# An Examination of Natural Background Noise on Spontaneous Speech Intelligibility in Parkinson's Disease

Mary Coleman, Amara Yaeger, B.S.*, Carrie Rountrey, Ph.D., CCC-SLP

University of Cincinnati

*University of Wisconsin - Madison

Abstract

Purpose: This study examined the effect of natural background noise ($BN_N$) on *conversational spontaneous speech intelligibility* (C-SSI in persons with Parkinson's Disease (PD). $BN_N$ characteristics included intensity (DB-SPL), frequency in the human speech range ($FREQ_{SP}$), consistency (CONS), presence of human speech (HU), presence of electronic noise (ELEC). This project is a first step toward improving our understanding of $BN_N$ effect on C-SSI in persons with PD.

Methods: 13 participants with PD and mild-mod hypokinetic dysarthria wore a recording device for 15 hours across 2 days. Recordings of the whole group and 2 clinically-relevant groups, low-intelligibility ($LOW_{C\text{-}SSI}$) and high-intelligibility ($HIGH_{C\text{-}SSI}$), were analyzed. Listeners transcribed 10 randomly selected sentences. Transcriptions were compared to a key for C-SSI. A 10ms clip in close temporal proximity to speech onset for each sentence was analyzed acoustically and visually for targeted $BN_N$ characteristics.

Results: There was no significant difference between or within subjects between dB-SPL and C-SSI, however, $HIGH_{C\text{-}SSI}$ and $LOW_{C\text{-}SSI}$ groups responded differently to dB-SPL ($p$=0.00463). When dB-SPL increased, C-SSI increased for $LOW_{C\text{-}SSI}$ speakers but decreased for $HIGH_{C\text{-}SSI}$ speakers. Further examination of the characteristics of $BN_N$ taken together showed that C-SSI decreased with increased dB-SPL. CONS related to a decrease C-SSI for the whole group, but an increase for $HIGH_{C\text{-}SSI}$ speakers. No other variables were significantly related to C-SSI.

Conclusions: Results suggest that $BN_N$ impacts C-SSI, and certain $BN_N$ characteristics may have different effects on speakers with PD overall. $HIGH_{C\text{-}SSI}$ and $LOW_{C\text{-}SSI}$ groups are impacted by $BN_N$ in different ways. Further research is needed.

## Introduction

Parkinson's Disease (PD) is a progressive neurodegenerative disorder usually found in individuals older than 50 years of age (Gelb et al., 1999). PD symptoms include bradykinesia, muscular rigidity, resting tremor, dystonia, including oromandibular dystonia, and hypokinetic dysarthria, (Darley et al., 1969; Gelb et al., 1999; Sveinbjornsdottir, 2016). These symptoms can greatly impact a speaker's intelligibility, which is a topic of great concern for PD patients. Intelligibility is defined as "the degree to which the speaker's intended message is recovered by the listener" (Kent et al., 1989). Previous studies show that naturally occurring background noise ($BN_N$) influences speech intelligibility and that the Lombard Effect impacts speech production, causing an increase in vocal intensity in speakers with PD, and impacting their intelligibility (Ishikawa et al., 2017; Stathopoulos et al., 2014). Previous research has studied speech elicited in a lab or clinic. The dataset used for this study is the first known dataset of recorded conversational speech samples from the home environment.

Speech production refers to the translation of thoughts into verbal output and is greatly impacted by PD as hypokinetic dysarthria changes aspects of production including articulation, phonation, prosody, and speech breathing (De Keyser et al., 2016). Perception of speech refers to the listener's ability to hear and understand verbal output. The current study highlights the impact of $BN_N$ on speech production, but recognizes that $BN_N$ will also influence speech perception, as perception and production are related within the language networks of the brain (De Keyser et al., 2016).

In previous analysis of this dataset, a significant relationship between intensity of $BN_N$ (dB-SPL) and intelligibility of spontaneous speech (C-SSI) was noted (Rountrey, 2015). Other studies have shown that noise characteristics besides dB-SPL can impact speech in healthy speakers. Characteristics were chosen based on possible impact on both speakers and listeners, especially keeping in mind the Lombard Effect. Frequency of $BN_N$ ($FREQ_{SP}$) has been found to impact speech recognition and intelligibility. This could be due to the need to separate speech signal from $BN_N$ for clear speech perception (Lee et al., 2015). The current study aims to determine which frequencies most impact C-SSI and how. It is interested specifically in frequency potentially inclusive of speech frequencies in the F1 and F2 space, falling in the mid value in this study. Consistency of $BN_N$ (CONS) is a common concern regarding intelligibility (Danhauer & Leppler, 1979; Lee et al., 2015; Stone et al., 2011). Previous studies have found that inconsistent $BN_N$ decreases C-SSI, while consistent $BN_N$ increases C-SSI (Stone et al., 2011), and one study found that when presented with 4 types of $BN_N$, speakers were least impacted by the consistent $BN_N$ samples (Danhauer & Leppler, 1979). $BN_N$ containing human speech (HU) has been shown to impact intelligibility and is of special interest due to research showing that non-stationary masking sounds are difficult to filter out and listeners have difficulty separating speech signals (Danhauer & Leppler, 1979; Darwin, 2008). $BN_N$ from electronic sources (ELEC) is of interest for several reasons, the primary being that some HU $BN_N$ is also ELEC, such as TV, phone, and radio speech. While no previous research has been done on the differences between ELEC and normal HU, ELEC speech has been found to be processed differently than normal, in-person speech (Liu et al., 2009).

From the literature, we developed several hypotheses. We hypothesized that $FREQ_{SP}$ in the speech range (100-8,000 Hz) will have the greatest impact on perception as previous research shows that speech $BN_N$ is more difficult to filter and has the same dB-SPL as the conversational partner (Darwin, 2008). When $BN_N$ is consistent, C-SSI will increase. The presence of HU will negatively impact C-SSI, as will the presence of ELEC $BN_N$.

# Methods

*Participants*

13 volunteers with PD and mild to moderate speech impairments characterized by hypokinetic dysarthria completed the study. 9 males and 4 females, ranging in age from 55-74 years with a median age of 66.9 years, participated. Participants did not have any other neurological disorders and per predetermined exclusion criteria, each participant scored at least 23/30 on a Mini-Mental State Examination (MMSE), which was used to determine that the individuals did not have dementia. All participants also had normal hearing upon screening.

*Data Collection*

Participants wore the Language Environment Analysis (LENA) speech recorder for 15 hours a day across 2 days to capture spontaneous speech and environmental background noise, (Ford, Baer, Xu, Yapanel, & Gray, 2008). 10 participant sentences were randomly selected from the recordings along with the NOISE that naturally accompanied the sentences, and set aside for analysis.

*Analysis*

A key was generated by researchers after listening to each randomly selected sentence as many times as needed at whatever volume was needed, and by listening to the surrounding sentences to determine the intention of the verbiage. The key was well matched for phonetic and linguistic predictability. Given that the participants were in the mild-moderate range of speech impairment, confidence of the researchers for the accuracy of the key was high. The sentences were then transcribed by 3 novel listeners. The transcripts were compared to the key to determine the percent of words intelligible for the variable of interest, C-SSI.

Data on average overall C-SSI was used to split participants into two naturally occurring and clinically relevant groups. $HIGH_{C\text{-}SSI}$ had an average C-SSI of over 75%. $LOW_{C\text{-}SSI}$ had an average C-SSI of less than 75%. The difference between these groups was measured with chi-square.

*Analysis*

For each sentence, a 10 ms clip of $BN_N$ in close temporal proximity to the onset of speech was captured. The intensity (dB-SPL) of this noise was measured in dB-SPL using Praat and compared to C-SSI via linear regression per participant. 5 noise characteristics emerged from literature searches and listening to the data samples. The researchers hypothesized these characteristics to have an effect on C-SSI and the 10ms of $BN_N$ were analyzed for them.

A group of 5 listeners (including 4 novel listeners) participated in analyzing the following characteristics: $FREQ_{SP}$, CONS, HU, and ELEC. $FREQ_{SP}$ was divided into 4 categories: white noise, low, mid, and high. White noise was defined as a broad spectrum noise with no density. Low was less than 500 Hz, mid was 500-2,000 Hz, and high was over 2,000 Hz. CONS was analyzed for whether $BN_N$ was consistent or continuous across time, or not. HU was analyzed for whether human speech was present in NOISE, including electronic human voices, such as a radio or TV. Listeners could also mark a NOISE sample as undetermined for HU. ELEC was analyzed for the presence of electronic or mechanical noise, such as a microwave, radio or fan. Listeners could also mark ELEC as undetermined. $BN_N$ that was determined by listeners to be an electronic version of a human voice, such as someone speaking on the radio or TV were marked as positive for both HU and ELEC.

These characteristics of $BN_N$ in each sentence were initially analyzed independently by 2 novel listeners. The listeners could listen to each clip as many times as needed and were able to change their answers later as they became more familiar with the task and could recognize the different characteristics more accurately. Interrater reliability (IRR) was calculated: 77.69% ($FREQ_{SP}$), 90.00% (CONS), 76.92% (HU), and 64.62% (ELEC), typical of the IRR of other perceptual measures. 2 authors reanalyzed the disagreements together in order to come to consensus. They independently spot-checked the agreements, but made no changes. The data ultimately used for the analysis and referred to in the results had an IRR of 100% for all characteristics. A Repeated Measures ANOVA was used to examine the characteristics of the natural background noise for the response variable C-SSI.

## Results

There was no significant difference found between or within subjects, but $HIGH_{C\text{-}SSI}$ and $LOW_{C\text{-}SSI}$ groups responded differently to the background noise ($p$=0.00463, $\chi^2$=8.0133,1). As dB-SPL increased, C-SSI of the $LOW_{C\text{-}SSI}$ group also increased. C-SSI of the $HIGH_{C\text{-}SSI}$ group decreased. Only CONS had a significant effect on C-SSI ($F$=7.966, $p$=0.00588). Participants as a whole showed decreased C-SSI with consistent noise.

When separated by group, $HIGH_{C\text{-}SSI}$ and $LOW_{C\text{-}SSI}$, CONS was significant when the group was $HIGH_{C\text{-}SSI}$ ($F$=4.337, $p$=0.0421). Participants in the $HIGH_{C\text{-}SSI}$ group had increased C-SSI with consistent noise. For the group $LOW_{C\text{-}SSI}$, only dB-SPL was significant ($F$=4.329, $p$=0.0475). Participants in the $LOW_{C\text{-}SSI}$ group showed increased C-SSI with increased **d**B-SPL.

## Discussion

In the first known study of its kind, the conversational spontaneous speech intelligibility (C-SSI) of speakers with Parkinson's Disease (PD) was analyzed, as well as the effects of natural background noise ($BN_N$) on C-SSI and its characteristics: intensity (dB-SPL), frequency ($FREQ_{SP}$), consistency (CONS), presence of human speech (HU), and presence of electronic noise (ELEC). Participants wore LENA recording devices to capture C-SSI and $BN_N$ samples. These samples were analyzed to determine the C-SSI of each participant. The samples were then analyzed for each of the $BN_N$ characteristics. These characteristics were examined for the response variable C-SSI for the population as whole and for 2 clinically relevant groups: speakers with high intelligibility ($HIGH_{C\text{-}SSI}$) and speakers with low-intelligibility ($LOW_{C\text{-}SSI}$).

From previous research on speech intelligibility led to hypotheses for each $BN_N$ characteristic. $FREQ_{SP}$ that fell in the speech range (100-8,000 Hz); for this study, the mid value was hypothesized to have the greatest impact, with a decreased C-SSI. When $BN_N$ was consistent, C-SSI was hypothesized to increase. The presence of HU was hypothesized to negatively impact C-SSI. The presence of ELEC $BN_N$ was also hypothesized to decrease C-SSI.

This study found that the presence of more intense $BN_N$ may positively impact $LOW_{C\text{-}SSI}$ speakers but that $HIGH_{C\text{-}SSI}$ speakers may not have the same benefit. These findings show that the $LOW_{C\text{-}SSI}$ group may have an increased sensitivity to the Lombard Effect, resulting in upregulation of the speech system. When dB-SPL increased, the $LOW_{C\text{-}SSI}$ group showed an increase in C-SSI, while the $HIGH_{C\text{-}SSI}$ group showed a decrease. This first finding also showed that $BN_N$ has varying impact on speakers, depending on the severity of their dysarthria. The $HIGH_{C\text{-}SSI}$ group's decreasing $LOW_{C\text{-}SSI}$ group with increased dB-SPL may be due to the impact of noise on the listener. These findings led to considering other noise characteristics that may affect C-SSI and the severity groups.

Interestingly, only CONS was found to significantly impact C-SSI. Results showed CONS to be related to a decrease in C-SSI for the sample as a whole. However, speakers with $\text{HIGH}_{\text{C-SSI}}$ increased intelligibility in the presence of consistent noise. The same effect was not seen for the $\text{LOW}_{\text{C-SSI}}$ group. Listeners may be more likely to adjust for consistent, or continuous, $\text{BN}_{\text{N}}$ when in conversation with a $\text{HIGH}_{\text{C-SSI}}$ speaker. The speaker may also be influenced by consistent $\text{BN}_{\text{N}}$ in that they may lose the Lombard Effect as they become more accustomed to the presence of the $\text{BN}_{\text{N}}$.

Several $\text{BN}_{\text{N}}$ characteristics were hypothesized to have effect based on previous research, but were not found in the current study. Of note, $\text{BN}_{\text{N}}$ overlapping common speech frequencies, which has been found to negatively impact speech intelligibility, did not show a significant effect on C-SSI. The presence of HU did not show a significant effect, even when combined with $\text{FREQ}_{\text{SP}}$. This was hypothesized to cause a decrease in C-SSI based on researcher's perception of listener and speaker distraction in presence of other human speech, as well as previous finding that human speech in $\text{BN}_{\text{N}}$ is more difficult to filter and has a negative impact on C-SSI. The presence of ELEC did not show a significant effect. This evidence nullifies our hypothesis that ELEC would have an effect base on the Lombard Effect.

Potential limitations of this study include the difficulty in obtaining clear samples of each $\text{BN}_{\text{N}}$ characteristic and in correctly identifying each characteristic during analysis, as some were undeterminable. The small dataset is another limitation as combinations of the $\text{BN}_{\text{N}}$ characteristics could impact the results, and this study did not consider this possible effect or analyze which characteristics occurred together. For example, it is possible that in this dataset, all occurrences of HU in $\text{BN}_{\text{N}}$ were in the same sample as consistent noise or the sound overall was marked as white noise, which could potentially impact our examination of these characteristics' influence on C-SSI.

The findings of this study suggest that $BN_N$ and its characteristics, specifically dB-SPL and CONS, may impact C-SSI in speakers with PD. Further research on C-SSI and $BN_N$ characteristics is needed for external validity and more information on each $BN_N$ characteristic. Similar research on speakers with other neurodegenerative disorders is also needed to better understand the impact of $BN_N$ on C-SSI.


## Acknowledgements

The University of Cincinnati Cincinnatus Scholarship provided tuition support during the time of the project for one author. The University of Cincinnati Communication Sciences and Disorders Department provided faculty support for one author.

References

Danhauer, J. L., & Leppler, J. G. (1979). Effects of four noise competitors on the California consonant test. *Journal of Speech and Hearing Disorders*, *44*(3), 354-362.

Darley, F. L., Aronson, A. E., & Brown, J. R. (1969). Differential diagnostic patterns of dysarthria. *Journal of speech and hearing research, 12*(2), 246-269.

Darwin, C. J. (2008). Listening to speech in the presence of other sounds. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*(1493), 1011-1021.

De Keyser, K., Santens, P., Bockstael, A., Botteldooren, D., Talsma, D., De Vos, S., ... & De Letter, M. (2016). The relationship between speech production and speech perception deficits in Parkinson's disease. *Journal of Speech, Language, and Hearing Research*, *59*(5), 915-931.

Ford, M., Baer, C. T., Xu, D., Yapanel, U., & Gray, S. (2008). The LENA™ language environment analysis system: Audio specifications of the DLP-0121. *Boulder, CO: Lena Foundation*.

Gelb, D. J., Oliver, E., & Gilman, S. (1999). Diagnostic criteria for Parkinson disease. *Archives of neurology*, *56*(1), 33-39.

Hornykiewicz, O. (1998). Biochemical aspects of Parkinson's disease. *Neurology*, *51*(2 Suppl 2), S2-S9.

Ishikawa, K., Boyce, S., Kelchner, L., Powell, M. G., Schieve, H., de Alarcon, A., & Khosla, S. (2017). The effect of background noise on intelligibility of dysphonic speech. *Journal of Speech, Language, and Hearing Research*, *60*(7), 1919-1929.

Kent, R. D., Weismer, G., Kent, J. F., & Rosenbek, J. C. (1989). Toward phonetic intelligibility testing in dysarthria. *Journal of Speech and Hearing Disorders*, *54*(4), 482-499.

Lee, J. Y., Lee, J. T., Heo, H. J., Choi, C. H., Choi, S. H., & Lee, K. (2015). Speech recognition in real-life background noise by young and middle-aged adults with normal hearing. *Journal of audiology & otology*, *19*(1), 39.

Liu, C., Fu, Q. J., & Narayanan, S. S. (2009). Effect of bandwidth extension to telephone speech recognition in cochlear implant users. *The Journal of the Acoustical Society of America*, *125*(2), EL77-EL83.

Rountrey, C. E. (2015). Speech Function in Persons with Parkinson’s Disease: Effects of Environment, Task and Treatment. JMU Scholarly Commons.

Stathopoulos, E. T., Huber, J. E., Richardson, K., Kamphaus, J., DeCicco, D., Darling, M., ... & Sussman, J. E. (2014). Increased vocal intensity due to the Lombard effect in speakers with Parkinson's disease: Simultaneous laryngeal and respiratory strategies. *Journal of communication disorders*, *48*, 1-17.

Stone, M. A., Füllgrabe, C., Mackinnon, R. C., & Moore, B. C. (2011). The importance for speech intelligibility of random fluctuations in “steady” background noise. *The Journal of the Acoustical Society of America*, *130*(5), 2874-2881.

Sveinbjornsdottir, S. (2016). The clinical symptoms of Parkinson's disease. *Journal of neurochemistry*, *139*, 318-324.

Yorkston, K. M., Strand, E. A., & Kennedy, M. R. (1996). Comprehensibility of dysarthric speech: Implications for assessment and treatment planning. *American Journal of Speech-Language Pathology*, 5(1), 55-66.