

Combining degradations: The effect of background noise on intelligibility of disordered speech

Sarah E. Yoho, and Stephanie A. Borrie

Citation: *The Journal of the Acoustical Society of America* **143**, 281 (2018);

View online: <https://doi.org/10.1121/1.5021254>

View Table of Contents: <http://asa.scitation.org/toc/jas/143/1>

Published by the *Acoustical Society of America*

Articles you may be interested in

[The impact of exploiting spectro-temporal context in computational speech segregation](#)

The Journal of the Acoustical Society of America **143**, 248 (2018); 10.1121/1.5020273

[Age effects on perceptual restoration of degraded interrupted sentences](#)

The Journal of the Acoustical Society of America **143**, 84 (2018); 10.1121/1.5016968

[Presenting and processing information in background noise: A combined speaker–listener perspective](#)

The Journal of the Acoustical Society of America **143**, 210 (2018); 10.1121/1.5020799

[Children's early bilingualism and musical training influence prosodic discrimination of sentences in an unknown language](#)

The Journal of the Acoustical Society of America **143**, EL1 (2018); 10.1121/1.5019700

[Auditory stream segregation of iterated rippled noises by normal-hearing and hearing-impaired listeners](#)

The Journal of the Acoustical Society of America **143**, 378 (2018); 10.1121/1.5021333

[Paradigmatic variation of vowels in expressive speech: Acoustic description and dimensional analysis](#)

The Journal of the Acoustical Society of America **143**, 109 (2018); 10.1121/1.5018433

Combining degradations: The effect of background noise on intelligibility of disordered speech

Sarah E. Yoho^{a)} and Stephanie A. Borrie

Department of Communicative Disorders and Deaf Education, Utah State University, Logan, Utah 84321, USA

(Received 19 December 2017; accepted 26 December 2017; published online 19 January 2018)

The effect of background noise on intelligibility of disordered speech was assessed. Speech-shaped noise was mixed with neurologically healthy (control) and disordered (dysarthric) speech at a series of signal-to-noise ratios. In addition, bandpass filtered control and dysarthric speech conditions were assessed to determine the effect of noise on both naturally and artificially degraded speech. While significant effects of both the amount of noise and the type of speech were revealed, no interaction between the two factors was observed, in either the broadband or filtered testing conditions. Thus, it appears that there is no multiplicative effect of the presence of background noise on intelligibility of disordered speech relative to control speech. That is, the decrease in intelligibility due to increasing levels of noise is similar for both types of speech, and both types of testing conditions, and the function for dysarthric speech is simply shifted downward due to the inherent source degradations of the speech itself. Last, large-scale online crowdsourcing via Amazon Mechanical Turk was utilized to collect data for the current study. Findings and implications for this data and data collection approach are discussed. © 2018 Acoustical Society of America.

<https://doi.org/10.1121/1.5021254>

[JFL]

Pages: 281–286

I. INTRODUCTION

The recognition of speech in everyday life typically occurs in sub-optimal listening conditions. A number of factors contribute to these adverse conditions. [Mattys et al. \(2012\)](#) have described a classification scheme that categorizes such factors according to *environmental* and *source* degradations. Environmental degradations refer to external factors acting upon the speech signal such as masking noise or filtering of the signal. For example, the intelligibility of speech is significantly reduced if the acoustic signal is presented in the presence of background noise ([Miller, 1947](#)), or if the speech of one talker is confused with the speech of other, concurrent talkers ([Kidd et al., 2005](#)). Source degradations, on the other hand, arise from the speech signal itself, with examples including the presence of a foreign accent or some type of speech disorder (e.g., dysarthria). There is a large body of literature in the area of speech perception detailing the independent effects of different types of listening adversity on speech intelligibility; however, the effects of simultaneous adversity have received much less attention. Yet, listeners in the real world are frequently required to perceive speech that has been degraded by several co-occurring factors. For example, deciphering the speech of a talker with dysarthria in the presence of background noise.

There has been some investigation into the effects of combined adversity, source and environmental, on intelligibility of the speech signal. [Adank et al. \(2009\)](#) showed that processing time is affected when listening to a non-native or unfamiliar accent in noise, and that processing time

increases as a function of decreasing signal-to-noise ratio (SNR). In addition, [Munro \(1998\)](#) had native speakers of American English listen to true/false statements produced in English by native speakers of American English and native speakers of Mandarin in both quiet and noisy conditions. The study found a larger increase in errors between quiet and noisy conditions for the foreign-accented speech relative to the native American speech. The author suggested that noise may produce a larger drop in intelligibility for non-native speech than for native speech (operationally defined here as a multiplicative effect), but indicated that more research on this speculation was required.

When the source degradation is due to the presence of a speech disorder, effects of the combined degradations inherent to signal production and external degradations such as noise have been observed. [McColl et al. \(1998\)](#) evaluated listeners' subjective impressions of tracheoesophageal speech—a surgical-prosthetic method of speaking after a patient has undergone a total laryngectomy and tracheoesophageal puncture—relative to healthy control speech in noisy conditions. The study involved presenting listeners with both types of speech at nine SNRs that varied widely (from +65 dB SNR, or effectively quiet, to –15 dB SNR). It was found that listeners rated transesophageal speech more negatively than control speech in all conditions except the most negative SNRs of –10 and –15 dB, where the ratings converged. The effect of noise on the intelligibility of dysphonic speech—a speech signal characterized by auditory perceptual features of disordered voicing including roughness, breathiness, and strain—has also been examined. [Ishikawa et al. \(2017\)](#) presented listeners with speech samples from speakers with typical speech and speakers with dysphonia, in quiet conditions and at two

^{a)}Electronic mail: sarah.leopold@usu.edu

SNRs (+5 and 0 dB). As expected, the dysphonic speech was significantly less intelligible than the typical speech, and there was a significant effect of SNR. Similarly, Lee *et al.* (2011) showed that spastic dysarthria is affected by background noise, with less favorable SNRs being more disruptive than more favorable SNRs. Last, the effect of noise on the intelligibility of hypokinetic dysarthric speech has also been examined (Dykstra *et al.*, 2012). In this study, the presence of background noise had a greater impact on intelligibility of the disordered speech as compared to the control speech, suggesting that there may have been a multiplicative effect when source and environmental degradations concurrently occur.

Regardless of the previously discussed findings, the specific nature of combined environmental and source degradation effects on intelligibility remain largely unclear. Despite an assumption that there is a multiplicative effect of these two types of degradations (e.g., Dykstra *et al.*, 2012), the existing literature is limited and does not entirely support this speculation. The majority of studies in this area appear to simply show a shift in intelligibility for source-degraded speech in noise which parallels the shift in intelligibility for the source-degraded speech in quiet conditions. Unfortunately, these results are often obscured by the presence of ceiling or floor effects in the data (e.g., Munro, 1998; Ishikawa *et al.*, 2017; Dykstra *et al.*, 2012).

The primary aim of the current study was to conduct a large, systematic evaluation of the combined effects of source and environmental degradation on intelligibility of speech in order to address the following research question: does systematically increasing the level of environmental degradation differentially influence the magnitude of intelligibility decline of disordered speech relative to healthy control speech? Given the lack of supporting evidence, we hypothesized that there is no multiplicative effect of these combined source and environmental degradations. To represent environmental degradation, we used speech-shaped noise as our initial test case for noise. The use of speech-shaped noise as masker, as opposed to babble or other forms of noise that involve informational masking, allowed us to avoid confounding variables such as the linguistic content impacting intelligibility of the target speech (e.g., Calandruccio *et al.*, 2010). To represent natural source degradation, we used dysarthric speech, a motor speech disorder arising from neurological origins (e.g., stroke, traumatic brain injury, Parkinson’s disease). Existing literature indicates that the presence of dysarthria significantly impacts intelligibility in otherwise optimal listening conditions (e.g., Borrie, 2015; Hustad, 2008). Finally, we used band-pass filtering to create artificially degraded speech conditions with both the healthy control and dysarthric speech. Restricting a speech signal via filtering is a commonly encountered environmental degradation (i.e., telephone communication) and is known to negatively impact intelligibility (Pollack, 1948). The purpose of the filtered conditions was two-fold: to allow for a representative comparison of intelligibility of dysarthric and control speech in quiet conditions, and to document the effects of both natural and artificial degradations (disordered and filtered speech, respectively). Last, we used semantically anomalous phrases to restrict top-down, cognitive influences on intelligibility. To determine if the effect of background noise on

disordered speech is more acute than the impact of background noise on control speech (in other words, if there is a multiplicative effect for disordered speech), the rate of magnitude of intelligibility decrease as a function of SNR was compared for each type of speech examined.

II. METHODS

A. Listener participants

A total of 260 adults (119 males and 141 females), 16 to 70 yr of age [$M = 36.62$, standard deviation (SD) = 10.29], participated as listeners in this study. All listener participants were native speakers of American English and living in the United States. Participants reported no history of speech, language, or hearing problems, and no significant prior contact with persons having neurogenic speech disorders. Demographic information regarding age, geographic region, and level of education of the participants is available in Table I.

Participants were recruited using the crowdsourcing website, Amazon Mechanical Turk¹ (MTurk; <http://www.mturk.com>). All participants were considered voluntary workers, protected through MTurk’s participation agreement and privacy notice. We used a number of setup options regarding participant prerequisites, limiting participation to individuals with a previous approval rate of greater than or equal to 99% and a confirmed status of U.S resident. This data collection method was approved by Utah State University Institutional Review Board (IRB).

B. Speech stimuli

The stimuli consisted of 80 syntactically plausible but semantically anomalous phrases (e.g., amend estate approach). Phrases were all six syllables in length and ranged from three

TABLE I. Demographic distribution data expressed in percentage scores for listener participants.

Gender	
Males	46
Females	54
Age	
≥50	13
40–49	19
30–39	42
≤29	26
Education	
Master’s	5
Bachelor’s	45
Attending College	10
High School Graduate	38
GED	1
Haven’t Graduated High School	1
Region	
Midwest	18
Northeast	26
Pacific	13
Rocky Mountain	2
Southeast	33
Southwest	8

to five words. These phrases, which reduce the influence of lexical cues on perceptual processing, were created specifically for examining speech perception in adverse conditions (Liss *et al.*, 1998) and have been used extensively in the study of perception of dysarthric speech (e.g., Borrie *et al.*, 2012; Borrie *et al.*, 2017a).

Two 72-yr-old male native talkers of American English, one with dysarthria and one age-matched neurologically healthy control, produced the stimuli for the study. The talker with dysarthria presented with a mild-moderate ataxic dysarthria secondary to cerebellar disease. His speech was characterized perceptually by excess and equal stress (scanning speech), prolonged phonemes and intervals, monotone, monoloudness, and imprecise articulation. The diagnosis was made by three independent Speech–Language Pathologists with expertise in differential diagnosis of motor speech disorders.

A speech-shape noise (SSN) was created for each of the two talkers independently. To do so, a 10-s white noise was shaped in MATLAB with a 1000-order FIR2 filter with the response characteristics of a 65 000-point, Hanning-windowed fast Fourier transform of the concatenated phrases from each individual talker. Prior to mixing with noise, all phrases were equated based on root-mean-square (rms) and a minimum of 100 ms of silence was added to the beginning and end of each phrase. The noise file was then looped to match the approximate length of the concatenated speech file, and the speech and noise files were mixed at the desired SNRs. The 48 k, 16-bit test stimuli were processed to create 13 testing conditions. Conditions were subdivided into two blocks: broadband and bandpass filtered (see Table II for summary of conditions). There were seven broadband conditions: both dysarthric and control speech mixed with SSN at 0, +3, and +6 dB SNR, as well as dysarthric speech in quiet. For the filtered conditions, the speech stimuli were bandpass filtered from 500 to 2500 Hz (500 order, FIR1 filter in MATLAB). There were six filtered conditions: both dysarthric and control speech in quiet, and mixed with SSN at +6 and +9 dB SNR. The SNRs and filter bandwidth were chosen based on pilot testing to ensure intelligibility was not at ceiling or floor for any condition.

C. Procedure

A brief description of the study task (including required use of headphones and completing the experiment in a quiet

TABLE II. Details of testing conditions. Listener participants were randomly assigned to one of 13 conditions ($n = 20$).

Control Speech	Dysarthric Speech
Broadband	
Quiet	
+ 6 dB SNR	+ 6 dB SNR
+3 dB SNR	+3 dB SNR
0 dB SNR	0 dB SNR
Filtered	
Quiet	Quiet
+9 dB SNR	+9 dB SNR
+6 dB SNR	+6 dB SNR

room with no distractions), time commitment, and remuneration (\$3 + \$2 bonus²) was posted on MTurk. Interested individuals were directed to a web page, loaded with a listener-perception application hosted on a secure university-based web server. Before beginning the study, individuals were required to read through the IRB approved consent form. By clicking “Agree,” individuals indicated that they had read and understood the information provided in the consent form and voluntarily agreed to participate. Participants were then required to complete a brief questionnaire regarding demographic information and questions related to inclusion/exclusion criteria. Upon completion of the questionnaire, listener participants were randomly assigned to one of 13 testing conditions ($n = 20$) before advancing to the experimental portion of the study.

Listener participants were told that they would be presented with 80 phrases that would be difficult to understand, either because the phrases would be produced by a person with a speech disorder and/or lots of background noise. They were also told that phrases contained real English words but would not make sense. Phrases were presented one at a time, and following each presentation, listeners were instructed to use the keyboard to type out exactly what they thought was being said. Listeners were strongly encouraged to make a guess at any words they did not recognize. Once they had finished typing their response, listeners were instructed to press the return key to move on to the next phrase. The self-paced experimental procedure took approximately 30 min to complete.

D. Transcript analysis

The total data set consisted of 260 listener transcripts, each containing 80 speech phrases. Transcripts were analyzed for correct words using previously established scoring criteria for the semantically anomalous phrases (Liss *et al.*, 1998; Borrie *et al.*, 2012) and an in-house computer program. The program automatically scored words as correct if they matched the intended target exactly or differed only by tense (-ed) or plurality (-s). Homophones and obvious spelling errors were also scored as correct. A percentage words correct (PWC) score was tabulated for each listener to reflect intelligibility performance. Twenty percent of the transcripts were randomly selected and reanalyzed by a human to examine reliability for coding words correct. Discrepancies between the computer and human revealed high agreement with Pearson correlation r score above 0.99.

III. RESULTS

Intelligibility scores, expressed as PWC, are shown for each broadband and filtered speech condition in Fig. 1 (top and bottom panels, respectively). As illustrated in the figure, intelligibility scores for both control and dysarthric speech decreased as a function of SNR. In addition, PWC scores for dysarthric speech were lower than for control speech. However, of important note is that the magnitude of decrease was comparable for dysarthric and control speech, indicating a lack of a multiplicative effect due to disordered speech. This

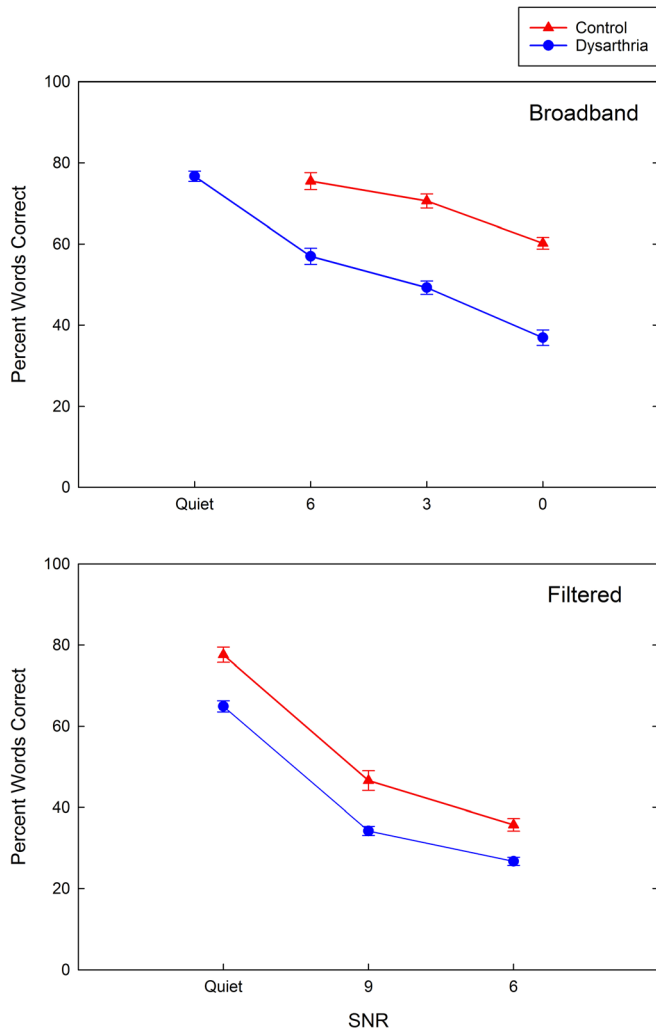


FIG. 1. (Color online) Average intelligibility, as measured by percent words correct, as a function of SNR for the broadband and filtered speech conditions (top and bottom panels, respectively; $n = 20$). Red triangles represent control speech and blue circles represent dysarthric speech. Error bars delineate ± 1 standard error of the mean.

was true for both broadband and filtered speech functions. These observations were confirmed with statistical analysis.

A two-way analysis of variance was performed for both the broadband and filtered speech conditions to examine the main effects of type of speech and SNR, as well as the interaction between the main effects on PWC scores. As anticipated, for the broadband conditions, the analysis revealed a significant main effect of the type of speech [$F(1, 114) = 202.7, p < 0.001$] a significant main effect of SNR [$F(2, 114) = 49.2, p < 0.001$]; however, the interaction was not significant. Similarly, for the filtered conditions, the analysis revealed a significant main effect of the type of speech [$F(1, 114) = 73.8, p < 0.001$], a significant main effect of SNR [$F(2, 114) = 332.8, p < 0.001$], and the interaction was not significant.

IV. DISCUSSION

The significant main effects of SNR and type of speech (control or dysarthric) indicate that both amount of noise and presence of a neurological speech disorder negatively impact

intelligibility. Although these findings are not surprising, the key comparisons of interest for the purposes of the current study are the interactions between amount of SNR and type of speech for each function. The non-significant interactions indicate that there is not a multiplicative effect of SNR on dysarthric speech relative to control speech. The decrease in intelligibility due to increasing levels of noise is similar for both dysarthria and control speech, and the function for dysarthric speech is simply shifted downward due to the inherent source degradations of the speech itself. This trend holds for both the broadband and filtered conditions.

The lack of interaction between amount of noise and type of speech for both the broadband and filtered functions indicates that not only is there no multiplicative effect of noise on neurologically disordered speech, but that there is similarly no effect even when an artificial distortion (filtering) is applied to the signal. For the filtered conditions, the magnitude of shift in intelligibility between control and dysarthric speech in quiet conditions is comparable to the magnitude of shift in noise, indicating that the effect of filtering alone may not produce a multiplicative effect on intelligibility either. Most importantly, it appears that in the conditions tested here, there is no multiplicative effect of noise on disordered or degraded speech relative to neurologically healthy speech.

Regardless of this lack of interaction, the impact of combined sources of degradations such as disordered speech and the presence of noise are important to consider. In a given suboptimal acoustic environment (i.e., a particular SNR), the intelligibility of disordered speech will be substantially lower than the intelligibility of neurologically healthy speech. While the focus of this study is not clinical application, these results do afford empirical evidence for a clinical strategy currently used in the management of dysarthria—educating the patient and their communication partners about the need to select a conducive speaking and listening environment (e.g., turn off the television, avoid noisy restaurants), particularly when scheduling important interactions (Duffy, 2005).

Despite a somewhat implicit assumption of a multiplicative effect of noise in listener perception of degraded speech, the results of the current study do not stand in obvious contrast to the results of other, related studies on the matter. However, it is difficult to conclude this entirely, as there have been several factors precluding a thorough and complete analysis of the effect of noise on degraded speech in previous findings, such as ceiling and/or floor effects (Munro, 1998; Rogers *et al.*, 2004; Dykstra *et al.*, 2012), no statistical analysis of the interaction effect (Ishikawa *et al.*, 2017), not controlling for SNR (Adams *et al.*, 2008), or simply that the study did not measure intelligibility directly (McCull *et al.*, 1998). The initial data available on the effect of noise on dysarthric speech specifically appears to be in agreement with the current findings. Lee *et al.* (2011) found a relationship between amount of noise and intelligibility of dysarthric speech, but there were no control conditions involving neurologically healthy speech to draw conclusions about the magnitude of intelligibility decline as a function of SNR. Conversely, Adams *et al.* (2008) did examine the magnitude of intelligibility decline between control and dysarthric speech (hypokinetic dysarthria); however,

the study did not control for SNR and results were substantially limited by the use of a very small number of experienced listener participants (two students in graduate school for Speech–Language Pathology). Despite this, a similar pattern of decrease was observed for both control and dysarthric speech, which is in accord with the findings of the current study. Alternatively, the pattern of results for the effect of noise on dysphonic speech appears to display a more multiplicative effect than that of noise on control speech (Ishikawa *et al.*, 2017). However, the absence of statistical analysis of this interaction renders it difficult to conclude this definitively. Similarly, data from Dykstra *et al.* (2012) appear to show a multiplicative effect of noise on hypokinetic dysarthric speech, but ceiling effects in the healthy speech control conditions make the results difficult to interpret as such.

The current study offers several factors that allow for a more conclusive and thorough analysis of the effect of noise on disordered speech. The inclusion of both naturally occurring source degradation (dysarthria) and artificially created environmental degradations (noise and filtering) permits a systematic evaluation of several combinations of degraded listening situations. While a direct comparison between the broadband and filtered speech conditions was not possible due to differing SNRs, the lack of interaction between healthy and dysarthric speech in either of the functions provides strong support that differing combinations of environmental and source degradations may not result in a multiplicative effect. In addition, SNRs were carefully chosen via pilot testing to preclude any possible floor or ceiling effects in the data. While an accurate examination of broadband control and dysarthric speech in quiet is not possible due to the optimal intelligibility of control speech in such conditions, utilizing bandpass filtering in the current study allowed for the inclusion of quiet conditions in the statistical analysis.

Although not a primary aim of the current study, the success of data collection via online crowdsourcing is an important factor to acknowledge. By crowdsourcing the experiment via MTurk, we were able to collect data from a large, diverse population. Whereas studies in speech perception are typically collected using convenience samples of the most readily-available population (i.e., young adult, college students), crowdsourcing allows for the rapid recruitment of a large heterogeneous sample that more closely represents the general population (see Table I), while still controlling for necessary variables (i.e., country of residence, native language, previous experience with dysarthric speech). Criticisms of such data collection methods have included lack of control over stimulus presentation levels and testing environment. However, compelling comparable results have been found with data collected via MTurk and data collected in the laboratory, including studies involving speech perception in adverse conditions, such as perception of disordered speech and speech in background noise (Cooke *et al.*, 2011; Lansford *et al.*, 2016; McAllister Byun *et al.*, 2015; Slote and Strand, 2016). As such, a number of studies in speech perception in adverse conditions have gone on to make use of data collection via MTurk (e.g., Borrie *et al.*, 2017a, 2017b). While we did not compare data collection environments, the data collected in the current study displayed a surprisingly small degree of variability

across the 20 listener participants in each testing condition (see error bars on Fig. 1). This result may be due in part to the use of a motivational bonus payment, which offered additional monetary compensation for participants whose transcripts showed evidence that they had correctly followed instructions and responded thoughtfully.

A final important factor in the current study was the use of speech-shaped noise rather than the commonly employed multi-talker babble (e.g., Adams *et al.*, 2008), which may have increased the degree of control in the comparisons between control and degraded speech. The effects of informational masking, or masking due to linguistic, semantic, and other similarities between the signal and the noise, are highly complex. For example, the use of non-native babble relative to native babble as a masker has been shown to decrease the amount of informational masking on speech (Van Engen and Bradlow, 2007). However, there is still much unknown about the differential effects of informational masking on disordered versus non-disordered speech. It is possible that the perceptual differences between disordered speech, such as dysarthria, and the neurologically healthy speech which generally makes up babble, may be enough to permit a release from masking. If so, observed differential effects of SNR on control and disordered speech may be due in part to differences in informational masking rather than the effects of noise more generally. Therefore, an examination of these effects is warranted before any conclusions can be made about the effect of babble noise on disordered speech.

One limitation of the current study is the use of only a single talker of each type of speech. Although this was specifically done to allow for more precise experimental control, a future investigation of the effects of different types and severities of dysarthria is necessary. The speaker with dysarthria who provided the speech stimuli in the current study presented with a mild-moderate speech disorder, with intelligibility on semantically anomalous sentences in quiet conditions at approximately 80% correct. It is plausible that with the inclusion of more severely degraded presentations of dysarthria, multiplicative effects of noise may emerge.

An additional direction for future investigation involves the effect of the listener on disordered, noisy speech. In addition to the source and environmental degradations that challenge speech intelligibility, there are receiver or listener limitations that can also play a substantial role (Mattys *et al.*, 2012). The introduction of sensorineural hearing impairment commonly results in two main perceptual consequences for the listener: a reduction in audibility and substantial difficulty understanding speech, particularly in the presence of noise (Moore, 1996). Therefore, the combined effects of both disordered hearing of the listener and disordered speech of the talker may be quite profound, as well as complex. Given that a primary concern of listeners with hearing loss is a difficulty understanding speech in noise, much work has gone in to the development of hearing aid technology to improve SNR and increase speech intelligibility (e.g., Healy *et al.*, 2013). However, studies with regard to the effect of source degradations on these noise reduction technologies are essential, particularly as the identification and segregation of disordered

speech is likely much more challenging than for typical speech (Seong *et al.*, 2014). Such studies would also address issues of ecological validity and clinical application, given that conditions such as presbycusis (age-related sensorineural hearing loss) and dysarthria generally occur in older adulthood.

V. CONCLUSIONS

In the current study, the impact of noise on degraded speech was systematically examined. Although there were significant effects of both the amount of noise and the type of speech, there was no interaction found between the two. Therefore, it can be concluded that there is no multiplicative effect of noise on dysarthric speech relative to healthy control speech. Instead, it appears that intelligibility is simply shifted downward as a function of the inherent source degradations arising from the presence of the neurological speech disorder of dysarthria. However, there remain many areas of investigation in this line of enquiry. In particular, the effects of combined listener and speaker limitations should be evaluated, given the complexities of both and possibilities for interactions. Last, the use of crowdsourcing to obtain perceptual data on the impact of noise on disordered speech appears to be an effective method, and should continue to be considered for future investigations.

ACKNOWLEDGMENTS

This paper was written with partial support from the National Institute of Deafness and Other Communication Disorders, National Institutes of Health Grant No. R21 DC 016084 (awarded to S.A.B.). We gratefully acknowledge our research assistants Paul Vicioso Osoria (P.V.O.) and Shea Long (S.L.), for development of the web-based application (P.V.O.) and reliability scoring of listener transcripts (S.L.). We would also like to thank Jee Eun Sung for personal correspondence concerning Y. Lee, H. S. Kim, and J. E. Sung (Lee *et al.*, 2011).

¹It has previously been demonstrated that results of studies looking at speech perception in adverse conditions (disordered speech and speech in noise) obtained by MTurk are comparable to those collected in the laboratory environment (Cooke *et al.*, 2011; Lansford *et al.*, 2016; McAllister Byun *et al.*, 2015; Slote and Strand, 2016).

²Participants were informed that they would be paid \$3 for the task but that they would receive a \$2 bonus payment if there was evidence that they had fully engaged in the phrase transcription task. All participants received the bonus.

Adams, S., Dykstra, A., Jenkins, M., and Jog, M. (2008). "Speech-to-noise levels and conversational intelligibility in hypophonia and Parkinson's disease," *J. Med. Speech Lang. Pathol.* **16**, 165–172.

Adank, P., Evans, B. G., Stuart-Smith, J., and Scott, S. K. (2009). "Comprehension of familiar and unfamiliar native accents under adverse listening conditions," *J. Exp. Psychol.: Hum. Percept. Perform.* **35**, 520.

Borrie, S. A. (2015). "Visual information: A help or hindrance to perceptual processing of dysarthric speech," *J. Acoust. Soc. Am.* **137**, 1473–1480.

Borrie, S. A., Baese-Berk, M., Van Engen, K., and Bent, T. (2017a). "A relationship between processing speech in noise and dysarthric speech," *J. Acoust. Soc. Am.* **141**, 4660–4667.

Borrie, S. A., Lansford, K. L., and Barrett, T. S. (2017b). "Generalized adaptation to dysarthric speech," *J. Speech, Lang. Hear. Res.* **60**, 3110–3117.

Borrie, S. A., McAuliffe, M. J., Liss, J. M., Kirk, C., O'Beirne, G. A., and Anderson, T. (2012). "Familiarisation conditions and the mechanisms that underlie improved recognition of dysarthric speech," *Lang. Cogn. Process.* **27**, 1039–1055.

Calandruccio, L., Dhar, S., and Bradlow, A. R. (2010). "Speech-on-speech masking with variable access to the linguistic content of the masker speech," *J. Acoust. Soc. Am.* **128**, 860–869.

Cooke, M., Barker, J., Lecumberri, M. L. G., and Wasilewski, K. (2011). "Crowdsourcing for word recognition in noise," in *Proceedings of the Twelfth Annual Conference of the International Speech Communication Association*, August 27–31, Florence, Italy, pp. 3049–3052.

Duffy, J. R. (2005). *Motor Speech Disorders: Substrates, Differential Diagnosis, and Management*, 2nd ed. (Elsevier, Amsterdam).

Dykstra, A. D., Adams, S. G., and Jog, M. (2012). "The effect of background noise on the speech intensity of individuals with hypophonia associated with Parkinson's disease," *J. Med. Speech Lang. Pathol.* **20**, 19–31.

Healy, E. W., Yoho, S. E., Wang, Y., and Wang, D. (2013). "An algorithm to improve speech recognition in noise for hearing-impaired listeners," *J. Acoust. Soc. Am.* **134**, 3029–3038.

Hustad, K. C. (2008). "The relationship between listener comprehension and intelligibility scores for speakers with dysarthria," *J. Speech Lang. Hear. Res.* **51**, 562–573.

Ishikawa, K., Boyce, S., Kelchner, L., Powell, M. G., Schieve, H., de Alarcon, A., and Khosla, S. (2017). "The effect of background noise on intelligibility of dysphonic speech," *J. Speech Lang. Hear. Res.* **60**, 1919–1929.

Kidd, G., Jr., Mason, C. R., and Gallun, F. J. (2005). "Combining energetic and informational masking for speech identification," *J. Acoust. Soc. Am.* **118**, 982–992.

Lansford, K. L., Borrie, S. A., and Bystricky, L. (2016). "Use of crowdsourcing to assess the ecological validity of perceptual training paradigms in dysarthria," *Am. J. Speech Lang. Pathol.* **25**, 233–239.

Lee, Y., Sim, H. S., and Sung, J. E. (2011). "Effects of the types of noise and signal to noise ratios on speech intelligibility in dysarthria," *Phon. Speech Sci.* **3**, 117–124.

Liss, J. M., Spitzer, S., Caviness, J. N., Adler, C., and Edwards, B. (1998). "Syllabic strength and lexical boundary decisions in the perception of hypokinetic dysarthric speech," *J. Acoust. Soc. Am.* **104**, 2457–2466.

McAllister Byun, T., Halpin, P. F., and Szeredi, D. (2015). "Online crowdsourcing for efficient rating of speech: A validation study," *J. Commun. Disord.* **53**, 70–83.

Mattys, S. L., Davis, M. H., Bradlow, A. R., and Scott, S. K. (2012). "Speech recognition in adverse conditions: A review," *Lang. Cogn. Process.* **27**, 953–978.

McColl, D., Fucci, D., Petrosino, L., Martin, D. E., and McCaffrey, P. (1998). "Listener ratings of the intelligibility of tracheoesophageal speech in noise," *J. Commun. Disord.* **31**, 279–289.

Miller, G. A. (1947). "The masking of speech," *Psychol. Bull.* **44**, 105–129.

Moore, B. C. (1996). "Perceptual consequences of cochlear hearing loss and their implications for the design of hearing aids," *Ear Hear.* **17**, 133–161.

Munro, M. J. (1998). "The effects of noise on the intelligibility of foreign-accented speech," *Stud. Second Lang. Acquis.* **20**, 139–154.

Pollack, I. (1948). "Effects of high pass and low pass filtering on the intelligibility of speech in noise," *J. Acoust. Soc. Am.* **20**, 259–266.

Rogers, C. L., Dalby, J., and Nishi, K. (2004). "Effects of noise and proficiency on intelligibility of Chinese-accented English," *Lang. Speech* **47**, 139–154.

Seong, W. K., Park, J. H., and Kim, H. K. (2014). "Reducing speech noise for patients with dysarthria in noisy environments," *IEICE Trans. Inf. Syst.* **97**, 2881–2887.

Slote, J., and Strand, J. F. (2016). "Conducting spoken word recognition research online: Validation and a new timing method," *Behav. Res. Methods* **48**, 553–566.

Van Engen, K. J., and Bradlow, A. R. (2007). "Sentence recognition in native- and foreign-language multi-talker background noise," *J. Acoust. Soc. Am.* **121**, 519–526.