Georgia Institute of Technology: Undergraduate Research Opportunities Program

Analysis of Speech Patterns in Alzheimer's Disease Literature Review

Dr. Elliot Moore II

Connor Hawley

I. Introduction

Given recent advances in computing and mobile devices, interest has been mounting regarding how these advances might be used to automatically collect and analyze data in a clinical setting, i.e. aid in the diagnosis of diseases. This review focuses specifically on the qualitative and quantitative relationships between speech and Alzheimer's disease pathology. Prosodic and acoustic aspects of speech in patients with Alzheimer's disease are of particular interest, and each study will be fully analyzed with attention to nuances in patient population, methodology, speech metrics, and metric analysis.

II. Background

A. Alzheimer's Disease Speech Pathology

Alzheimer's disease is most notorious for its cognitive effects on its victims, but linguistic deficits are still considered a secondary symptom [1]. Linguistic deficits would be more useful in discriminating those afflicted by Alzheimer's disease from normal elderly adults, but language deficits are a more primary symptom of other dementias, i.e. aphasia. This is usually presented as a dichotomy, where aphasic patients are unable to express themselves but understand their surroundings, whereas patients with Alzheimer's disease are generally able to speak normally, but they do not have a good understanding of their environment.

In Wasilewski et al. [1], an interesting hypothesis for the development of speech issues in Alzheimer's patients was discussed: stating that whereas most aphasia patients experience focal brain damage around Broca's area, Alzheimer's disease speech impediments originate from multifocal neuropathological lesions. All forms of dementia appear to share speech impediments, including "considerably reduces fluency of speech, hearing comprehension deficits, and the reduced ability to maintain spontaneous speech, grammar, intonation, and pronunciation, and ability to repeat." [1] However, "[Alzheimer's disease] patients show increasing difficulty in reading, naming, and defining. Moreover, they tend to have problems with word choice, and their utterances become 'bizarre' and devoid of content." The early stage of communication impediment in Alzheimer's disease is briefly discussed. Patients "have difficulty finding the right word" and "persistent repetition of the same statement". Patients will also substitute semantically related, but incorrect, words in sentences.

Of quantitative interest, Wasilewski et al. [1] mention that "patients tend to use low-frequency words," and perhaps, this could be quantified with natural language processing in a simple word frequency analysis. Another important detail: phonology is largely unaffected until the late stages of Alzheimer's disease, but communication will grow increasingly nonsensical. It would appear from [1] that there is little qualitative evidence for speech deficits in Alzheimer's disease, versus other dementias.

Chapman et al. [2] touched on the differences in speech in patients Alzheimer's disease and aphasia. After comparing performance in these areas across groups, Chapman et al. [2] drew a few conclusions: "The aphasia group received significantly lower scores than both the Alzheimer's disease and normal control groups on linguistic formulation. The patients with Alzheimer's disease exhibited significant difficulties on the pragmatic domain of drawing inferences as compared to the aphasia and normal control groups. For the majority of tasks, there were no significant group differences on the communicative intentions measure. However, there were significant differences on the domain of language-information balance for the aphasia and Alzheimer's disease groups on most tasks as compared to the normal control group." Thus, this work paints a picture of the qualitative differences between aphasia and Alzheimer's disease groups, and how they compare to normal control groups.

Chapman et al. [2] claim that most methodologies for characterizing communication have been designed to diagnose aphasia, leaving a large bias that makes identifying other dementing disorders more challenging. In general, "[p]atients with aphasia are commonly noted to communicate better than they talk and patients with Alzheimer's disease are observed to talk better than they communicate (Holland, 1982)." Essentially, linguistic competence is not a fair metric of communicative competence. Chapmen et al. [2] examined both linguistic function and "pragmatics" with discourse measures to further elucidate patients' conditions.

From Chapman et al. [2], aphasia and Alzheimer's disease patients share some verbal communication deficits, including comprehension deficits, word-finding deficits, verbosity, empty content, and incoherent responses. These deficits show up the most in discourse measures. However, "language disturbances are pervasive even in the earliest stages of Alzheimer's disease (Bayles et al., 1989; Henderson, 1996; Hodges et al., 1996; Kontiola et al., 1990)." It might be worthwhile investigating the cited papers to learn more about the language disturbances that occur in the early stages of Alzheimer's disease, but the authors do provide some detail: "impairments are localized predominantly in the semantic domain marked by semantic paraphasic errors (Bayles, 1982) and by a reduced vocabulary which is particularly evident on confrontation naming tasks (Hier et al., 1985; Martin & Fedio, 1983)." It is also worth noting that syntax is known to be simplified in Alzheimer's disease, according to "(Hier et al., 1985; Ulatowska et al., 1988)." Chapman et al. [2] postulate that the abnormalities in language between the two groups may be attributable to the fact that there are a limited number of ways language can break down. Without these metrics for distinguishing Alzheimer's disease from aphasia, they can still be distinguished over time because they "have separate etiologies, divergent progressions, and disparity between cognitive and linguistic ability." The paper hopes to distinguish "whether the discourse similarities are superficial, reflecting different underlying mechanisms in each patient population."

After laying the groundwork for the study, Chapman et al. [2] turn to propose a "conceptual framework for discourse production." Chapman et al. [2] go on to define discourse as "a unit of language used to convey a message where the expression of the message is governed by (a) linguistic facility in expressing the information (b) pragmatic aspects that relate to paralinguistic phenomena, and (c) cognitive function required to effectively manipulate information to communicate a message." These aspects of discourse are notably not exhaustive but were selected based on previous work, and compared to previous work, this paper's model expands significantly on the pragmatic aspect of discourse, which is difficult to quantify, but is still useful to consider qualitatively if pragmatic measures of speech become practical to quantify in the future.

In a more detailed discussion of Chapman et al.'s [2] results, it is worth noting that in the pragmatic domain of communicative intentions, the Alzheimer's disease patients had a significantly more difficult time understanding that they were supposed to generate a story from a single frame picture presented to them. The most significant way to distinguish Alzheimer's disease patients from aphasia and normal control patients was in the pragmatic domain of drawing inferences. The Alzheimer's disease patients were generally unable to draw inferences between the texts presented to them and real-world knowledge, demonstrating a general lack of understanding relative to other subjects. aphasia patients were much better at selecting the meaning of a text in a multiple-choice format, whereas Alzheimer's disease patients continued to struggle.

Overall, "the present findings support the conclusion of Blanken et al. (1987) that the linguistic disturbances in Alzheimer's disease do not represent an impairment of lexicalization, but rather an inadequacy in accessing the appropriate meaning through impaired inferencing." [2]

B. Metrics for Discriminating Alzheimer's Disease from Healthy Populations

Meilán et al. [3] claim to be able to discriminate patients with Alzheimer's disease from adults with normal aging with an accuracy of 84.8% using measures of speech, specifically: "variations in the percentage of voice breaks, number of periods of voice, number of voice breaks, shimmer (amplitude perturbation quotient), and noise-to-harmonics ratio, characterize people with Alzheimer's disease." These characteristics were extracted using the Praat software.

Meilán et al. [3] begin by describing some of the early characteristics of Alzheimer's Disease. Notably, logopenic progressive aphasia is described as being indicative of Alzheimer's disease at the early stage, where patients speak slowly and have difficulty producing speech because of their difficulty finding words, claimed by a paper by "Kemper et al.", which could be worth investigation. Other works are cited as having evidence for some prosodic speech features being able to discriminate Alzheimer's disease patients from normal aging. Once source found higher hesitation ratios in patients with Alzheimer's disease, and "Tosto et. al [18] found prosodic impairment in Alzheimer's disease: features of speech such as emphasis placed on certain syllables, changes in tempo or timing, and differences in pitch and intonation." Meilán et al. [3] also speak of a work that examined spoken language in patients with mild cognitive impairment. These works observed that patients with mild cognitive impairment had higher percentages of voiceless segments, mean duration of pauses, standardized phonation time, and verbal rate. Though phonation and articulation are preserved in patients with Alzheimer's disease, prosody, temporal, and acoustic measures have discriminating power. Meilán et al. [3] actually suggest a mapping of quantitative measures to describe qualitative differences between control population and patients with early Alzheimer's disease: "temporal aspects of the speech sample and interruption of sound for alterations in rhythm; analysis of the fundamental frequency, periods of voice, and fluctuation in frequency for slowness, phonological errors and articulatory apraxias, and fluctuation in amplitude for loudness." Essentially, the study carried out analysis on these parameters to elucidate how well these parameters describe patients with Alzheimer's disease.

As with any work, this analysis comes with strings attached, as described in the methodology. The most glaring hole in this work is its lack of comparison to any other neurodegenerative disorder, e.g. aphasia. The only participants in the study were control and early Alzheimer's disease, which calls into question the study's ability to discriminate Alzheimer's disease from Aphasia patients. Another notable difference is the study was conducted with "a professional Fostex FR-2 LE recording equipment, with 24-byte resolution, a sampling rate of 48 kHz, and an AKG D3700S cardioid microphone. The microphone was placed on a stand 8 cm from the participant at an angle of 45 degrees to the patient's mouth to decrease aerodynamic noise from the mouth." The participants also read a controlled, familiar passage and were shown the script in the exact same way. After calculating a number of parameters on the speech, linear discriminant analysis was applied to unearth the most significant measures.

Using the linear discriminant analysis, Meilán et al. [3] were able to explain 100% of the data variance with five factors: percentage of voice breaks, number of periods, number of voice breaks, shimmer, and noise-to-harmonics ratio. For each of the parameters, a Wilks' lambda was ascribed, which the closer to zero the Wilks' lambda is, the more a variable contributed to discriminating the data. Based on this, the noise-to-harmonics ratio contributed the most to discriminating the data. Even after all these steps, Meilán et al. [3] could only correctly classify 84.8% of the patients with their methods.

In discussion, Meilán et al. [3] interestingly remark that the fundamental frequency and pitch modulation of study participants with Alzheimer's disease was lower than the control group, a result also supported by Horley et al. [4] as well, perhaps pointing to a physiological indicator of early Alzheimer's disease.

However, Meilán et al.'s [3] conclusions differ from the conclusions of similar work, perhaps because of methodology, i.e. the patients read a familiar passage, so prosody is more likely to be preserved. Meilán et al.'s [3] list three studies and what parameters they each found to be most relevant. Meilán et al. [5] aimed to characterize the impairment of Alzheimer's disease associated with the onset of the disease. Meilán et al. [5] examined 21 elderly patients with a series of acoustic characteristics. The most significant result was that 34% of the variance in the data was able to be explained with increases in the fraction of voiceless segments in patient speech, i.e. Meilán et al. [5] were some of the first to demonstrate a strong relationship between voicing and Alzheimer's disease. Roark et. al [6], in a task where the participants were asked to re-tell a story immediately after it had been told to them, found that the standardized pause rate, phonation rate, and many linguistic complexity measures were useful in discriminating between healthy elderly participants and participants with mild cognitive impairment. However, Singh et al. [7] reported that the mean duration of pauses, standardized phonation time, and verbal rate were useful in discriminating between healthy elderly participants and patients with Alzheimer's disease." It appears that different methodology will yield different discrimination tactics, but it is worth noting that each of the studies mentioned all discriminate with features measuring very similar phenomena.

In another work, Bucks et al. [8] set out to analyze patients with probable dementia of Alzheimer's type with control subjects of similar age, gender, and education in a quantitative, easily replicable way. Quantitative analysis of patient speech was nothing new, but at its time few works had applied quantitative analysis to conversational, spontaneous speech for DAT patients. Bucks et al. [8] went for a truly quantitative approach: all conversations were continued until the patients had spoken 1000 words, and then these conversations were recorded and transcribed, where linguistic features were extracted.

The linguistic features measured were the following: noun rate, pronoun rate, verb rate, adjective rate, clause-like semantic unit rate (all per 100 words), and three lexical richness measures: type token ratio, Brunét's Index, and Honoré's statistic.

When discussing the background and motivation for the work, Bucks et al. [8] highlighted an interested dimension to Alzheimer's disease, onset: "Interestingly they found that object naming was worse in late onset DAT, but spontaneous speech was worse in early onset DAT participants." The takeaway from this is twofold: patients with speech issues are more likely to be early onset, and late onset Alzheimer's disease will be more difficult to distinguish from normal aging by speech analysis alone. This was probably what was implied in Khodabaksh et al. [9], that the part of the brain cortex that processes language abilities is one of the earliest parts to be affected by the disease.

Bucks et al. [8] uses similar methodology to previous works that quantitatively analyzed aphasia, but now applies these methods to Alzheimer's disease patients. Bucks et al. [8] had 24 participants, with 16 healthy controls and 8 individuals diagnosed with moderate-severe Alzheimer's disease. Differences in age and gender between the two groups was insignificant, but the years of education between the two groups proved to be significant, which was corrected for statistically, even though other works have not found the metrics used in Bucks et al. [8] to be correlated with education level. To collect the speech data, the participants were asked a series of open-ended questions and encouraged to talk about themselves.

As for results of Bucks et al. [8], great significance ($p < 0.01$) was seen for noun rate, pronoun rate, Honoré's Statistic, Brunét's Index, and type token ratio between the control group and Alzheimer's disease patients, and adjective rate and verb rate showed significance ($p < 0.05$). Another metric employed, CSU-rate, turned out not to be significant. In particular, noun and pronoun count were especially useful in discriminating control from Alzheimer's disease, similar to what was replicated in Khodabakhsh et al. [9]

Other than correlations and significance tests, Bucks et al. [8] also performed principal components analysis (PCA) and linear discriminant analysis (LDA) on the data. In the principal components analysis, the first principal component represented

64.1% of the variance, and appeared to represent lexical richness and phrase making ability. Linear discriminant analysis was carried out with all 8 variables and was able to reach 100% accuracy. Fortunately, Bucks et al. [8] realized this severe overfitting problem, and with the leave-one-out method, they achieved an average 87.5% accuracy.

Overall, the main takeaways from Bucks et al. [8] were that control patients performed much better on lexical richness measures, and patients with Alzheimer's disease used significantly more pronouns and many fewer nouns.

Unlike other studies, Khodabakhsh et al. [9] set out to predict Alzheimer's disease with just recordings of unstructured conversations, analyzing both linguistic features, from transcripts of the conversations, and prosodic features, mathematically extracted from the audio files themselves. Overall, prosodic features were much better predictors of whether a patient was Alzheimer's disease or healthy. Khodabakhsh et al. [9] support their approach by claiming that the part of the brain that processes language is one of the first areas to be affected by Alzheimer's disease, citing Bucks et al. [8].

For each feature, predictive power was evaluated by using the feature in support vector machines (SVM), nearest-neighbor (NN) classifiers, naïve Bayes classifiers, and classification trees (CTrees). For three prosodic features, accuracy of 80% was achieved. The features were: silence ratio, average silence count, and average continuous word count.

The 18 extracted linguistic features investigated will not be full listed here, but it is worth noting that many of the features were exceptional because of the Turkish language's structure. Most of the features were focused on capturing the speaker's richness of vocabulary and expression, and a few were focused on cohesiveness.

For prosodic features, 20 were extracted, and they fell into three categories: voice activity-related features, articulation-related features, and rate of speech-related features. Each feature was calculated over the locution of the subject, being defined as "the total response period of the subject which is the sum of all the subject's speech turns." [9]

As for Khodabakhsh et al.'s [9] data, speech was recorded for 32 patients and 51 age and education-matched control subjects, but 4 patients' data were not used because of how little they spoke, leaving 28 patients in the study. The data were collected over about 10 minutes of conversation recorded using a high-quality microphone. Speech segmentation and transcription were done manually. When testing the accuracy of the classifiers, 10 random subsamplings would be used, with equal representation from both populations, and average performance was reported. Because of the small sample size, a leave-one-out strategy was used for training the classifiers. Khodabakhsh et al. [9] then analyze the results, controlling for one variable at a time.

In age-controlled results, the only part of speech features that were significant were the noun and pronoun-related features, and most other linguistic features were not significant. Prosodic features were mostly still significant, even when controlling for age. In education-controlled results, it is again seen that pronoun and noun measures are still some of the only significantly discriminating linguistic features. Voice activity-related features remained strongly discriminating even when controlling for education.

Overall, Khodabakhsh et al. [9] found that a couple of linguistic features and many prosodic features were useful in discriminating Alzheimer's disease from the control group. However, using multiple features together, Khodabakhsh et al. [9] were not able to generalize better.

III. Future Study

A. Interesting Leads from Cited Material

Bucks et al. [8] cites a work, "Boller et al., 1991", that demonstrated the severity of language impairment is predictive of the rate of decline the patient suffers, which could be another interesting implication for predicting Alzheimer's disease with voice.

The difference in symptoms between late and early onset Alzheimer's disease patients led Bucks et al. [8] to discussion on breaking Alzheimer's disease up into two separate diseases. However, it might be worth polling what the contemporary opinion on this matter is, as even in this work, Bucks et al. [8] caution that there is not consensus on the relationship between language impediment and the onset of the disease.

B. Unknown Aspects of Speech's Relationship to Alzheimer's Disease

Another interesting quantitative observation from Meilán et al. [3] is that patients with Alzheimer's disease exhibited a higher number of "voice breaks", which are imperceptible to the human ear, but in large number may be noted as "bubbles or tremors" in the voice. It would be interesting to see if this phenomenon is significant in patients in earlier stages of the disease, and at different onsets.

In Khodabakhsh et al. [9], the most useful features with the highest discriminating power were the silence ratio, average silence count, and average continuous word count. Since the patients in this study were late stage, it would be incredibly interesting to extract some of the same kind of information and see how well the data could discriminate for early stage Alzheimer's disease or MCI patients.

For all the worked cited, most had a relatively small sample size, and even fewer were longitudinal studies. Any longitudinal, large-scale studies on the speech of patients with Alzheimer's disease and healthy elderly adults would greatly elucidate current understanding. For prosodic and acoustic features, in previous works, very high-quality recording equipment was used, and it would also be interesting to investigate the minimum quality audio needed to preserve the information useful for discriminating Alzheimer's disease.

IV. References

[1] T. P. Wasilewski and H. Kachaniuk, "Speech disturbances in Alzheimer's disease," *Neuropsychiatria i Neuropsychologia*, vol. 3, pp. 93–98, 2016.

[2] S. B. Chapman, A. P. Highley, and J. L. Thompson, "Discourse in fluent aphasia and Alzheimers disease: Linguistic and pragmatic considerations," *Journal of Neurolinguistics*, vol. 11, no. 1-2, pp. 55–78, 1998.

[3] J. J. G. Meilán, F. Martínez-Sánchez, J. Carro, D. E. López, L. Millian-Morell, and J. M. Arana, "Speech in Alzheimers Disease: Can Temporal and Acoustic Parameters Discriminate Dementia?" *Dementia and Geriatric Cognitive Disorders*, vol. 37, no. 5-6, pp. 327–334, 2014.

[4] K. Horley, A. Reid, and D. Burnham, "Emotional Prosody Perception and Production in Dementia of the Alzheimers Type," *Journal of Speech, Language, and Hearing Research*, vol. 53, no. 5, pp. 1132–1146, 2010.

[5] J. J. G. Meilán, F. Martínez-Sánchez, J. Carro, J. A. Sánchez, and E. Pérez, "Acoustic Markers Associated with Impairment in Language Processing in Alzheimer's Disease," *The Spanish Journal of Psychology*, vol. 15, no. 2, pp. 487–494, Jun. 2011.

[6] B. Roark, M. Mitchell, J.-P. Hosom, K. Hollingshead, and J. Kaye, "Spoken Language Derived Measures for Detecting Mild Cognitive Impairment," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 19, no. 7, pp. 2081–2090, 2011.

[7] S. Singh, R. S. Bucks, and J. M. Cuerden, "Evaluation of an objective technique for analysing temporal variables in DAT spontaneous speech," *Aphasiology*, vol. 15, no. 6, pp. 571–583, 2001.

[8] R. S. Bucks, S. Singh, J. M. Cuerden, and G. K. Wilcock, "Analysis of spontaneous, conversational speech in dementia of Alzheimer type: Evaluation of an objective technique for analysing lexical performance," *Aphasiology*, vol. 14, no. 1, pp. 71–91, 2000.

[9] A. Khodabakhsh, F. Yesil, E. Guner, and C. Demiroglu, "Evaluation of linguistic and prosodic features for detection of Alzheimer's disease in Turkish conversational speech," *EURASIP Journal on Audio, Speech, and Music Processing*, vol. 2015, no. 1, Mar. 2015.