ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) | 978-1-6654-0540-9/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICASSP43922.2022.9746627

REPETITION ASSESSMENT FOR SPEECH AND LANGUAGE DISORDERS: A STUDY OF THE LOGOPENIC VARIANT OF PRIMARY PROGRESSIVE APHASIA

R'mani Haulcy[1], Katerina Placek[2], Brian Tracey[2], Adam Vogel[3], James Glass[1]

[1]Massachusetts Institute of Technology, [2]Takeda Pharmaceutical Company,
[3]University of Melbourne, [3]Redenlab Inc

[1]{rhaulcy, glass}@mit.edu, [2]{katerina.placek, brian.tracey}@takeda.com, [3]vogela@unimelb.edu.au

ABSTRACT

Impaired repetition is a characteristic of several speech and language disorders, including certain variants of Primary Progressive Aphasia (PPA). People with the logopenic variant of PPA (lvPPA) can present with impaired repetition abilities and repetition tasks can be used to distinguish lvPPA speakers from healthy controls. In this paper, we propose a novel technique for quantifying the quality of repetition in speech recordings and demonstrate the utility of the technique by using it to distinguish between healthy speakers and lvPPA speakers. We train several classifiers on features extracted from the repetition recordings. The best classifier distinguishes the lvPPA speakers with impaired repetition from the healthy speakers with 85.7% accuracy and classifies all healthy speakers with perfect accuracy. Although we evaluate the method on lvPPA detection, we believe that the method has potential utility for a range of tasks and speech disorders where repetition occurs.

Index Terms— Primary Progressive Aphasia, Cognitive Impairment, Repetition Assessment, Speech Processing

1. INTRODUCTION

It is no secret that speech can be used to detect cognitive impairment in patients with dementia and other neurocognitive disorders [1, 2, 3, 4, 5, 6, 7]. Spoken repetition tasks have been used for decades to detect cognitive and language impairment in children and adults. Non-word repetition tasks have been used to detect language impairment in children [8, 9, 10, 11] and word/sentence repetition tasks have been used to distinguish healthy controls from patients with various forms of Primary Progressive Aphasia (PPA) and Alzheimer's Disease [12, 13, 14, 15, 16]. In previous research, repetition scores have been manually assigned to each speaker (e.g. by computing how many syllables a speaker said correctly) by clinicians and graduate students that were trained to assess the performance of the patients on certain tasks. This process can be biased by the scorer's expectations [17] and can be tedious to complete. For this reason, an automatic way of quantifying repetition quality could be beneficial to the research and clinical communities.

Previous research has explored ways of detecting repetition in audio. Early work on unsupervised pattern discovery in speech were able to find repeating word-like sequences in speech signals without prior knowledge of the words or the language being spoken [18, 19]. These techniques involved using spectral distance matrices, segmental dynamic-time-warping (SDTW) and graph-based clustering methods to identify reoccurring sequences. We were motivated by this work to develop a repetition detection method that could be used to quantify the quality of repetition in a speech recording.

In this paper, we describe our repetition detection method and evaluate its ability to distinguish subjects with the logopenic variant of Primary Progressive Aphasia (lvPPA) from healthy subjects. The speech of lvPPA subjects is characterized by a decline in word retrieval, poor repetition, and phonemic paraphasias [20]. Speakers with poor repetition abilities are unable to repeat other people's speech correctly, often producing phonological speech errors. For this reason, recordings of lvPPA subjects completing repetition tasks are particularly illuminating when compared to the recordings of healthy speakers completing the same tasks. Our work differs from previous work by providing an approach for computing a metric for repetition quality without the need for manual evaluation and without needing to know how many syllables/words are present beforehand. In the following sections, we explain and demonstrate the feasibility of our approach by training classification models using the extracted metrics as inputs.

2. DATASETS

The first dataset used for analysis and classification was the Crowdsourced Language Assessment Corpus (CLAC) [21], a speech corpus that consists of audio recordings for several speech and language tasks from 1,832 speakers that were presumed healthy. As part of the data collection protocol, speakers completed a repetition task in which they were instructed to repeat multisyllabic words ("artillery", "catastrophe", and "impossibility") five times in succession. The repetition recordings for a subset of the speakers in the CLAC corpus were used as healthy controls during analysis and classification.

The second dataset consisted of speech from subjects with various forms of Frontotemporal Dementia (FTD) and PPA completing speech tasks similar to those found in the CLAC corpus, including the repetition tasks. The FTD dataset has been used in previous research to explore which speech characteristics are most salient for the detection of the behavioral variant of FTD (bvFTD) [5]. More information about the language assessment of the subjects in the dataset can be found in [5]. The repetition recordings associated with the lvPPA subjects were used for analysis and classification.

3. SIGNAL PROCESSING FOR REPETITION DETECTION

Each repetition waveform was processed in several steps. First, silence was detected and removed from the beginning and end of each recording. This was accomplished with Pydub [22] with a minimum silence length of 250 ms, and a silence threshold set to the volume of the recording in dB relative to full scale (dBFS) minus 16.

The next step was to extract 13 Mel Frequency Cepstral Coefficients (MFCCs) from each waveform, using a window length of 25 ms and analysis rate of 10 ms. The MFCCs were used as the basis for computing distance matrices described in the next section.

3.1. Self-Distance Matrices

In the general SDTW framework, distance matrices correspond to frame-level distances between two speech waveforms, with each element being a distance between two individual frames. In the repetition task however, each waveform contains multiple occurrences of the same word or syllable, so a (square) self-distance matrix is an appropriate representation to capture the self-similarity between successive repetitions. For this work, we use the Euclidean distance metric to compute frame-level MFCC distances. In Figure 1, an example of a self-distance matrix for a recording with unimpaired repetition (top) versus the matrix for a recording with impaired repetition (bottom) can be seen. In both cases, the matrix diagonal is zero (blue), since we are comparing a recording to itself. However, in the unimpaired repetition matrix, we can also see diagonal-like stripes at regular intervals. These "stripes" correspond to low-distance alignments of successive repetitions. The first off-diagonal corresponds to matching successive repetitions in the waveform (e.g., 1v2, 2v3, 3v4, 4v5), the second off-diagonal corresponds to matching repetitions spaced two repetitions apart (1v3, 2v4, 3v5), and so on. In this way, the off-diagonal structure neatly summarizes the distances between each pair of repetitions.

In contrast to the unimpaired repetition, a poorly spoken repetition recording will not exhibit the same degree of off-diagonal structure. Aside from the main diagonal, we would expect to see more random distribution of distances, corresponding to Euclidean distances between random speech frames. In the impaired repetition example, we see some degree of off-diagonal structure, but it is clearly weaker and more sporadic than the unimpaired repetition example.

The self-distance matrix is a useful representation of the self-similarity between repeating speech patterns. It has the advantage of being agnostic to the chosen word, syllable or even language being spoken, and requires no task-specific training.

3.2. Normalized Diagonal Sum Profile

In order to determine the optimal alignment between successive repetitions, an algorithm such as SDTW should be used [18]. In doing so, we can establish an optimal warping path and an associated alignment cost. In our initial work however, we chose to approximate the warp by assuming an unimpaired repetition alignment would be nearly diagonal, and that the alignment cost could be reasonably represented by the normalized sum of distances along the diagonal.

As a way of characterizing the overall self-distance matrix, we transformed it into a one-dimensional profile, with each element corresponding to a normalized sum of distances along a particular diagonal. An example of a diagonal sum profile is shown in Figure 2, which shows the profile for the unimpaired repetition and impaired repetition examples from Figure 1. The profile is plotted as a function of time, which represents the offset in seconds between two alignments (1s corresponds to 100 frames) in the self-distance matrix. From the sum profile, we can easily identify good alignments, as they correspond to local minima. For an unimpaired repetition (Figure 2a), the minima occur at regular intervals and have consistent magnitude and drop from the average profile value. For an impaired repetition (Figure 2b), the minima are more sporadic and are rather insignificant compared to the average profile value.

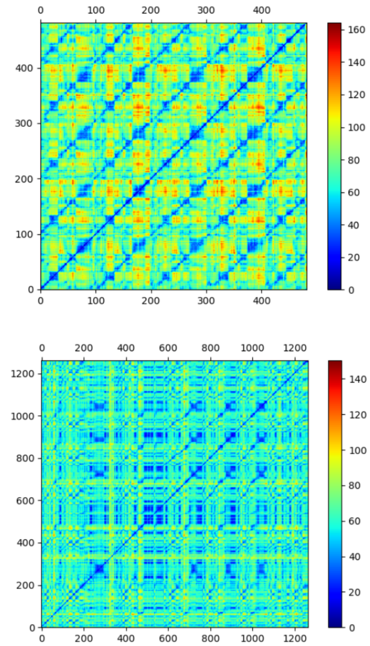


Fig. 1: Self-distance matrices for an unimpaired (top) and impaired (bottom) repetition of five instances of the word "catastrophe".

3.2.1. Sum Profile Features

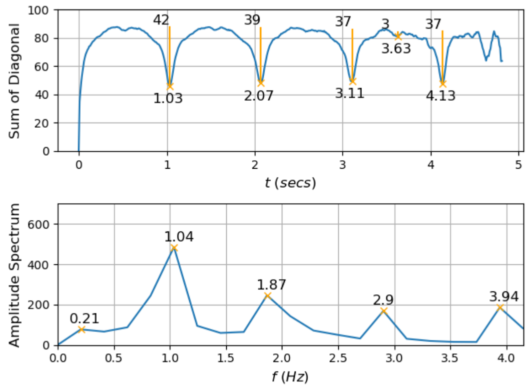
Although the self-distance matrix or the sum profile could have been used directly as a feature representation for the repetition classification model, we wanted to extract a more compact set of features due to the lack of data from lvPPA speakers. Since information about the local minima in the sum profile seemed important, we extracted information about the offset time and magnitude of the local minima. To accomplish this, the find_peaks function in Scipy's signal processing library [23] was used to find the minima in each sum profile. An example of the detected minima can be seen in Figure 2. The detected time offsets are shown beneath each minima and the associated range (the magnitude drop of the local minima) are shown above the lines representing the magnitude.

For each recording, the first three local minima were found and the minima with the largest range was used to represent the recording, along with the corresponding offset time. Each speaker had at least two repetition recordings. After each recording had an associated minima range and time, the minima with the smallest range was selected as the representation for the speaker. In other words, each speaker was represented by their worst repetition, since we expected that healthy speakers would perform well in all their recordings whereas struggling speakers might not.

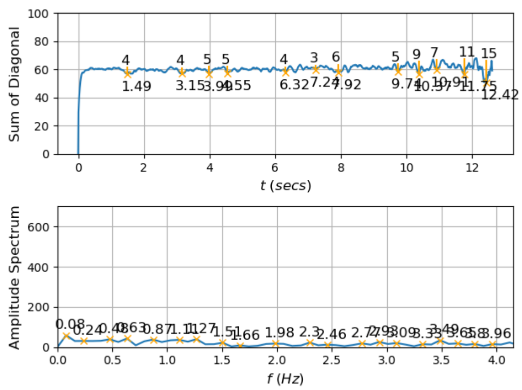
3.2.2. Fourier Analysis Features

As an alternative to the direct feature extraction of local minima information on the sum profile, we also examined the use of a Fourier analysis. We hypothesized that applying a discrete Fourier Transform (DFT) to the sum profile could help us identify the regularity of the local minima structure in the sum profile.

Example DFTs can be seen in Figures 2a and 2b (bottom). In Figure 2a, there are clear DFT peaks at several frequencies. Scipy's



(a) Unimpaired Repetition Example



(b) Impaired Repetition Example

Fig. 2: The diagonal sum profile and associated spectral amplitudes for the unimpaired (a) and impaired (b) repetitions of Figure 1.

find_peaks function was also used to find the peaks in the DFT. As can be seen in the Figure, the first harmonic occurs at 1.04 Hz, which, given the DFT resolution, approximately corresponds to the regular minima in the sum profile.

The DFT in Figure 2a also shows that significant harmonics have higher amplitude values (in contrast to Figure 2b). This suggests that the frequency and amplitude values associated with the harmonics in the DFT may provide a useful representation for analyzing repetitions. For this reason, we extracted the frequency and amplitude for the largest harmonic in the DFT and used that to represent the repetition quality in each recording. As before, we represented each speaker by the smallest amplitude from all of their recordings (representing the speaker's worst performance) and used the frequency and amplitude of that peak to represent the speaker.

4. ANALYSIS

The processing steps described in Section 3 were applied to the audio recordings of healthy and lvPPA speakers and used to extract four features that were described previously (sum profile local minima time and range, DFT first harmonic frequency and amplitude). Since the CLAC corpus contains a wide demographic, a subset of speakers older than 44 was selected as the healthy subset in order to match the lvPPA speaker age range. The data thus consisted of recordings from 354 healthy speakers (aged 53.3 ± 8 years; 199 female speakers) and

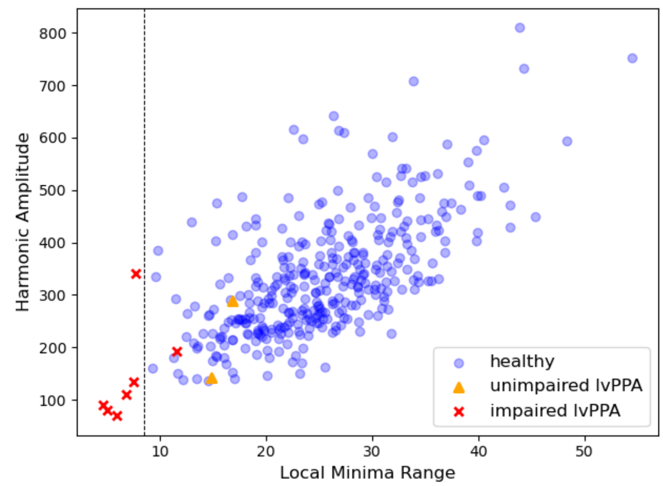


Fig. 3: A scatter plot of the sum profile local minima range and DFT harmonic amplitude for healthy and lvPPA speakers.

9 lvPPA speakers (aged 64.1 ± 7 years; 5 female speakers).

We experimented with different combinations of the four features to see which were the most salient for distinguishing between speakers with unimpaired and impaired repetitions. One approach that we used to make predictions about which combination of features would be most salient was to plot one feature as a function of another for each speaker and visualize the separation. Figure 3 shows one such visualization that plots the local minima range as a function of the harmonic amplitude for lvPPA and healthy speakers.

While lvPPA speakers are known to have difficulty with repetition, there are some that perform the task well. To better analyse the results, a speech and language pathologist independently scored all lvPPA speakers on the multisyllabic word repetition task in terms of speed, precision/accuracy and consistency, before assigning each speaker an overall score. These scores were used to divide the lvPPA speakers into those with relatively unimpaired repetition abilities, and those with relatively impaired repetition abilities.

In Figure 3, the unimpaired lvPPA speakers tend to fall among the cluster of healthy speakers while all but one impaired lvPPA speaker is separated from the healthy speakers. The one impaired lvPPA speaker that is among the healthy speakers is dysfluent throughout the recordings to a lesser degree than the other impaired lvPPA speakers. For this reason, we think it's reasonable for that speaker to be close to the border between the clusters.

4.1. Classification

In order to explore how useful the features we extracted were for distinguishing between unimpaired and impaired repetition recordings, we trained five classification models on several different combinations of features. The five classifiers that we trained were the Linear Discriminant Analysis (LDA) classifier, the Decision Tree (DT) classifier, the K-Nearest Neighbors (KNN) classifier, the Random Forest (RF) classifier, and the Support Vector Machine (SVM) classifier. Because of the small size of the data, we used Leave-One-Subject-Out (LOSO) to train the models. Due to space limitations, we only share the classification results for the KNN classifier trained on all the different combinations of features (Table 1), and we also share the best LOSO results for all five models and the combination of

Features	Acc.	Spec.	Sens.	Imp. Sens.
[MR]	**0.992**	**1**	**0.667**	**0.857**
[HA]	0.989	1	0.556	0.714
[MT, MR]	**0.992**	**1**	**0.667**	**0.857**
[MT, HF]	0.975	1	0	0
[MR, HA]	0.989	1	0.556	0.714
[HF, HA]	0.989	1	0.556	0.714
[HF, MR]	0.989	1	0.556	0.714
[MT, MR, HF, HA]	0.989	1	0.556	0.714

Table 1: KNN LOSO results for healthy vs. lvPPA speakers with different feature combinations (MT: Minima Time, MR: Minima Range, HF: Harmonic Frequency, HA: Harmonic Amplitude).

Classifier	Features	Acc.	Spec.	Imp. Sens.
LDA	[HF, MR]	0.983	1	0.429
DT	[HF, HA]	0.986	0.994	0.857
KNN	[MR], [MT, MR]	**0.992**	**1**	**0.857**
RF	[MT, MR], [HF, MR]	**0.992**	**1**	**0.857**
SVM	[MR], [HF, MR]	**0.992**	**1**	**0.857**

Table 2: The best LOSO results for each classifier.

features that the models were trained on (Table 2).

Each model output a prediction of "healthy" or "lvPPA". Since there was a large disparity in the number of healthy speakers versus lvPPA speakers, we report the specificity (the percentage of healthy speakers that were correctly classified as healthy) and sensitivity (the percentage of lvPPA speakers that were correctly classified as lvPPA) results in addition to the average accuracy across all splits. For sensitivity, we were particularly interested in the sensitivity for the impaired lvPPA speakers, as we expected the unimpaired lvPPA speakers to be misclassified as healthy. For this reason, we report the impaired lvPPA sensitivity (the percentage of impaired lvPPA speakers that were classified as lvPPA) in both tables. Table 2 shows that the best-performing classifiers had an average accuracy of 0.992, a specificity of 1, and an impaired lvPPA sensitivity of 0.857.

5. DISCUSSION

The classification results demonstrate the feasibility of our approach by distinguishing between speakers with unimpaired and impaired repetition (healthy and impaired lvPPA speakers). Only one impaired lvPPA speaker was misclassified as healthy. The misclassified speaker is the impaired lvPPA speaker that is among the healthy speakers in Figure 3, which accounts for the misclassification. The misclassified speaker was also early in their disease progression (0 years since diagnosis when the task was completed), which likely made their repetition impairment less severe compared to the other impaired lvPPA speakers who were all further along in the progression of the disease (the other impaired lvPPA speakers had an average of 3 years since diagnosis).

5.1. Combination Of Features

Table 2 shows that the MR feature is present in the feature combinations of all but one classifier when the best performance for that classifier is achieved. For two classifiers (KNN and SVM), using only the MR feature achieves the best performance (Table 2). We can conclude therefore that MR is an important feature for quantifying repetition in audio, perhaps the most important feature. However, only two classifiers achieved the best performance using only MR features. The other classifiers used a combination of MR and MT or HF. This suggests having a feature that represents the time the repetition occurred is best for good performance across classifiers.

5.2. Benefits Of Our Approach

There are several benefits associated with our approach. The first is that our approach is agnostic to accent and language. Since an utterance is being compared to itself when the distance matrix is computed, the approach does not need to be augmented based on what language someone speaks or what accent they have. Our approach simply tries to capture and measure repeating sounds. This means that the word the speaker repeats does not matter.

Our approach can also potentially be used to identify repeating sounds that are not words, like repeating environmental sounds or non-word repetition tasks, like the popular pataka task that is used to study patients with Parkinson's disease [24]. Our approach can be applied to a variety of problems in addition to the detection of language impairment. Our method is also agnostic to the data collection method that is used. Recordings can be captured in a lab setting, where a proctor is present (e.g. FTD dataset). Recordings can also be collected by the speaker themselves without a proctor present, via phone or computer (e.g. CLAC dataset). The task is simple enough for people to complete on their own from anywhere.

Variation in the recording environment is also not a problem. The concern associated with having recordings that have been collected in vastly different environments using different microphones is alleviated by the fact that self-distance matrices are collected and distances are not being computed between the recordings with vastly different recording environments. Lastly, our approach does not depend on the amount of input data we have and can be applied to a single recording. This is useful in situations where data tends to be limited (e.g. when working with patient data in the health domain).

6. CONCLUSIONS

In this paper, we present a novel approach for measuring the quality of repetition in a recording. We demonstrate the feasibility of this approach by using it to distinguish between healthy and lvPPA speakers using classification. We discuss the many benefits of our approach, including the fact that it is a general repetition measurement approach that can be applied to research problems in a variety of areas. We envision our approach being used to create an application that allows speakers to record themselves completing a simple repetition task (e.g. "repeat the word 'Artillery' five times") and get a repetition score in real time. In the future, we would like to explore how repetition score changes over time. We anticipate using this approach to measure changes in the quality of repetition over time, thus measuring disease progression. We would also like to have a baseline that directly uses speech features as input for disorder assessment while also exploring how the number of repetitions affects our method. Our approach is minimally invasive, inexpensive, and uses recordings that are quick/easy to record, making it appealing for use in clinical trials. The language-independent nature of the approach also makes it potentially useful for global clinical trials.

7. ACKNOWLEDGEMENTS

The authors would like to thank Millennium Pharmaceuticals, Inc. (a subsidiary of Takeda Pharmaceuticals) for funding this research.

8. REFERENCES

[1] Laureano Moro-Velazquez, Jorge A Gomez-Garcia, Julian D Arias-Londoño, Najim Dehak, and Juan I Godino-Llorente, "Advances in parkinson's disease detection and assessment using voice and speech: A review of the articulatory and phonatory aspects," *Biomedical Signal Processing and Control*, vol. 66, pp. 102418, 2021.

[2] Catarina Botelho, Francisco Teixeira, Thomas Rolland, Alberto Abad, and Isabel Trancoso, "Pathological speech detection using x-vector embeddings," *arXiv preprint arXiv:2003.00864*, 2020.

[3] R'mani Haulcy and James Glass, "Classifying Alzheimer's disease using audio and text-based representations of speech," *Frontiers in Psychology*, vol. 11, pp. 3833, 2021.

[4] Sofia de la Fuente Garcia, Craig Ritchie, and Saturnino Luz, "Artificial intelligence, speech, and language processing approaches to monitoring alzheimer's disease: A systematic review," *Journal of Alzheimer's Disease*, , no. Preprint, pp. 1–27, 2020.

[5] Adam P Vogel, Matthew L Poole, Hugh Pemberton, Marja WJ Caverlé, Frederique MC Boonstra, Essie Low, David Darby, and Amy Brodtmann, "Motor speech signature of behavioral variant frontotemporal dementia: Refining the phenotype," *Neurology*, vol. 89, no. 8, pp. 837–844, 2017.

[6] Thea Nygaard Grimstvedt, Jeanette Ullmann Miller, Marleen Regina van Walsem, and Kristin J Billaud Feragen, "Speech and language difficulties in huntington's disease: A qualitative study of patients' and professional caregivers' experiences," *International Journal of Language & Communication Disorders*, 2021.

[7] Adam P Vogel, Christopher Shirbin, Andrew J Churchyard, and Julie C Stout, "Speech acoustic markers of early stage and prodromal huntington's disease: a marker of disease onset?," *Neuropsychologia*, vol. 50, no. 14, pp. 3273–3278, 2012.

[8] Chris Dollaghan and Thomas F Campbell, "Nonword repetition and child language impairment," *Journal of Speech, Language, and Hearing Research*, vol. 41, no. 5, pp. 1136–1146, 1998.

[9] Katharine Graf Estes, Julia L Evans, and Nicole M Else-Quest, "Differences in the nonword repetition performance of children with and without specific language impairment: A meta-analysis," 2007.

[10] Vera F Gutiérrez-Clellen and Gabriela Simon-Cereijido, "Using nonword repetition tasks for the identification of language impairment in spanish-english-speaking children: Does the language of assessment matter?," *Learning Disabilities Research & Practice*, vol. 25, no. 1, pp. 48–58, 2010.

[11] Susan Ellis Weismer, J Bruce Tomblin, Xuyang Zhang, Paula Buckwalter, Jan Gaura Chynoweth, and Maura Jones, "Nonword repetition performance in school-age children with and without language impairment," *Journal of Speech, Language, and Hearing Research*, vol. 43, no. 4, pp. 865–878, 2000.

[12] Cristian E Leyton, Sharon Savage, Muireann Irish, Samantha Schubert, Olivier Piguet, Kirrie J Ballard, and John R Hodges, "Verbal repetition in primary progressive aphasia and alzheimer's disease," *Journal of Alzheimer's Disease*, vol. 41, no. 2, pp. 575–585, 2014.

[13] Sladjana Lukic, Maria Luisa Mandelli, Ariane Welch, Kesshi Jordan, Wendy Shwe, John Neuhaus, Zachary Miller, H Isabel Hubbard, Maya Henry, Bruce L Miller, et al., "Neurocognitive basis of repetition deficits in primary progressive aphasia," *Brain and language*, vol. 194, pp. 35–45, 2019.

[14] Michael F Bonner, Sharon Ash, and Murray Grossman, "The new classification of primary progressive aphasia into semantic, logopenic, or nonfluent/agrammatic variants," *Current neurology and neuroscience reports*, vol. 10, no. 6, pp. 484–490, 2010.

[15] Maria Luisa Gorno-Tempini, Argye E Hillis, Sandra Weintraub, Andrew Kertesz, Mario Mendez, Stefano F Cappa, Jennifer M Ogar, JD Rohrer, Steven Black, Bradley F Boeve, et al., "Classification of primary progressive aphasia and its variants," *Neurology*, vol. 76, no. 11, pp. 1006–1014, 2011.

[16] M-Marsel Mesulam, Christina Wieneke, Cynthia Thompson, Emily Rogalski, and Sandra Weintraub, "Quantitative classification of primary progressive aphasia at early and mild impairment stages," *Brain*, vol. 135, no. 5, pp. 1537–1553, 2012.

[17] Emma Louise Clark, Catherine Easton, and Sarah Verdon, "The impact of linguistic bias upon speech-language pathologists' attitudes towards non-standard dialects of english," *Clinical Linguistics & Phonetics*, vol. 35, no. 6, pp. 542–559, 2021.

[18] Alex S Park and James R Glass, "Unsupervised pattern discovery in speech," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 16, no. 1, pp. 186–197, 2007.

[19] Aren Jansen, Kenneth Church, and Hynek Hermansky, "Towards spoken term discovery at scale with zero resources," in *Eleventh Annual Conference of the International Speech Communication Association*, 2010.

[20] Matthew L Poole, Amy Brodtmann, David Darby, and Adam P Vogel, "Motor speech phenotypes of frontotemporal dementia, primary progressive aphasia, and progressive apraxia of speech," *Journal of Speech, Language, and Hearing Research*, vol. 60, no. 4, pp. 897–911, 2017.

[21] R'mani Haulcy and James Glass, "CLAC: A Speech Corpus of Healthy English Speakers," in *Proc. Interspeech 2021*, 2021, pp. 2966–2970.

[22] James Robert, Marc Webbie, et al., "Pydub," 2018.

[23] Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan J. van der Walt, Matthew Brett, Joshua Wilson, K. Jarrod Millman, Nikolay Mayorov, Andrew R. J. Nelson, Eric Jones, Robert Kern, Eric Larson, C J Carey, İlhan Polat, Yu Feng, Eric W. Moore, Jake VanderPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E. A. Quintero, Charles R. Harris, Anne M. Archibald, Antônio H. Ribeiro, Fabian Pedregosa, Paul van Mulbregt, and SciPy 1.0 Contributors, "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python," *Nature Methods*, vol. 17, pp. 261–272, 2020.

[24] Pedro Gómez-Vilda, Andrés Gómez-Rodellar, Daniel Palacios-Alonso, and Athanasios Tsanas, "Performance of monosyllabic vs multisyllabic diadochokinetic exercises in evaluating parkinson's disease hypokinetic dysarthria from fluency distributions," 2021.