Language Analysis of Speakers with Dementia of the Alzheimer's Type

Curry Guinn and Anthony Habash

Department of Computer Science, University of North Carolina Wilmington
guinnc@uncw.edu

Abstract

This research is a discriminative analysis of conversational dialogs involving individuals suffering from dementia of Alzheimer's type. Several metric analyses are applied to the transcripts of the Carolina Conversation Corpus (Pope and Davis 2011) in order to determine if there are significant statistical differences between individuals with and without Alzheimer's disease. Results from the analysis indicate that go-ahead utterances, certain fluency measures, and paraphrasing provide defensible means of differentiating the linguistic characteristics of spontaneous speech between healthy individuals and those with Alzheimer's disease. Several machine learning algorithms were used to classify the speech of individuals with and without dementia of the Alzheimer's type.

Introduction

"The Alzheimer's Association [2009] reports that 13% of Americans over the age of 65 … present with AD [Alzheimer's Disease]. During the course of the disease, individuals lose cognitive-communication skills in predictable ways" (Fried-Oaken et al. 2009). This decline in language facility can result in decreased social interaction and life satisfaction for persons with AD and their caregivers. In particular, persons with AD begin to feel a loss of their personal identity.

Individuals suffering from dementia of Alzheimer's type are characterized as being afflicted with a loss in cognitive and communicative functionality (Bucks 2000). This condition is often reflected within their powers of communication with 88 – 95% of individuals with Alzheimer's (Thompson 1987) portraying some degree of aphasia (language disability) and cognitive failure including the inability to grasp concepts, events of their past, or the ability to recognize individuals.

Spontaneous speech, the focus of this research, is characterized as allowing for no pre-emptive planning or memorization of a response and demanding the highest level of cognitive action and memory contemplation to produce accurate and effective responses. This research attempts to detect quantitative signs of degradation in speech and cognitive capacities within individuals suffering from Alzheimer's.

Speech Pathologies in Dementia of the Alzheimer's Type

A vast majority of individuals with Alzheimer's are characterized by the degradation of their language and cognitive functionality, resulting in significant complications in vocal communication. How to properly analyze and treat individuals with Alzheimer's is an ongoing question. In the majority of cases of individuals diagnosed with Alzheimer's disease, the patient has likely suffered the condition for several years before the symptoms became evident enough for testing (Ray 2007). This makes a comparison of previous cognitive and communication ability within the patient with Alzheimer's prior to the onset of the disease extremely difficult if not impossible. Therefore a broad, general comparison of the communication abilities of individuals with Alzheimer's to those of individuals who do not suffer the condition might prove more accessible in determining signs of cognitive and linguistic degradation.

Computational Models of Speech in Individuals with Alzheimer's

Prior quantitative studies of the pathologies (in general) of speech include the use of pauses, fillers, formulaic speech, restarts, repeats, incomplete statements and speech disfluencies (Davis 2009, Snover 2004, Bortfeld 2001, Yaruss 1998). The degradation of speech capacity exhibited within individuals with Alzheimer's should include these factors. Also, past research suggests that persons with Alzheimer's exhibit less lexical richness (Yaruss 1998).

There is a complication in using these quantitative measures for contrasting individuals with and without dementia of the Alzheimer's type: namely, it can be hard to gauge an individual's pre-existing speech capacity. Also, pre-existing studies (Bucks 2000) which claim effective discrimination of individuals with Alzheimer's and healthy individuals using these computational models have done so using a limited and size-controlled corpus for analysis, as well as using interview-style dialog instead of spontaneous speech. For instance, the Bucks et al. study used 8 individuals with dementia of the Alzheimer's type. The experiment described in this paper uses a substantially larger corpus (n = 31) than previous research.

Literature Review and Analysis Methods

Several studies have used computational methods to examine the speech of individuals with dementia of the Alzheimer's type. Several projects have examined the level of speech degradation in relevance to age, gender, relationships, topics, role, and stages of dementia (Bucks 2000, Bortfeld 2001). Most of this research has focused on two dimensions of speech: lexical richness (Bortfeld 2001, Bucks 2000, Singh 1996, Singh 1997) and/or speech fluency (Yaruss 1998, Singh 2000, Davis 2009, Snover 2004).

Corpus Annotation

The corpus of spontaneous speech conversations involving individuals with Alzheimer's that will be used in this study is the Carolina Conversations Collection (CCC). The CCC Corpus, developed in a partnership between the Medical University of South Carolina and the University of North Carolina Charlotte, was constructed in an effort to provide useful data concerning speech patterns with regards to age, gender, social identities, health and illness stories, and explanatory models of disease (Pope and Dave 2011).

The CCC corpus consists of over 400 transcribed conversations with 125 multiethnic, older individuals suffering from any number of possible conditions that have been individually categorized (Pope and Davis 2011). These conversations were originally recorded in only audio format and eventually transcribed into text using the methods defined by Ten Have (2007). Of those transcripts, we use 80 conversations where one participant had been diagnosed with Alzheimer's disease. Within this subset, there are 33 subjects suffering from Alzheimer's disease with most subjects participating in more than one interview. To compensate for multiple transcripts involving the same person, all data gathered from transcripts involving the same Patient or Interviewer will be averaged. A dialog fragment from the CCC is given in Figure 1.

Interviewer	When you was little did you work on the farm, in the fields?
Subject	Uh huh.
Interviewer	What did you do in the fields?
Subject	Cotton cotton cotton and horses and all kinds of stuff.
Interviewer	Did you mind picking cotton?
Subject	We, we just had to do it.
Interviewer	Yeah Okay. What did you do with the horses?
Subject	Fields, some where they uh, but they uh, but other--
Interviewer	Mmm hmm animals. What other animals were there? Like some pigs and hogs and chickens?
Subject	Um hmm.
Interviewer	Okay. Did y'all grow any types of like corn or something?
Subject	Mmm Hmm.

Figure 1 A fragment of a conversation within the Carolina Conversations Collection. Subject has dementia of the Alzheimer's type. Interviewer is a graduate student eliciting some life events from the subject.

For contrast, we also did a comparative analysis with the Switchboard Corpus, a collection of transcribed, spontaneous telephone conversations between healthy individuals (Godfrey 1992). This corpus was formulated as part of a project by Texas Instruments that was meant to address a growing need for large multispeaker databases of telephone bandwidth speech. Like with the CCC Corpus, conversations for the Switchboard Corpus were originally recorded via audio and then transcribed by hand into text using the methods defined by Ten Have (2007).

Syntactic Modeling

Identifying the parts of speech within a speaker's dialog allows for the break-down and categorization of the cognitive strength and capacity based on their lexical richness. This is achieved using the Natural Language Toolkit, a series of analytical programs designed to systematically parse and tag raw text (Bird 2009). As the patient with Alzheimer's cognitive capacities fail, their ability to grasp facts such as names, places, and actions deteriorates. Past research reports a rise in the use of pronouns, and a decrease in the use of proper nouns, verbs, and adjectives (Singh 1996).

Semantic Modeling and Lexical Richness

Several studies have established that a form of communication degradation caused by the onset of

Alzheimer's is the weakening of an individual's vocabulary (Bortfeld 2001, Bucks 2000, Singh 1997).

Three different forms of linguistic measurement are applied to the corpus to analyze the vocabulary richness employed by the speakers: Type Token Ratio (TTR), Brunét's Index (BI), and Honore's Statistic (HS) (Bucks 2000, Singh 1996). These weighted measurements provide an applied approach to measuring the lexical richness of a dialog by providing algorithms for weighing the significance of unique vocabulary versus word count and total vocabulary.

Type-Token Ratio (TTR) provides a comparison to the total vocabulary used in a dialogue (V) to the total word count (N) of the dialogue.

$$\mathbf{TTR = V/N}$$

However Type-Token Ratio does not account for the variations in word count, which is an uncontrolled factor in spontaneous conversation. Brunét's Index (BI) is unique from Type-Token Ratio in that it attempts to quantify the vocabulary used without considering the word count. Brunét's Index will then likely show a more applicable result when applied.

$$\mathbf{BI = N^{V(-0.165)}}$$

Honore's Statistic (HS) attempts a deeper analysis by accounting for words that are only used once (V_1), indicating a higher lexical richness.

$$\mathbf{HS = (100 \log N) / (1 - V_1/V)}$$

Disfluency Modeling

Distinct from a language's lexicon, disfluencies reflect a failure of concept more than a failure of vocabulary, where the speaker is uncertain, unclear, or doubtful of what they are trying to communicate. Spontaneous speech by its nature is commonly very disfluent (Bortfeld 2001). Our analysis will track several features associated with disfluency.

Filler Words

Fillers are non-word and short phrase utterances that serve a communicative purpose with several possible meanings depending on their place (Bortfeld 2001). When placed at the beginning or end of a dialog they could possibly be "hints" given by the speaker in order to indicate that they had trouble understanding something or that they desire input. They could also serve to indicate that the speaker has misspoken and desires to take back what was said or reword it, which is shown more the case when they occur in the middle of a dialog than at the ends of it (Bortfeld 2001). In general though, they do indicate some form of cognitive lapse, where the speaker fails to communicate properly, hence we theorize that the rate of fillers would be higher for individuals with Alzheimer's. In the following example from a speaker with Alzheimer's disease, we see repeated use of the filler "uh":

> *But, uh, potato bread, we, uh,-- uh- pared the potatoes the night before. And, uh,- then,- uh, cut it all up into little pieces like that.*

Repetitions

Repetition is a notable coping mechanism for cognitive failure where the speaker will repeat a stated word or words in order to allow the information provided within the discourse to become more evident within their cognitive process. Repeats usually occur surrounding small pauses within an utterance where the speaker recaps or reaffirms to themselves what they are trying to communicate and then continue forth by picking up where they left off as in the following example:

> *"That was a very… very intriguing speech we just heard."*

Repeats are a natural coping method for individuals who experience a form of cognitive lapse. As such it is arguable that individuals with Alzheimer's whose cognitive functionality has suffered degradation would be prone to use repeats more often than healthy individuals.

Incomplete Words

Incomplete words are instants within the corpus where the speaker begins to pronounce a word but then inexplicably stops before completing the pronunciation, indicated within the CCC corpus by the use of a dash or tilde at the end of a word as in the following example:

> "Until my Mom got home and uh, as ***gro~*** as I was growing up, it was, I had a great mother."

The cause of this behavior is creditable as a lapse in cognitive ability where the speaker intercepts their thought process before saying something wrong, or it can serve as an indicator that the speaker is unsure of what they are trying to communicate and stop altogether.

Syllables per Minute

The rate of speech is provided within the CCC corpus using the metric of syllables-per-minute for both the Interviewers and the Subjects with Alzheimer's disease. This form of measurement is indicated to be an effective means of comparing the speech capacity of individuals (Yaruss 1998).

Pragmatic Features

Part of our study is to investigate dialog moves that may facilitate conversation with persons with dementia of the Alzheimer type. Pope and Davis (2011) have suggested that the use of short (often two-syllable) go-ahead phrases and increased use of paraphrasing will increase the fluency of dialogs with persons with dementia.

Go-ahead utterances

Go-ahead utterances are instances in dialog in which a speaker provides responses that do not add anything in a conversation beyond a minimal response. Go-ahead utterances usually serve as an indicator by one of the speakers that they either have nothing to input within a conversation or wish for another speaker to continue. Go-aheads also serve as means of validation that a person is listening to what someone is saying, or that they agree or disagree with what is being said as in the following dialog example:

Ms. April: *It was a nonprofit organization.*
Ms. March: *Uh huh.*
Ms. April: *And they didn't provide any of that.*
Ms. March: *Awww.*
Ms. April: *So, all my health coverage is still my …*

Paraphrasing

Paraphrasing is the act of taking what has been previously stated and reusing it in one's own context. People paraphrase all the time, be it something they read or something they heard prior to a situation in which they restate it in their own words. Paraphrasing is also a natural occurrence within conversation, where an idea or topic is passed back and forth between speakers with each adding their own knowledge in part to the matter and commenting on what the other has expressed.

Our analysis will detect three different types of paraphrasing: direct, reflexive, and indirect. An example of a **direct** paraphrase would be to use the same word (perhaps after lemmatization) as the previous speaker. A **reflexive** form of paraphrasing would be cases where the speaker inverts the pronoun that the previous speaker used (for example, using "I" when the previous speaker used "you"). An **indirect** paraphrase of a word can occur when the speaker uses a term related to the previous word, for example, if previous speaker uses the word "farm" and the current speaker uses the word "grow". Our system uses WordNet to compute a distance score based on the semantic distance between two terms. To compute a paraphrase score for a sentence, our system compares each word in an utterance to the words in the previous utterance and computes an average level of paraphrase for each word. If comparing to a previous utterance, we call this backward paraphrasing; if comparing to the next utterance, we call this forward paraphrasing.

Results and Analysis

For each feature, we calculated the average for each subject with Alzheimer's and for each conversational partner (for shorthand, we will call the person *without* Alzheimer's the "Interviewer" even though there is no formal interview being conducted). There are 80 total transcripts with 31 different individuals with Alzheimer's disease and 57 different interviewers.

In Table 1, we present a number of metrics comparing speakers without Alzheimer's disease (Interviewers) and speakers with dementia of the Alzheimer's type (Subjects) along with the p value associated with a T-test comparing the two groups. The part-of-speech metrics are computed per word (in other words, 40.7% of the words spoken by the Interviewers are nouns). Incomplete words, filler phrases, and repeats are also calculated per word. Percentage of go-ahead utterances and forward and backward paraphrasing are averaged per utterance.

Table 1 Metrics contrasting persons without Alzheimer's disease (Interviewers) with subject with Alzheimer's disease

Metric	Interviewers (Mean & SD)	Subject w/ Alz. (Mean & SD)	T-Test (p value)
Noun Rate	0.407 (0.054)	0.397 (0.050)	0.1978
Verb Rate	0.080 (0.017)	0.083 (0.014)	0.2010
Adjective Rate	0.157 (0.013)	0.020 (0.019)	0.1434
Pronoun Rate	0.162 (0.030)	0.179 (0.025)	0.1166
Type-Token Ratio	0.414 (0.140)	0.406 (0.114)	0.3865
Brunet's Index	13.23 (1.896)	13.21 (1.823)	0.4740
Honore's Statistic	615.1 (93.7)	623.2 (93.9)	0.3508
% of Go-Ahead Utt.	**34.88 (15.4)**	**41.18 (16.3)**	**0.0384**
Repeats	**0.0035 (0.004)**	**0.0068 (0.006)**	**0.0014**
Incomplete Words	**0.0067 (0.007)**	**0.0114 (0.010)**	**0.0054**
Filler phrases	**0.012 (0.017**	**0.026 (0.025))**	**0.0702**
Forward Paraphrasing	**0.286 (0.085)**	**0.256 (0.067)**	**0.0492**
Backward Paraphrasing	0.280 (0.087)	0.259 (0.083)	0.1416
Syllables per minutes	**216.1 (42.1)**	**199.0 (70.6)**	**0.0801**

As can be seen from Table 1, part-of-speech was not a good metric for distinguishing between speakers with and without dementia of the Alzheimer's type. This finding contrasts with Singh's earlier work suggesting that persons with Alzheimer's use more pronouns (Singh 2000). Our analysis does indicate a slightly higher use of pronouns; however, the p-value of the T-test makes this statistic inconclusive.

Also, unlike Singh, the semantic complexity of the two speakers was not statistically different as measured by the three lexical richness measures: type-token ratio, Brunet's index, and Honore's statistic. However, when we compared the lexical richness of the dialogs within the CCC to dialogs within the Switchboard corpus, there was a significant difference in Honore's statistic: (615.1 for the CCC corpus; 652.3 for the Switchboard corpus, T-test p-value = 0.0053). We conjecture that the interviewer in the CCC corpus may be trying to roughly match the lexical richness of the person with Alzheimer's; however, we would need a collection of dialogs with the interviewer and persons not afflicted with Alzheimer's disease in order to determine if this may be the case. The CCC corpus does not support this analysis.

The statistically significant differences between the two groups of speakers are related to the use of go-ahead phrases, several fluency measures, and paraphrasing. The Interviewer in the CCC corpus uses significantly more go-ahead phrases than the subject with Alzheimer's. The person with Alzheimer's also exhibits more overt signs of difficulty in fluency with higher rates of incomplete words, filler words, repeated words, and slower rate of speech. The Interviewer also exhibits higher levels of paraphrasing.

Learning Algorithms for Classifying Alzheimer's versus non-Alzheimer's Speech

Using the features that were observed to be statistically significant, several learning algorithms were employed to automatically categorize a speaker's set of utterances based on whether the speaker may or may not have dementia of the Alzheimer's type. We have examined three different classification algorithms: k-nearest neighbors clustering, a decision tree learning algorithm (C4.5), and support vector machines. For all three of these learning algorithms, training and testing was performed using leave-one-out testing (i.e., training on all data but one, testing on that one piece of data, and rotating in this manner for all data points). Scaling of the data was performed for both k-neighbors and the SVM. Table 2 presents these results.

Table 2 Classification of Alzheimer's versus non-Alzheimer's speech using several learning algorithms

Learning Algorithm (31 Alz; 57 non-Alz)	% correct identifying speech of Alzheimer's	% correct identifying speech of non-Alzheimer's	Combined
k-neighbors (k = 1)	41.9	73.7	62.5
k-neighbors (k = 3)	32.3	73.7	59.1
k-neighbors (k=5)	29.0	77.2	60.2
Decision Tree	41.9	100.0	79.5
Support Vector Machine	37.5	94.7	75.0

First, note that the data set is quite unbalanced with almost twice as many non-Alzheimer speakers as Alzheimer's speakers (because the CCC corpus occasionally has the same subject in multiple dialogs with different interviewers). Thus, a k-neighbors algorithm will be particularly vulnerable to this imbalance as *k* increases.

Both the decision tree and support vector machine algorithms achieve high accuracy in classifying speech as non-Alzheimer's. However, all three algorithms have a tendency to classify the speech of persons with Alzheimer's disease as resembling the speech of someone without dementia. Thus, these classifiers have a somewhat high false negative rate, but a low false positive rate. From a diagnosis standpoint, this is not necessarily a bad feature – these classifiers rarely say someone exhibits signs of dementia when, in fact, they do not have dementia.

Conclusions and Future Work

We have shown that a number of features in the speech of persons with dementia of the Alzheimer's type are statistically significantly different from the speech of healthy individuals. Using these features, we have employed a number of classifiers to classify the speech of an individual as either indicating the symptoms of Alzheimer's disease or not. These classifiers had a very lower false positive right – they rarely classified speech as indicating the symptoms of Alzheimer's disease. However, the false negative rate was around 60% for the decision tree and SVM classifiers.

Several issues remain open questions. As noted in our paper, the lexical richness of both participants in the conversation in the CCC corpus was lower than the lexical richness of the participants in the Switchboard corpus. Prior research suggests that persons with dementia of the Alzheimer's type display lower lexical richness than healthy individuals. Are the conversational participants lowering than natural lexical richness to match the lexical richness of the person with Alzheimer's?

Our next main goal is to determine what dialog behaviors enhance the fluency of individuals with Alzheimer's disease. Does the use of go-aheads and paraphrasing by the healthy conversational participant increase the fluency and general conversational ability of the person with Alzheimer's disease? This future study will involve looking at more than just syntactic measures as it will require looking at the relative "success" of each dialog turn.

References

Alzheimer's Association. (2009). *2009 Alzheimer's Disease Facts and Figures.* Downloaded on 4/30/09 from www.alz.org.

Bird, S.; Klein, E. & Loper, E. (2009), *Natural Language Processing with Python: Analyzing Text with the Natural Language Toolkit*, O'Reilly, Beijing.

Bortfeld, H.; Leon, S.; Bloom, J.; Schober, M. & Brennan, S. (2001), 'Disfluency rates in conversation: Effects of age, relationship, topic, role, and gender', *Language and Speech* **44**(2), 123--147.

Bucks, R. S.; Singh, S.; Cuerden, J. M. & Wilcock, G. K. (2000), 'Analysis of spontaneous, conversational speech in dementia of Alzheimer type: Evaluation of an objective technique for analysing lexical performance', *Aphasiology* **14**(1), 71-91.

Davis, B. & Maclagan, M. (2009), 'Examining pauses in Alzheimer's discourse', *American journal of Alzheimer's Disease and other dementias* **24**(2), 141--154.

Fried-Oken, M., Rowland, C., Baker, G., Dixon, M., Mills, C., Schultz, D., and Oken, B. (2009). The Effect of Voice Output on AAC-Supported Conversations of Persons with Alzheimer's Disease. *ACM Transactions on Accessible Computing*, Vol. 1, No. 3, Article 15.

Pope, C. & Davis, B. (2011), 'Finding a balance: The Carolinas Conversation Collection', *Corpus Linguistics and Linguistic Theory* 7(1), 143--161.

Ray, S.; Britschgi, M.; Herbert, C.; Takeda-Uchimura, Y.; Boxer, A.; Blennow, K.; Friedman, L. F.; Galasko, D. R.; Jutel, M.; Karydas, A.; Kaye, J. A.; Leszek, J.; Miller, B. L.; Minthon, L.; Quinn, J. F.; Rabinovici, G. D.; Robinson, W. H.; Sabbagh, M. N.; So, Y. T.; Sparks, D. L.; Tabaton, M.; Tinklenberg, J.; Yesavage, J. A.; Tibshirani, R. & Wyss-Coray, T. (2007), 'Classification and prediction of clinical Alzheimer's diagnosis based on plasma signaling proteins', *Nat Med* **13**(11), 1359--1362.

Singh, S. (1996), 'Neural Networks In Spontaneous Speech Assessment Of Dysphasic Patients', Ph.D. Dissertation, University of the West of England, Bristol, UK.

Singh, S. & Bookless, T. (1997), 'Analysing spontaneous speech in dysphasic adults', *International Journal of Applied Linguistics* **7**(2), 165--181.

Snover, M.; Dorr, B. & Schwartz, R. (2004), A lexically-driven algorithm for disfluency detection, *in* 'Proceedings of HLT-NAACL 2004: Short Papers', pp. 157--160.

Ten Have, P. (1999), *Doing conversation analysis*, Sage Publications Ltd.

Thompson, I. M. (1987). Language in dementia. International Journal of Geriatric Psychiatry, 2,
145-161.

Yaruss, J. (1998), 'Real-time analysis of speech fluency: Procedures and reliability training', *American Journal of Speech-Language Pathology* **7**(2), 25.