ASHA American Speech-Language-Hearing Association

JSLHR

Research Article

Beyond Speech Intelligibility: Quantifying Behavioral and Perceived Listening Effort in Response to Dysarthric Speech

Annalise R. Fletcher,[a] Alan A. Wisler,[b] Emily R. Gruver,[c] and Stephanie A. Borrie[a]

[a] Department of Communicative Disorders and Deaf Education, Utah State University, Logan [b] Department of Mathematics and Statistics, Utah State University, Logan [c] Department of Audiology and Speech-Language Pathology, University of North Texas, Denton

ARTICLE INFO

Article History:
Received March 1, 2022
Revision received May 17, 2022
Accepted July 8, 2022

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

https://doi.org/10.1044/2022_JSLHR-22-00136

ABSTRACT

Purpose: This study investigated whether listener processing of dysarthric speech requires the recruitment of more cognitive resources (i.e., higher levels of listening effort) than neurotypical speech. We also explored relationships between behavioral listening effort, perceived listening effort, and objective measures of word transcription accuracy.
Method: A word recall paradigm was used to index behavioral listening effort. The primary task involved word transcription, whereas a memory task involved recalling words from previous sentences. Nineteen listeners completed the paradigm twice, once while transcribing dysarthric speech and once while transcribing neurotypical speech. Perceived listening effort was rated using a visual analog scale.
Results: Results revealed significant effects of dysarthria on the likelihood of correct word recall, indicating that the transcription of dysarthric speech required higher levels of behavioral listening effort relative to neurotypical speech. There was also a significant relationship between transcription accuracy and measures of behavioral listening effort, such that listeners who were more accurate in understanding dysarthric speech exhibited smaller changes in word recall when listening to dysarthria. The subjective measure of perceived listening effort did not have a statistically significant correlation with measures of behavioral listening effort or transcription accuracy.
Conclusions: Results suggest that cognitive resources, particularly listeners' working memory capacity, are more taxed when deciphering dysarthric versus neurotypical speech. An increased demand on these resources may affect a listener's ability to remember aspects of their conversations with people with dysarthria, even when the speaker is fully intelligible.

People with dysarthria frequently report communication challenges when interacting with others in everyday situations. However, assessments of speech intelligibility demonstrate only weak to moderate relationships with reported communicative participation difficulties (Barnish et al., 2017; Borrie et al., 2022; Börjesson et al., 2021; Spencer et al., 2020; Yorkston et al., 2017). Thus, it is likely that additional aspects of the communication process, not captured by intelligibility, contribute to the communicative participation restrictions experienced by this population. One likely contributor is listener effort. Listeners often report increased mental effort and allocation of cognitive resources when listening to dysarthric speech (Beukelman et al., 2011; Connaghan et al., 2021; Whitehill & Wong, 2006). Furthermore, this perceived mental exertion cannot be fully explained by the level of the speaker's intelligibility impairment (Whitehill & Wong, 2006). Increased listener effort is likely to affect broader communicative situations (Peelle, 2018) and may be one reason why listeners reportedly avoid more challenging topics when talking to people with dysarthria (Brady et al., 2011).

Correspondence to Annalise R. Fletcher: Annalise.Fletcher@usu.edu.
Disclosure: *The authors have declared that no competing financial or nonfinancial interests existed at the time of publication.*

Copyright © 2022 American Speech-Language-Hearing Association

Downloaded from: https://pubs.asha.org Stephanie Borrie on 11/18/2022, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

The Framework for Understanding Effortful Listening highlights that listening effort is the deliberate allocation of mental resources to overcome obstacles in goal pursuit when carrying out a listening task (Pichora-Fuller et al., 2016). However, the term *listening effort* can be challenging to define due to the range of ways in which it has been described across the literature (Francis & Love, 2020; Herrmann & Johnsrude, 2020; Strand et al., 2020). For example, the term has been used to refer to listeners' subjective feelings of mental exertion, as well as the cognitive resources that are employed when parsing an incoming acoustic signal (Herrmann & Johnsrude, 2020). In addition, different assessments of listening effort are not necessarily well correlated; thus, it can be beneficial to consider multiple categories of listening effort to best understand its underlying dimensions (Alhanbali et al., 2019; Strand et al., 2020). In this article, consistent with Strand et al. (2020), we will use the term *behavioral listening effort* to represent objective changes in listener behavior that occur during or after a challenging listening task. Behavioral listening effort measures assume that an increased use of cognitive resources will result in quantifiable changes in listener behavior, including longer reaction times, reduced performance in secondary tasks, and a reduced ability to encode speech material into memory (Peelle, 2018). Consistent with other speech disorder literature, the term *perceived listening effort* will be used to refer to an individual's overall experience of mental exertion following the recruitment of these cognitive resources, a subjective construct that is commonly indexed through self-report measures (Nagle & Eadie, 2018).

The degree of listening effort employed in speech perception is dependent on task demands. As discussed in the Ease of Language Understanding (ELU) model, in optimal listening conditions, stored lexical representations are rapidly unlocked upon hearing a stream of phonological information (Rönnberg et al., 2008). In contrast, when listening conditions are suboptimal, immediate matches to stored lexical representations cannot be made, and speech processing requires explicit processing and storage capacity. Studies of listener effort have used various paradigms to create these suboptimal listening conditions. Commonly, studies have focused on speech in noise (e.g., Houben et al., 2013; Sarampalis et al., 2009; Wendt et al., 2018) or other situations in which hearing acuity is reduced (see Ohlenforst et al., 2017, for a review), but there has been little research of naturally degraded speech signals that result from speech disorders.

Naturally degraded speech signals vary in the challenges that they present. For example, phonemes are not necessarily missing or inaudible but are often imprecise and ambiguous. Phonetic ambiguity can cause the listener to rely more on word frequency and context for speech recognition (Norris & McQueen, 2008). Indeed, it has been posited that listeners might depend on preexisting lexical knowledge to a greater degree when processing dysarthric speech as compared to some types of speech in noise (Fletcher & McAuliffe, 2021). A change in speech processing strategies will likely result in different allocations of cognitive resources, which could affect measures of behavioral listening effort. Thus, further examination of listening effort in the context of naturally degraded speech signals is needed.

Broadly speaking, an individual's listening effort will depend on their availability of cognitive resources, as well as their attitudes and motivation, which determine how cognitive resources are allocated (Francis & Love, 2020; Herrmann & Johnsrude, 2020; Pichora-Fuller et al., 2016). However, unraveling these influences can be challenging, as highlighted by the lack of clear correlation between perceived and behavioral listening effort measures (e.g., Alhanbali et al., 2019; Lau et al., 2019; McGarrigle et al., 2021). It is possible that the cognitive resources used to complete a listening task may require different levels of subjective effort to recruit, depending on the individual listener. However, recent work has also suggested that listeners may struggle to accurately rate perceived effort because it is an abstract and difficult-to-quantify concept. For example, it is suggested that many listeners may automatically substitute an easier question when they are asked to rate perceived effort (e.g., "How well did I perform on the task?"; Moore & Picou, 2018). Despite these concerns, understanding perceived listening effort, in addition to behavioral listening effort, could provide important insight into communication barriers. For example, behavioral listening effort measures will likely be predictive of a listener's ability to manage multiple tasks and remember information in a conversation. In contrast, perceived listening effort might be more predictive of their choice to continue participating in a communicative interaction. Thus, an understanding of the relationship between these two measures could provide further insights into why listeners sometimes choose to disengage.

To date, the small body of speech disorder literature has mostly focused on understanding perceived listening effort rather than quantifying behavioral changes. Studies have consistently found that listeners report increased effort when exposed to disordered speech, and perceived listening effort tends to have a strong negative correlation with a speaker's overall speech intelligibility (e.g., Connaghan et al., 2021; Eadie et al., 2021; Landa et al., 2014; Nagle & Eadie, 2012, 2018; Panico & Healey, 2009; Whitehill & Wong, 2006). However, listeners can also experience increases in perceived effort while achieving perfect speech transcription accuracy (Nagle & Eadie, 2018).

It is unclear whether these findings are true of behavioral listening effort. Thus far, it has been reported that dysarthric speech requires increased reaction times to process (Cote-Reschny & Hodge, 2010; Fletcher et al., 2019) and may also cause greater autonomic arousal and attention (Farahani et al., 2020). Some studies have

Downloaded from: https://pubs.asha.org Stephanie Borrie on 11/18/2022, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

reported a positive relationship between the severity of speech symptoms and levels of behavioral listening effort (Farahani et al., 2020; Fletcher et al., 2019). However, these prior studies have not directly compared disordered and neurotypical speech patterns (Cote-Reschny & Hodge, 2010; Farahani et al., 2020) or have been unable to identify significant differences between neurotypical speech and mild levels of speech impairment (Fletcher et al., 2019).

This study explores whether processing dysarthric speech evokes greater levels of behavioral and perceived listening effort as compared to neurotypical speech. To measure behavioral listening effort, we utilize a word recall paradigm. In this paradigm, listeners are asked to both transcribe (primary task) and remember (secondary task) a set of words read by a speaker. A reduced ability to recall words from previous sentences suggests that the transcription task requires more cognitive resources, indicative of increased behavioral listening effort. To be most sensitive to the effects of increased listening effort, a secondary task should demand similar types of cognitive resources as the primary task (Guttentag, 1989). In this study, the word recall task utilizes listeners' working memory, as this cognitive resource has been consistently identified as important in the processing of dysarthric speech (e.g., Ingvalson et al., 2017; Lee et al., 2014). To measure perceived listening effort, we use a visual analog rating scale (Nagle & Eadie, 2018; Whitehill & Wong, 2006).

This study addresses the following three key research questions: (a) Does deciphering dysarthric speech demand more behavioral listening effort than deciphering neurotypical speech? (b) Does deciphering dysarthric speech demand more perceived listening effort than deciphering neurotypical speech? (c) Are there relationships between behavioral listening effort, perceived listening effort, and ability to decipher dysarthric speech? We hypothesize that both perceived and behavioral listening effort will be sensitive to the effects of dysarthria, with increased use of cognitive resources to decipher dysarthric speech relative to neurotypical speech. Furthermore, individual listener differences are anticipated, including evidence of greater perceived and behavioral listening effort from listeners who are less accurate in deciphering dysarthric speech.

Method

Participants

This study received institutional review board approval from the University of North Texas. Twenty-one speakers of American English initially completed the listening tasks. Data from two participants were later removed due to disruptions during the listening task and/or a failure to follow task directions. The remaining participants consisted of three men and 16 women between the ages of 19 and 53 years (M = 26.3, SD = 8.7). Due to COVID-19 protocols, listeners' hearing was not directly screened prior to participation. However, all listeners reported that they had normal hearing with no known history of speech, language, or hearing difficulties.

Speech Stimuli

Speech stimuli consisted of audio-recorded productions of 100 Revised Speech Perception in Noise (SPIN-R) sentences (Bilger, 1994), with 50 sentences elicited from a female speaker with moderate spastic dysarthria (aged 63 years) and 50 sentences elicited from a female speaker with no neurological history or speech impairment (aged 66 years). Sentence lists for each speaker were unique, but sentences were of a similar length, and the final word was always a monosyllabic noun with a mid-range word frequency (see Elliott et al., 1995, for details). Sentences were also balanced for predictability, with each set containing an equal number of highly predictable final words and unpredictable final words, as described in Kalikow et al. (1977). Both talkers were speakers of Western American English. The speech stimuli were digitally recorded in a quiet room, at 48 kHz with 16 bits of quantization.

Procedure

Task Order

Listeners were asked to complete a word recall paradigm, generating a measure of speech intelligibility (words correct) and a metric reflecting behavioral listening effort (difference in the number of words recalled). In addition, listeners rated speech for a measure of perceived listening effort (subjective ratings). For these tasks, participants were seated alone in a quiet room and presented with sentences via a custom-designed MATLAB program. Audio was presented using Sennheiser HD 598 closed-back headphones. Participants were given the opportunity to adjust the volume of the speech signal to whatever level they desired prior to beginning the listening tasks, and this volume was then held constant throughout all tasks. In the process of adjusting the audio, listeners were exposed to an example sentence from each speaker (these sentences were kept constant and were not repeated in the experimental tasks). Listeners then completed both the word recall paradigm and the effort rating task with stimuli from one speaker before repeating each procedure with stimuli from the second speaker (the order in which speakers were presented was counterbalanced across listeners). The word recall paradigm always preceded the effort rating task. Listeners were offered a short break between speakers and tasks, if desired. Together, the tasks took an average of 30 min to complete.

Downloaded from: https://pubs.asha.org Stephanie Borrie on 11/18/2022, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

Word Recall Procedure

The procedures used in the word recall experiment closely follow the methods described in Sarampalis et al. (2009) and Pichora-Fuller et al. (1995), in which listeners are required to transcribe audio sentence productions while simultaneously holding words from previous sentences in their working memory for subsequent retrieval. Forty-eight SPIN-R sentences from each speaker were used in the main experiment (with an equal number of predictable and unpredictable final words). The presentation order of the 48 sentences was randomized to create a unique order of sentences for every listener in the study. This procedure ensured that there were no systematic differences in the order of unpredictable versus predictable sentences across the two listening conditions. Each sentence was played only once. After each sentence was presented, the listeners were prompted to type the last word they thought they heard into the experiment's graphical user interface. Listeners were encouraged to always make a guess if they were uncertain of the final word, and consequently, there were no missing or blank responses. The participants were also informed that they would need to remember and recall their responses later. After every eight sentences, listeners were cued to type as many of the previously reported words as they could from that block of eight. Listeners were told they could type the words in any order and were able to move to the next sentence at their own pace. Six sets of eight sentences were played from each speaker.

Effort Rating Task

Following the word recall procedure, listeners were asked to provide an effort rating scale judgment of six sentence stimuli from each speaker. Listeners provided one rating following each sentence. These sentences were the same for each listener, but the order of presentation was randomized. Listeners were asked to rate "How easy is this speech to understand?" based on previous publications that have examined listener effort in dysarthria with a focus on "ease of understanding" (e.g., Landa et al., 2014; Stipancic et al., 2021). This study used a visual analog with the words "easy" and "difficult" printed at each end. The language used was slightly adapted from the Landa et al. (2014) study, to focus more on the degree of perceived difficulty without references to overall accuracy or performance in the transcription task. To assess interrater reliability, intraclass correlations (ICCs) were calculated (as described in Sheard et al., 1991). The obtained ICC(2,*k*) coefficients were 0.984, 95% CI [0.967, 0.995], indicating that excellent levels of listener reliability were obtained with this prompt.

Data Analysis

Speech Intelligibility Scoring

Following data collection, transcription accuracy was automatically calculated using the open-source tool, Autoscore, which compares the word spoken by the talker to the word that was transcribed by the listener (http://autoscore.usu.edu/; Borrie et al., 2019). Listener responses were considered correct when they were identical to the word spoken by the talker. Additional Autoscore scoring rules including the acceptable spell, tense, and plural rules were also applied, as per previous studies examining perception of dysarthric speech (e.g., Borrie et al., 2021; Lansford et al., 2019). Briefly, these rules allow for words to be considered correct if they were common homophones or a misspelling of the target word, if they differed from the target word only by the addition or omission of *–d* or *–ed* or if they differed from the target word only by the addition or omission of *–s* or *–es*. Following the use of Autoscore, a research assistant manually screened the files for any common spelling errors that were not detected by Autoscore (e.g., "theif" vs. "thief"). If an incorrect word was clearly attributable to a spelling mistake, the response was manually recoded as correct. The same scoring system was also used to compare the string of eight words typed by listeners across trials (i.e., the target) to the string of words they listed when prompted to recall their earlier responses (i.e., the response). In this case, word order did not affect their score.

Results

Speech Intelligibility

A mixed-effects logistic regression model was used to assess the effect of dysarthria on the likelihood of accurate word identification. The dependent variable was a binary measurement of whether a given word was correctly identified by a given listener, and a random intercept for listener was included in the model to account for repeated measures from the same listener. In addition, since word predictability can have a strong influence on both accurate word identification and listening effort (Hunter, 2021), we included a binary factor for final-word predictability (i.e., predictable vs. unpredictable) to quantify and control for this variable. Block number of the stimuli was also included as a main effect, to examine any learning or familiarization effects that might occur during the experiment. The model revealed the expected significant effects of dysarthria on the likelihood of correct word identification ($\beta = 3.600$, $SE = 0.419$, $p < .001$). This effect size indicates that the odds of a listener correctly identifying a word were 36.6 times higher when deciphering neurotypical speech, as opposed to dysarthric speech. There was also a statistically significant effect of final-word predictability, demonstrating that the predictable final words were more likely than the unpredictable final words to be accurately transcribed ($\beta = 1.583$, $SE = 0.209$, $p < .001$).

Downloaded from: https://pubs.asha.org Stephanie Borrie on 11/18/2022, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

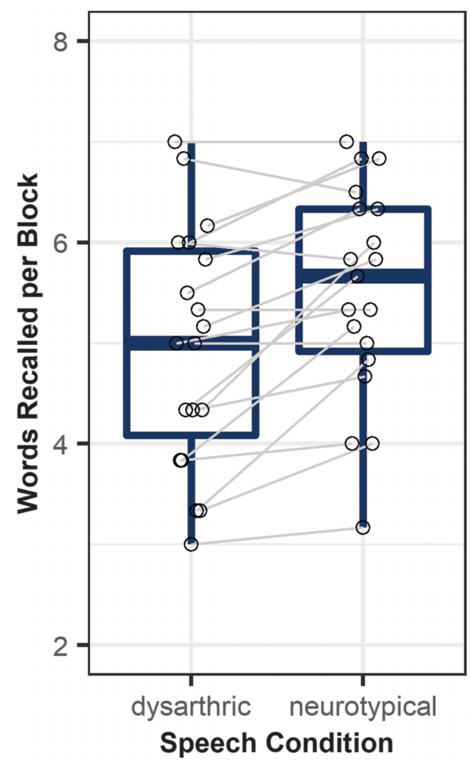
This corresponds to predictable words having 4.87 times higher odds of being transcribed correctly. There was no statistically significant effect of stimuli block on word identification accuracy (β = 0.003, *SE* = 0.049, *p* = .950). On average, across all listeners, speech intelligibility was 82% and 99% for the speaker with dysarthria and neurotypical speaker, respectively. Thus, as anticipated, the dysarthric speech was less intelligible than the neurotypical speech.

Behavioral Listening Effort

Our first research question asked whether deciphering dysarthric speech demanded more behavioral listening effort than deciphering neurotypical speech. Behavioral listening effort was assessed based on the proportion of words that were accurately recalled by the listener. As described in the Method section, words were considered accurately recalled based on whether they matched what was previously transcribed by the listener (i.e., a word could be considered correctly recalled even if the listener originally misidentified the word). A mixed-effects logistic regression model was used to assess the effect of dysarthric speech, relative to neurotypical speech, on the likelihood of accurate word recall. Stimuli block and final-word predictability were also added to the model as fixed effects. A random intercept was included for each listener. Results of the model revealed significant effects of dysarthria on the likelihood of correct word recall (β = 0.336, *SE* = 0.103, *p* = .001). This effect size indicates that the statistical odds of a listener correctly recalling a word were 1.40 times higher when listening to the neurotypical speaker, as opposed to the speaker with dysarthria. There was also a statistically significant effect of final-word predictability, demonstrating that the predictable final words were more likely than the unpredictable final words to be accurately recalled (β = 0.278, *SE* = 0.103, *p* = .007). There was no statistically significant effect of stimuli block on word recall (β = −0.029, *SE* = 0.030, *p* = .333). On average, listeners were able to recall 4.96 of eight (62%) words produced by the speaker with dysarthria and 5.51 of eight (69%) words spoken by the neurotypical speaker (see Figure 1). Thus, deciphering dysarthric speech affected performance in the secondary working memory task, suggesting that the task demands more behavioral listening effort than deciphering neurotypical speech.

To assess whether differences in word recall of dysarthric versus neurotypical speech persisted when stimuli were fully intelligible, we created one further model. This model focused on the subset of words that were accurately identified (i.e., 100% correct across trials) in both speech types. Again, a mixed-effects logistic regression model was used to assess the effect of dysarthria on word recall, and a random intercept for listener was included. Fixed effects of final-word predictability and stimuli block were also included. Results of the model revealed that there was still a significant effect of dysarthria on the likelihood of correct word recall (β = 0.336, *SE* = 0.108, *p* = .002). This effect size indicates that the odds of a listener correctly recalling a fully intelligible word were still 1.40 times higher when listening to the neurotypical speaker, as opposed to the speaker with dysarthria. The effect of final-word predictability was also still present (β = 0.325, *SE* = 0.108, *p* = .003), while stimuli block continued to have no statistically significant effect on the likelihood of word recall (β = −0.023, *SE* = 0.032, *p* = .465). Thus, this model highlighted that the presence of dysarthria affected word recall even for intelligible words.

Figure 1. Average number of words each listener recalled per block, reflecting listener behavioral listening effort. Gray lines show differences in each listener's recall across the two listening conditions.



Perceived Listening Effort

Our second research question asked whether deciphering dysarthric speech demanded more perceived listening effort than deciphering neurotypical speech. A mixed-effects linear regression was used to model ratings provided along the visual analog scale. We examined the fixed effect of dysarthric speech (vs. neurotypical speech), as well as final-word predictability on rating scores. Since only six sentences per speaker were included in the assessment of perceived listening effort, we did not include stimuli block as a variable. However, a random intercept for listener was added, to account for repeated ratings provided by the

Downloaded from: https://pubs.asha.org Stephanie Borrie on 11/18/2022, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

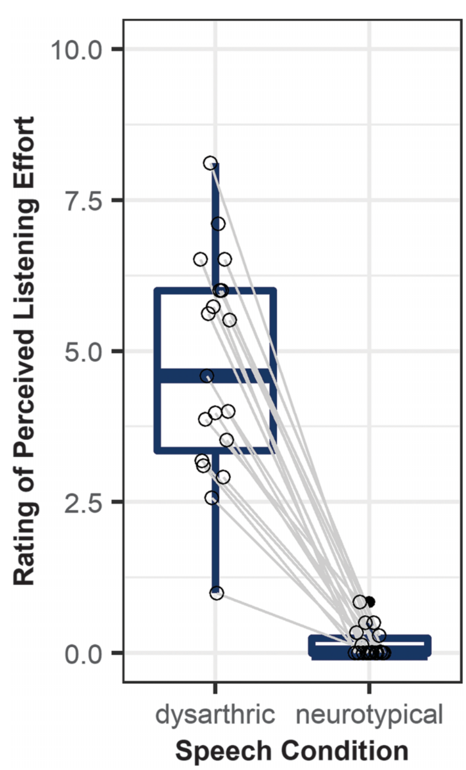
same listener. The speaker with dysarthria was perceived as being significantly more difficult to understand (β = 4.520, *SE* = 0.250, $p < .001$). This result indicated that listeners' subjective experiences of effort were, on average, 4.5 points (out of 10) further along the perceived effort scale when deciphering dysarthric speech. Figure 2 displays the relatively large differences in ratings of the speakers along the visual analog scale. Note that the ends of the scale ranged from 0 (minimum effort) to 10 (maximum effort). The effect of final-word predictability was also statistically significant, indicating that sentences with predictable final words were also perceived as easier to understand than sentences with unpredictable final words (β = 0.742, *SE* = 0.263, $p = .005$).

Behavioral Listening Effort, Perceived Listening Effort, and Transcription Accuracy

Our final research question sought to better understand the relationships between measures of a listener's behavioral listening effort, perceived listening effort, and ability to decipher dysarthric speech. To assess this, behavioral listening effort was calculated as the average difference in word recall across speech conditions, for each listener. Changes in perceived listening effort and the ability to decipher dysarthric speech were based on the average differences in transcription accuracy and effort ratings provided by each listener across speech conditions. Since there were 19 listeners in this data set, these procedures resulted in 19 measurements representing each listener's change in behavioral listening effort, 19 measurements representing changes in perceived listening effort, and 19 changes in transcription accuracy (i.e., ability to decipher dysarthric speech).

A Pearson correlation coefficient demonstrated that there was no statistically significant relationship between changes in perceived listening effort and changes in behavioral listening effort, $r(17) = .174$, $p = .475$. There was also no statistically significant relationship between changes in a listener's transcription accuracy and changes in perceived effort ratings, $r(17) = .019$, $p = .938$. However, changes in transcription accuracy were significantly related to changes in listeners' behavioral listening effort, such that listeners who were more accurate in understanding dysarthric speech had better recall for that speech, $r(17) = −.489$, $p = .033$, indicating fewer cognitive resources were required. These correlations are displayed in Figures 3, 4, and 5.

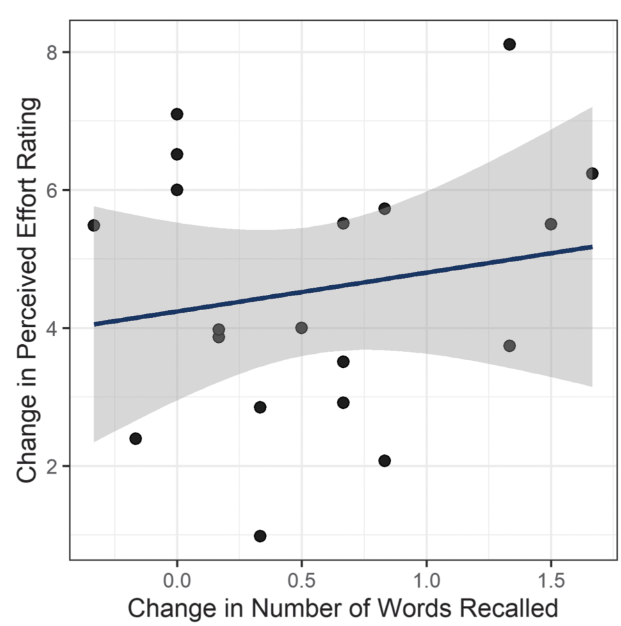
Figure 2. Average rating of listener effort provided by each listener. Gray lines show differences in a listener's ratings across the two listening conditions. The rating scale was continuous with end points equal to 0 and 10.



Discussion

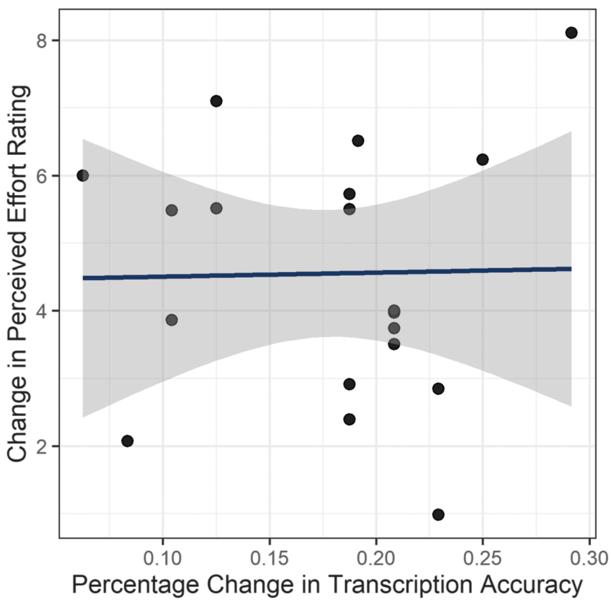
The results of this study provide evidence that deciphering dysarthric speech increases both listener behavioral listening effort and perceived listening effort relative to deciphering neurotypical speech and that these two constructs

Figure 3. Correlation between changes in a listener's perceived effort ratings and changes in their word recall across the two speech conditions.



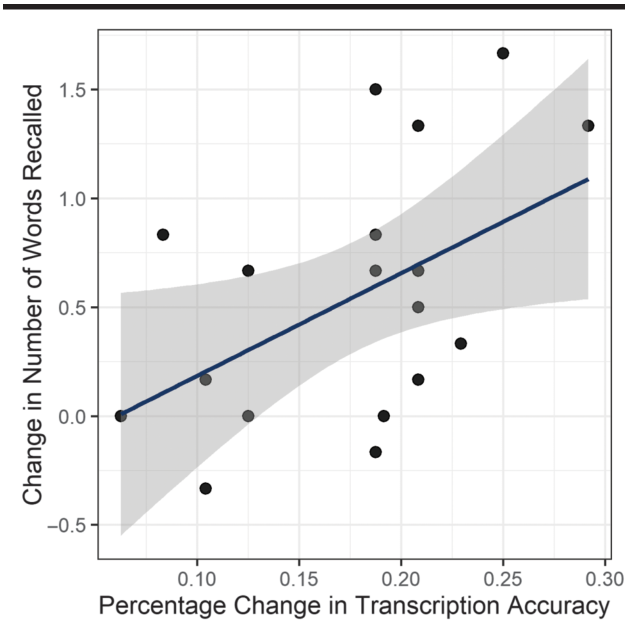
Downloaded from: https://pubs.asha.org Stephanie Borrie on 11/18/2022, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

Figure 4. Correlation between changes in a listener's transcription accuracy and changes in their perceived listening effort across the two speech conditions.



are distinct. Increased behavioral listening effort indicates that greater levels of cognitive resources are required during speech perception, whereas changes in perceived effort reflect listeners' personal experiences of mental exertion. Few studies have simultaneously examined these constructs in the perception of speakers with dysarthria (Cote-Reschny & Hodge, 2010; Farahani et al., 2020).

Figure 5. Correlation between changes in a listener's transcription accuracy and changes in their word recall across the two speech conditions.



This study employed a word recall paradigm to index behavioral listening effort. When transcribing dysarthric speech, the odds of listeners recalling a word 1.40 times lower than when transcribing neurotypical speech. Furthermore, these word recall differences between dysarthric and neurotypical speech were not solely attributable to difficulties remembering unintelligible words. When considering trials that contained only correctly transcribed words, the odds of a listener recalling a word were still 1.40 times lower after transcribing the speaker with dysarthria. This suggests that more cognitive resources are being allocated to the task of listening to and transcribing dysarthric speech, even when the speech is intelligible.

Although few studies have examined listeners' behavioral listening effort while deciphering dysarthric speech, prior research has suggested that behavioral listening effort is increased when listening to people with dysarthria. For example, Fletcher et al. (2019) found that listeners had longer reaction times when responding to statements from people with dysarthria, as compared to identical statements from neurotypical speakers. Farahani et al. (2020) examined listener responses to speech produced by individuals with spasmodic dysphonia and found evidence of greater peak pupil dilation when listening to more strained voice qualities, indicating increased attention and arousal (Francis & Love, 2020; Strand et al., 2020), though Farahani et al. did not directly compare these results to neurotypical speech patterns. Hence, this study goes one step further in establishing a statistically significant increase in behavioral listening effort when deciphering dysarthric speech, as compared to neurotypical speech patterns. Furthermore, this study provides evidence that differences in behavioral listening effort occur when speech stimuli are fully intelligible, a finding that was not previously established in Fletcher et al.

The measure of behavioral listening effort used in this study taps into the listener's working memory resources when deciphering dysarthric speech. Although many cognitive resources are thought to be recruited during speech perception, working memory resources are particularly important when speech is degraded. As described in the ELU model, when the acoustic signal is unclear, the string of acoustic information cannot be rapidly and automatically matched with a specific item in the listener's lexicon, so the listener must store acoustic details while integrating broader contextual information (Rönnberg et al., 2008). Working memory capacity is thought to directly modulate a listener's ability to apply this "top-down" linguistic knowledge to the process of lexical selection (Janse & Jesse, 2014; Wingfield et al., 1994). Prior studies have demonstrated that listeners with larger working memories tend to be more successful in transcribing

Downloaded from: https://pubs.asha.org Stephanie Borrie on 11/18/2022, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

dysarthric speech (e.g., Ingvalson et al., 2017; Lee et al., 2014). There is also evidence that working memory capacity is more closely linked to listener comprehension when speech is naturally altered (i.e., in the case of accented speech) rather than masked by noise, which suggests that listeners' cognitive resources are allocated in different ways depending on how the speech signal is altered (Francis et al., 2021). The results of this study provide further evidence of the specific role of working memory in understanding dysarthric speech by measuring the degree to which working memory resources are taxed during speech transcription.

A reduced working memory capacity may result in additional challenges for listeners during communicative interactions. To interpret complex messages and build a connection to a speaker, the ability to recall details from conversations is important. Findings from this study show that word recall is negatively affected by dysarthric speech, even when words are fully intelligible. This may lead to challenges remembering instructions, names, or broader contextual information provided by speakers with dysarthria. Prior research has found weak relationships between the accurate transcription of dysarthric speech and the ability of listeners to later interpret the meaning of these messages (Hustad, 2008). Thus, it is suggested that broader comprehension of dysarthric speech is multifaceted and not solely related to a listener's ability to accurately identify spoken words. We postulate that difficulties with word recall may be an important factor that contributes to some of these broader listener comprehension difficulties.

This study also found evidence of increased effort ratings when listening to dysarthric speech relative to neurotypical speech. When asked to rate listening effort, participants rated dysarthric speech an average of 4.5 points higher than neurotypical speech along a 10-point scale, indicating a large difference between the speech conditions. This result aligns with prior research, which has consistently found that listeners report increased effort when exposed to disordered speech (e.g., Eadie et al., 2021; Landa et al., 2014; Nagle & Eadie, 2012, 2018; Panico & Healey, 2009; Whitehill & Wong, 2006).

A strong effect of dysarthria on perceived listening effort is not surprising when considering the cognitive–perceptual processes involved in deciphering dysarthria. As discussed earlier, when processing unclear speech patterns, listeners may be forced to rely more on "top-down" linguistic knowledge to decode the acoustic signal. Although working memory plays an important role in this process, other factors like information-processing speed and multiple cognitive resources related to attention, inhibitory control, and long-term memory retrieval are likely to be involved in the storage and retrieval of lexical knowledge (Eckert et al., 2008; Peelle, 2018). Thus, perceived effort may offer broader insight into the listener's experience of deciphering dysarthric speech, as opposed to the allocation of a single cognitive resource, like working memory.

We hypothesized that an individual's increases in perceived listening effort would be correlated with changes in their behavioral listening effort when transcribing dysarthria. However, when examining changes in perceived effort ratings across listeners, there was not a statistically significant relationship between increases in perceived effort (in response to hearing dysarthric speech) and increases in the listener's behavioral listening effort when transcribing dysarthric speech. There are several reasons why differences in perceived listening effort and behavioral listening effort might not be strongly correlated. As discussed in the previous paragraph, the full range of cognitive resources required to understand degraded speech is unlikely to be perfectly indexed through any single behavioral listening effort measurement. Additionally, the resources used by listeners may require different levels of subjective effort to recruit, depending on the individual (Herrmann & Johnsrude, 2020). Indeed, the subjectivity of the effort rating task may also result in values that more closely resemble a listener's broader motivation and attitude toward people with speech disorders (Connaghan et al., 2021). For example, in cases where people exhibit equal levels of behavioral listening effort, a higher perceived effort level from one listener may indicate that the interaction is viewed less favorably, and the listener is less willing to continue participating. Thus far, little is known about the consequences of perceived listening effort on communicative participation. However, this is an important line of inquiry, and future studies should investigate such relationships by incorporating broader measurements of listeners' attitudes, motivation, and willingness to participate in interactions with speakers with dysarthria.

Increases in perceived listening effort when hearing dysarthric speech did not have a statistically significant correlation with reductions in the listeners' ability to decipher dysarthric speech. Several prior studies have found relationships between the intelligibility of people with speech disorders and the perceived listening effort experienced by listeners (e.g., Connaghan et al., 2021; Landa et al., 2014; Nagle & Eadie, 2018; Whitehill & Wong, 2006). It should be noted, however, that these studies focused on correlations between the intelligibility of a particular speaker, or speech sample, and the corresponding effort ratings given to that speaker or stimuli. In these cases, as the speech sample becomes less intelligible, the amount of perceived effort required to understand the sample tends to increase. In contrast, our study focuses on comparing how different listeners respond to the same set of speech samples. When exploring these interlistener

Downloaded from: https://pubs.asha.org Stephanie Borrie on 11/18/2022, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

variations in perceived effort, the results demonstrate that increases in self-reported effort do not have a statistically significant correlation with reductions in a person's ability to decipher speech relative to other listeners. In other words, when listeners were asked to rate their effort, or their difficulty understanding a speech sample, increased effort in response to dysarthria was not correlated with how "well" they performed the task relative to other listeners (i.e., their reduction in transcription accuracy in response to dysarthric speech). Thus, even when a listener becomes skilled in understanding dysarthric speech, they may still perceive the speech as being just as effortful to understand (at least in cases where they do not know how well they are performing relative to others). This idea is also supported by the results in Connaghan et al. (2021), which found no difference in the effort ratings of speech-language pathologists and less experienced listeners when listening to dysarthric speech.

In contrast to perceived effort, a person's level of behavioral listening effort was significantly correlated with their ability to decipher dysarthric speech. Listeners who experienced larger declines in their word recall performance tended to also exhibit greater reductions in their transcription of dysarthric speech. This pattern suggests that certain listeners are more cognitively challenged by the process of deciphering dysarthric speech, resulting in a greater recruitment of cognitive resources and reduced performance in the primary speech transcription task. This finding does not necessarily mean that increased behavioral listening effort is an undesirable behavior. If the recruitment of cognitive resources was reduced among these listeners, performance in the speech transcription task would likely be more negatively affected. Nevertheless, the results do raise questions concerning whether this level of behavioral listening effort is sustainable. In prior research, we have found evidence of increased behavioral listening effort even for speakers with very low levels of intelligibility (Fletcher et al., 2019). However, over time, applying increased cognitive resources to a speech signal and achieving limited success in deciphering that signal may cause fatigue or frustration. For this reason, it has been speculated that behavioral listening effort may reduce if a task is perceived to be "too difficult" to complete (Herrmann & Johnsrude, 2020). Evidence of listener disengagement at low levels of intelligibility has been found in some speech-in-noise literature (Wendt et al., 2018). Thus, identifying whether behavioral listening effort in response to disordered speech becomes negatively affected by low levels of motivation is an important avenue for future research.

We acknowledge that this study is limited in the conclusions it can draw from a single speaker with dysarthria. To better interpret how behavioral listening effort is affected by dysarthric speech, speech samples from more speakers should be considered, including a wider spectrum of dysarthria severities. It is also possible that different disordered-speech features may tax listeners' cognitive resources in different ways (i.e., a slower rate of speech may affect how a listener stores acoustic features in working memory), so broader representations of dysarthria are also warranted. In future studies, it will be important to establish what levels of listener behavioral listening effort are sustainable for communicative partners, to enable them to understand disordered speech without excessive cognitive fatigue. To accomplish this, future research should consider a broader variety of listening tasks and behavioral listening effort measures, to better reflect the task demands of everyday communicative interactions.

Summary and Conclusions

Deciphering dysarthric speech can result in increased perceived effort and behavioral listening effort, even in cases where the speech signal is relatively intelligible. However, the simultaneous examination of behavioral listening effort and perceived effort highlights that there are differences between the two constructs. Measurements of behavioral listening effort appear to be more closely related to a listener's ability to decipher dysarthric speech than to perceived listening effort. It is hypothesized that measurements of behavioral listening effort provide unique insight into communication difficulties encountered by listeners, including difficulties retaining and recalling aspects of their communicative interaction. Whereas ratings of perceived listening effort may be more sensitive to differences in a listener's attitude toward the speaker and the listening task. Ultimately, communicative participation relies on contributions from both the speaker and the listener. Thus, further exploration of behavioral listening effort and perceived effort is needed to understand the participation barriers faced by speakers with dysarthria.

Data Availability Statement

Anonymized listener data, analysis code, and model outputs associated with this work are available on request from the authors.

Acknowledgments

This research was supported by National Institute on Deafness and Other Communication Disorders Grant R21DC018867, awarded to Stephanie A. Borrie.

Downloaded from: https://pubs.asha.org Stephanie Borrie on 11/18/2022, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

References

Alhanbali, S., Dawes, P., Millman, R. E., & Munro, K. J. (2019). Measures of listening effort are multidimensional. *Ear and Hearing, 40*(5), 1084–1097. https://doi.org/10.1097/AUD.0000000000000697

Barnish, M. S., Horton, S. M., Butterfint, Z. R., Clark, A. B., Atkinson, R. A., & Deane, K. H. (2017). Speech and communication in Parkinson's disease: A cross-sectional exploratory study in the UK. *BMJ Open, 7*(5), Article e014642. https://doi.org/10.1136/bmjopen-2016-014642

Beukelman, D. R., Childes, J., Carrell, T., Funk, T., Ball, L. J., & Pattee, G. L. (2011). Perceived attention allocation of listeners who transcribe the speech of speakers with amyotrophic lateral sclerosis. *Speech Communication, 53*(6), 801–806. https://doi.org/10.1016/j.specom.2010.12.005

Bilger, R. C. (1994). *Authorized Version Revised Spin Test (Revised Speech Perception in Noise Test)*. The University of Illinois Press.

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 the Acoustical Society of America, 145*(1), 392–399. https://doi.org/10.1121/1.5087276

Borrie, S. A., Lansford, K. L., & Barrett, T. S. (2021). A clinical advantage: Experience informs recognition and adaptation to a novel talker with dysarthria. *Journal of Speech, Language, and Hearing Research, 64*(5), 1503–1514. https://doi.org/10.1044/2021_JSLHR-20-00663

Borrie, S. A., Wynn, C. J., Berisha, V., & Barrett, T. S. (2022). From speech acoustics to communicative participation in dysarthria: Toward a causal framework. *Journal of Speech, Language, and Hearing Research, 65*(2), 405–418. https://doi.org/10.1044/2021_JSLHR-21-00306

Börjesson, M. S., Hartelius, L., & Laakso, K. (2021). Communicative participation in people with amyotrophic lateral sclerosis. *Folia Phoniatrica et Logopaedica, 73*(2), 101–108. https://doi.org/10.1159/000505022

Brady, M. C., Clark, A. M., Dickson, S., Paton, G., & Barbour, R. S. (2011). The impact of stroke-related dysarthria on social participation and implications for rehabilitation. *Disability and Rehabilitation, 33*(3), 178–186. https://doi.org/10.3109/09638288.2010.517897

Connaghan, K. P., Wertheim, C., Laures-Gore, J. S., Russell, S., & Patel, R. (2021). An exploratory study of student, speech-language pathologist and emergency worker impressions of speakers with dysarthria. *International Journal of Speech-Language Pathology, 23*(3), 265–274. https://doi.org/10.1080/17549507.2020.1768286

Cote-Reschny, K. J., & Hodge, M. M. (2010). Listener effort and response time when transcribing words spoken by children with dysarthria. *Journal of Medical Speech-Language Pathology, 18*(4), 24–35.

Eadie, T. L., Durr, H., Sauder, C., Nagle, K., Kapsner-Smith, M., & Spencer, K. A. (2021). Effect of noise on speech intelligibility and perceived listening effort in head and neck cancer. *American Journal of Speech-Language Pathology, 30*(3S), 1329–1342. https://doi.org/10.1044/2020_AJSLP-20-00149

Eckert, M. A., Walczak, A., Ahlstrom, J., Denslow, S., Horwitz, A., & Dubno, J. R. (2008). Age-related effects on word recognition: Reliance on cognitive control systems with structural declines in speech-responsive cortex. *Journal of the Association for Research in Otolaryngology, 9*(2), 252–259. https://doi.org/10.1007/s10162-008-0113-3

Elliott, L. L. (1995). Verbal auditory closure and the Speech Perception in Noise (SPIN) Test. *Journal of Speech and Hearing Research, 38*(6), 1363–1376. https://doi.org/10.1044/jshr.3806.1363

Farahani, M., Parsa, V., Herrmann, B., Kadem, M., Johnsrude, I., & Doyle, P. C. (2020). An auditory-perceptual and pupillometric study of vocal strain and listening effort in adductor spasmodic dysphonia. *Applied Sciences, 10*(17), 5907. https://doi.org/10.3390/app10175907

Fletcher, A., & McAuliffe, M. (2021). Comparing lexical cues in listener processing of dysarthria and speech in noise. *American Journal of Speech-Language Pathology, 30*(3S), 1572–1579. https://doi.org/10.1044/2020_AJSLP-20-00137

Fletcher, A. R., Risi, R. A., Wisler, A., & McAuliffe, M. J. (2019). Examining listener reaction time in the perceptual assessment of dysarthria. *Folia Phoniatrica et Logopaedica, 71*(5–6), 297–308. https://doi.org/10.1159/000499752

Francis, A. L., Bent, T., Schumaker, J., Love, J., & Silbert, N. (2021). Listener characteristics differentially affect self-reported and physiological measures of effort associated with two challenging listening conditions. *Attention, Perception, & Psychophysics, 83*(4), 1818–1841. https://doi.org/10.3758/s13414-020-02195-9

Francis, A. L., & Love, J. (2020). Listening effort: Are we measuring cognition or affect, or both? *Wiley Interdisciplinary Reviews: Cognitive Science, 11*(1), Article e1514. https://doi.org/10.1002/wcs.1514

Guttentag, R. E. (1989). Age differences in dual-task performance: Procedures, assumptions, and results. *Developmental Review, 9*(2), 146–170. https://doi.org/10.1016/0273-2297(89)90027-0

Herrmann, B., & Johnsrude, I. S. (2020). A model of listening engagement (MoLE). *Hearing Research, 397,* 108016. https://doi.org/10.1016/j.heares.2020.108016

Houben, R., van Doorn-Bierman, M., & Dreschler, W. A. (2013). Using response time to speech as a measure for listening effort. *International Journal of Audiology, 52*(11), 753–761. https://doi.org/10.3109/14992027.2013.832415.

Hunter, C. R. (2021). Dual-task accuracy and response time index effects of spoken sentence predictability and cognitive load on listening effort. *Trends in Hearing, 25*. https://doi.org/10.1177/23312165211018092

Hustad, K. C. (2008). The relationship between listener comprehension and intelligibility scores for speakers with dysarthria. *Journal of Speech, Language, and Hearing Research, 51*(3), 562–573. https://doi.org/10.1044/1092-4388(2008/040)

Ingvalson, E. M., Lansford, K. L., Fedorova, V., & Fernandez, G. (2017). 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., & Jesse, A. (2014). Working memory affects older adults' use of context in spoken-word recognition. *Quarterly Journal of Experimental Psychology, 67*(9), 1842–1862. https://doi.org/10.1080/17470218.2013.879391

Kalikow, D. N., Stevens, K. N., & Elliott, L. L. (1977). Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. *The Journal of the Acoustical Society of America, 61*(5), 1337–1351. https://doi.org/10.1121/1.381436

Landa, S., Pennington, L., Miller, N., Robson, S., Thompson, V., & Steen, N. (2014). Association between objective measurement of the speech intelligibility of young people with dysarthria and listener ratings of ease of understanding.

Downloaded from: https://pubs.asha.org Stephanie Borrie on 11/18/2022, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

International Journal of Speech-Language Pathology, 16(4), 408–416. https://doi.org/10.3109/17549507.2014.927922
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
Lau, M. K., Hicks, C., Kroll, T., & Zupancic, S. (2019). Effect of auditory task type on physiological and subjective measures of listening effort in individuals with normal hearing. *Journal of Speech, Language, and Hearing Research, 62*(5), 1549–1560. https://doi.org/10.1044/2018_JSLHR-H-17-0473
Lee, Y., Sung, J. E., & Sim, H. (2014). Effects of listeners' working memory and noise on speech intelligibility in dysarthria. *Clinical Linguistics & Phonetics, 28*(10), 785–795. https://doi.org/10.3109/02699206.2014.904443
McGarrigle, R., Rakusen, L., & Mattys, S. (2021). Effortful listening under the microscope: Examining relations between pupillometric and subjective markers of effort and tiredness from listening. *Psychophysiology, 58*(1), Article e13703. https://doi.org/10.1111/psyp.13703
Moore, T. M., & Picou, E. M. (2018). A potential bias in subjective ratings of mental effort. *Journal of Speech, Language, and Hearing Research, 61*(9), 2405–2421. https://doi.org/10.1044/2018_JSLHR-H-17-0451
Nagle, K. F., & Eadie, T. L. (2012). Listener effort for highly intelligible tracheoesophageal speech. *Journal of Communication Disorders, 45*(3), 235–245. https://doi.org/10.1016/j.jcomdis.2012.01.001
Nagle, K. F., & Eadie, T. L. (2018). Perceived listener effort as an outcome measure for disordered speech. *Journal of Communication Disorders, 73,* 34–49. https://doi.org/10.1016/j.jcomdis.2018.03.003
Norris, D., & McQueen, J. M. (2008). Shortlist B: A Bayesian model of continuous speech recognition. *Psychological Review, 115*(2), 357–395. https://doi.org/10.1037/0033-295X.115.2.357
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
Panico, J., & Healey, E. C. (2009). Influence of text type, topic familiarity, and stuttering frequency on listener recall, comprehension, and mental effort. *Journal of Speech, Language, and Hearing Research, 52*(2), 534–546. https://doi.org/10.1044/1092-4388(2008/07-0238)
Peelle, J. E. (2018). Listening effort: How the cognitive consequences of acoustic challenge are reflected in brain and behavior. *Ear and Hearing, 39*(2), 204–214. https://doi.org/10.1097/AUD.0000000000000494
Pichora-Fuller, M. K., Kramer, S. E., Eckert, M. A., Edwards, B., Hornsby, B. W., 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
Pichora-Fuller, M. K., Schneider, B. A., & Daneman, M. (1995). How young and old adults listen to and remember speech in noise. *The Journal of the Acoustical Society of America, 97*(1), 593–608. https://doi.org/10.1121/1.412282
Rönnberg, J., Rudner, M., Foo, C., & Lunner, T. (2008). Cognition counts: A working memory system for ease of language understanding (ELU). *International Journal of Audiology, 47*(Suppl. 2), S99–S105. https://doi.org/10.1080/14992020802301167
Sarampalis, A., Kalluri, S., Edwards, B., & Hafter, E. (2009). Objective measures of listening effort: Effects of background noise and noise reduction. *Journal of Speech, Language, and Hearing Research, 52*(5), 1230–1240. https://doi.org/10.1044/1092-4388(2009/08-0111)
Sheard, C., Adams, R. D., & Davis, P. J. (1991). Reliability and agreement of ratings of ataxic dysarthric speech samples with varying intelligibility. *Journal of Speech and Hearing Research, 34*(2), 285–293. https://doi.org/10.1044/jshr.3402.285
Spencer, K., Friedlander, C., & Brown, K. (2020). Predictors of health-related quality of life and communicative participation in individuals with dysarthria from Parkinson's disease. *International Journal of Neurodegenerative Disorders, 3,* 14. https://doi.org/10.23937/2643-4539/1710014
Stipancic, K. L., Palmer, K. M., Rowe, H. P., Yunusova, Y., Berry, J. D., & Green, J. R. (2021). "You say severe, I say mild": Toward an empirical classification of dysarthria severity. *Journal of Speech, Language, and Hearing Research, 64*(12), 4718–4735. https://doi.org/10.1044/2021_JSLHR-21-00197
Strand, J. F., Ray, L., Dillman-Hasso, N. H., Villanueva, J., & Brown, V. A. (2020). Understanding speech amid the jingle and jangle: Recommendations for improving measurement practices in listening effort research. *Auditory Perception & Cognition, 3*(4), 169–188. https://doi.org/10.1080/25742442.2021.1903293
Wendt, D., Koelewijn, T., Książek, P., Kramer, S. E., & Lunner, T. (2018). Toward a more comprehensive understanding of the impact of masker type and signal-to-noise ratio on the pupillary response while performing a speech-in-noise test. *Hearing Research, 369,* 67–78. https://doi.org/10.1016/j.heares.2018.05.006
Whitehill, T. L., & Wong, C. C.-Y. (2006). Contributing factors to listener effort for dysarthric speech. *Journal of Medical Speech-Language Pathology, 14*(4), 335–342.
Wingfield, A., Alexander, A. H., & Cavigelli, S. (1994). Does memory constrain utilization of top-down information in spoken word recognition? Evidence from normal aging. *Language and Speech, 37*(3), 221–235. https://doi.org/10.1177/002383099403700301
Yorkston, K., Baylor, C., & Mach, H. (2017). Factors associated with communicative participation in amyotrophic lateral sclerosis. *Journal of Speech, Language, and Hearing Research, 60*(6S), 1791–1797. https://doi.org/10.1044/2017_JSLHR-S-16-0206

Downloaded from: https://pubs.asha.org Stephanie Borrie on 11/18/2022, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions