1044/2017_AJSLP-17-0090

Lee, Y., Sung, J. E., & Sim, H. (2014). Effects of listeners' working memory and noise on speech intelligibility in dysarthria. *Clinical Linguistic & Phonetics, 28*(10), 785–795. https://doi.org/10.3109/02699206.2014.904443

Liss, J. M., Spitzer, S. M., Caviness, J. N., Adler, C., & Edwards, B. (1998). Syllabic strength and lexical boundary decisions in the perception of hypokinetic dysarthric speech. *The Journal of the Acoustical Society of America, 104*(4), 2457–2466. https://doi.org/10.1121/1.423753

Liss, J. M., Spitzer, S. M., Caviness, J. N., Adler, C., & Edwards, B. W. (2000). Lexical boundary error analysis in hypokinetic and ataxic dysarthria. *The Journal of the Acoustical Society of America, 107*(6), 3415–3424. https://doi.org/10.1121/1.429412

Liss, J. M., White, L., Mattys, S. L., Lansford, K., Lotto, A. J., Spitzer, S. M., & Caviness, J. N. (2009). Quantifying speech rhythm abnormalities in the dysarthrias. *Journal of Speech, Language, and Hearing Research, 52*(5), 1334–1352. https://doi.org/10.1044/1092-4388(2009/08-0208)

Lunner, T. (2003). Cognitive function in relation to hearing aid use. *International Journal of Audiology, 42*(Suppl. 1), S49–S58. https://doi.org/10.3109/14992020309074624

Lunner, T., & Sundewall-Thorén, E. (2007). Interactions between cognition, compression, and listening conditions: Effects on speech-in-noise performance in a two-channel hearing aid. *American Academy of Audiology, 18*(7), 604–617. https://doi.org/10.3766/jaaa.18.7.7

Mattys, S. L., Davis, M. H., Bradlow, A. R., & Scott, S. K. (2012). Speech recognition in adverse conditions: A review. *Language and Cognitive Processes, 27*(7–8), 953–978. https://doi.org/10.1080/01690965.2012.705006

Mattys, S. L., White, L., & Melhorn, J. F. (2005). Integration of multiple speech segmentation cues: A hierarchical framework. *Journal of Experimental Psychology: General, 134*(4), 477–500. https://doi.org/10.1037/0096-3445.134.4.477

McAuliffe, M. J., Gibson, E. M. R., Kerr, S. E., Anderson, T., & LaShell, P. J. (2013). Vocabulary influences older and younger listeners' processing of dysarthric speech. *The Journal of the Acoustical Society of America, 134*(2), 1358–1368. https://doi.org/10.1121/1.4812764

McAuliffe, M. J., Kerr, S. E., Gibson, E. M. R., Anderson, T., & LaShell, P. J. (2014). Cognitive-perceptual examination of remediation approaches to hypokinetic dysarthria. *Journal of Speech, Language, and Hearing Research, 57*(4), 1268–1283. https://doi.org/10.1044/2014_JSLHR-S-12-0349

McAuliffe, M. J., Wilding, P. J., Rickard, N. A., & O'Beirne, G. A. (2012). Effect of speaker age on speech recognition and perceived listening effort in older adults with hearing loss. *Journal of Speech, Language, and Hearing Research, 55*(3), 838–847. https://doi.org/10.1044/1092-4388(2011/11-0101)

Neger, T. M., Rietveld, T., & Janse, E. (2014). Relationship between perceptual learning in speech and statistical learning in younger and older adults. *Frontiers in Human Neuroscience, 8,* Article 628. https://doi.org/10.3389/fnhum.2014.00628

Ohlenforst, B., Zekveld, A. A., Jansma, E. P., Wang, Y., Naylor, G., Lorens, A., Lunner, T., & Kramer, S. E. (2017). Effects of hearing impairment and hearing aid amplification on listening effort: A systematic review. *Ear and Hearing, 38*(3), 267–281. https://doi.org/10.1097/AUD.0000000000000396

O'Neill, E. R., Kreft, H. A., & Oxenham, A. J. (2019). Cognitive factors contribute to speech perception in cochlear-implant users and age-matched normal-hearing listeners under vocoded conditions. *The Journal of the Acoustical Society of America, 146*(1), 195–210. https://doi.org/10.1121/1.5116009

Pichora-Fuller, M. K. (2003). Cognitive aging and auditory information processing. *International Journal of Audiology, 42*(Suppl. 2), 26–32. https://doi.org/10.3109/14992020309074641

Pichora-Fuller, M. K., Kramer, S. E., Eckert, M. A., Edwards, B., Hornsby, B. W. Y., Humes, L. E., Lemke, U., Lunner, T., Matthen, M., Mackersie, C. L., Naylor, G., Phillips, N. A., Richter, M., Rudner, M., Sommers, M. S., Tremblay, K. L., & Wingfield, A. (2016). Hearing impairment and cognitive energy: The Framework for Understanding Effortful Listening (FUEL). *Ear and Hearing, 37*(1), 5S–27S. https://doi.org/10.1097/AUD.0000000000000312

Powell, T. W. (2006). A comparison of English reading passages for elicitation of speech samples from clinical populations. *Clinical Linguistics & Phonetics, 20*(2–3), 91–97. https://doi.org/10.1080/02699200400026488

R Core Team. (2022). *R* (Version 4.2.0) [Computer software]. R Foundation. https://www.r-project.org/

Rönnberg, J. (2003). Cognition in the hearing impaired and deaf as a bridge between signal and dialogue: A framework and a model. *International Journal of Audiology, 42*(Suppl. 1), 68–76. https://doi.org/10.3109/14992020309074626

Rönnberg, J., Holmer, E., & Rudner, M. (2019). Cognitive hearing science and ease of language understanding. *International Journal of Audiology, 58*(5), 247–261. https://doi.org/10.1080/14992027.2018.1551631

Rudner, M., Foo, C., Rönnberg, J., & Lunner, T. (2009). Cognition and aided speech recognition in noise: Specific role for cognitive factors following nine-week experience with adjusted compression settings in hearing aids. *Scandinavian Journal of Psychology, 50*(5), 405–418. https://doi.org/10.1111/j.1467-9450.2009.00745.x

Rudner, M., Foo, C., Sundewall-Thorén, E., Lunner, T., & Rönnberg, J. (2008). Phonological mismatch and explicit cognitive processing in a sample of 102 hearing-aid users.

International Journal of Audiology, 47(Suppl. 2), S91–S98. https://doi.org/10.1080/14992020802304393

Rudner, M., & Rönnberg, J. (2008). Explicit processing demands reveal language modality-specific organization of working memory. *The Journal of Deaf Studies and Deaf Education, 13*(4), 466–484. https://doi.org/10.1093/deafed/enn005

Salthouse, T. A. (2004). What and when of cognitive aging. *Current Directions in Psychological Science, 13*(4), 140–144. https://doi.org/10.1111/j.0963-7214.2004.00293.x

Samuel, A. G., & Kraljic, T. (2009). Perceptual learning for speech. *Attention, Perception, & Psychophysics, 71*(6), 1207–1218. https://doi.org/10.3758/APP.71.6.1207

Scharenborg, O., Weber, A., & Janse, E. (2015). The role of attentional abilities in lexically guided perceptual learning by older listeners. *Attention, Perception, & Psychophysics, 77*(2), 493–507. https://doi.org/10.3758/s13414-014-0792-2

Sjoberg, D. D., Whiting, K., Curry, M., Lavery, J. A., & Larmarange, J. (2021). Reproducible summary tables with the gtsummary package. *The R Journal, 13*(1), 570–580. https://doi.org/10.32614/RJ-2021-053

Souza, P., & Arehart, K. (2015). Robust relationship between reading span and speech recognition in noise. *International Journal of Audiology, 54*(10), 705–713. https://doi.org/10.3109/14992027.2015.1043062

Tiego, J., Testa, R., Bellgrove, M. A., Pantelis, C., & Whittle, S. (2018). A hierarchical model of inhibitory control. *Frontiers in Psychology, 9*. Article 1339. https://doi.org/10.3389/fpsyg.2018.01339

Weintraub, S., Dikmen, S. S., Heaton, R. K., Tulsky, D. S., Zelazo, P. D., Bauer, P. J., Carlozzi, N. E., Slotkin, J., Blitz, D., Wallner-Allen, K., Fox, N. A., Beaumont, J. L., Mungas, D., Nowinski, C. J., Richler, J., Deocampo, J. A., Anderson, J. E., Manly, J. J., Borosh, B., . . . Gershon, R. C. (2013). Cognition assessment using the NIH toolbox. *Neurology, 80*(11, Suppl. 3), S54–S64. https://doi.org/10.1212/wnl.0b013e3182872ded

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

Wingfield, A., Tun, P. A., Koh, C. K., & Rosen, M. J. (1999). Regaining lost time: Adult aging and the effect of time restoration on recall of time-compressed speech. *Psychology and Aging, 14*(3), 380–389. https://doi.org/10.1037//0882-7974.14.3.380

Zecker, S. G., Hoffman, H. J., Frisina, R., Dubno, J. R., Dhar, S., Wallhagen, M., Kraus, N., Griffith, J. W., Walton, J. P., Eddins, D. A., Newman, C., Victorson, D., Warrier, C. M., & Wilson, R. H. (2013). Audition assessment using the NIH toolbox. *Neurology, 80*(11, Suppl. 3), S45–S48. https://doi.org/10.1212/WNL.0b013e3182872dd2