AUTOMATIC SPEECH ASSESSMENT FOR APHASIC PATIENTS BASED ON SYLLABLE-LEVEL EMBEDDING AND SUPRA-SEGMENTAL DURATION FEATURES

Ying Qin†, *Tan Lee*† *and Anthony Pak Hin Kong*‡

† Department of Electronic Engineering, The Chinese University of Hong Kong, Hong Kong SAR, China
‡ Department of Communication Sciences and Disorders, University of Central Florida, Orlando, FL, USA
†yingqin@link.cuhk.edu.hk, †tanlee@cuhk.edu.hk, ‡antkong@ucf.edu

ABSTRACT

Aphasia is a type of acquired language impairment resulting from brain injury. Speech assessment is an important part of the comprehensive assessment process for aphasic patients. It is based on the acoustical and linguistic analysis of patients' speech elicited through pre-defined story-telling tasks. This type of narrative spontaneous speech embodies multi-fold atypical characteristics related to the underlying language impairment. This paper presents an investigation on automatic speech assessment for Cantonese-speaking aphasic patients using an automatic speech recognition (ASR) system. A novel approach to extracting robust text features from erroneous ASR output is developed based on word embedding methods. The text features can effectively distinguish the stories told by an impaired speaker from those by unimpaired ones. On the other hand, a set of supra-segmental duration features are derived from syllable-level time alignments produced by the ASR system, to characterize the atypical prosody of impaired speech. The proposed text features, duration features and their combination are evaluated in a binary classification experiment as well as in automatic prediction of subjective assessment score. The results clearly show that the text features are very effective in the intended task of aphasia assessment, while using duration features could provide additional benefit.

***Index Terms*—** Speech assessment, aphasia, Cantonese, ASR, word embedding

1. INTRODUCTION

Aphasia refers to a type of acquired language impairment resulting from focal brain damage, which is commonly caused by stroke. Symptoms of aphasia adversely affect multiple modalities of language: listening, speaking, reading and writing. The impairment could also span over various levels and components of the language system, including phonology, lexicon, syntax, and semantics [1]. Speakers with aphasia have difficulties in recalling names of objects and/or putting words together into sentences [2]. Assessment is a critical process in speech and language rehabilitation. It aims at determining the type and/or severity of impairment and identifying specific aspects of disability. Conventionally assessment of aphasic patients is carried out by trained speech pathologists, based on subjective evaluation of various abilities of language communication [3]. Subjective assessment of speech from aphasic patients is a challenging task because it requires not only clinical knowledge about the disease but also good understanding of relevant linguistic and cultural background. There is clearly a strong demand for reliable and effective methods of objective speech assessment for aphasic patients.

This paper presents our recent effort on applying state-of-the-art automatic speech recognition (ASR) technology to automatic assessment of impaired speech from Cantonese-speaking aphasic patients. A dedicated Cantonese ASR system is developed to facilitate acoustical and linguistic analysis of patients' speech, which are elicited through story-telling tasks on prescribed topics. In our earlier work [4], it was shown that supra-segmental duration features derived by forced alignment with a general-purpose ASR system could capture the impaired fluency of speech and effectively distinguish an impaired speaker from unimpaired ones. In the present study, the duration features used for automatic speech assessment are obtained from the time alignment of input speech given by the ASR system. On the other hand, we propose to extract robust text features from the ASR output using word embedding techniques. A continuous bag-of-word (CBOW) model [5] is trained to generate syllable-level embeddings from the stories told by unimpaired subjects. Given an input story (impaired or unimpaired), the ASR system produces a hypothesized syllable sequence, from which a story-level vector is obtained by assembling (averaging) the constituting syllable embeddings. The proposed text features aim to capture the discrepancies between an impaired story and unimpaired ones. Fig. 1 shows the proposed system of automatic speech assessment. The ASR system adopts the DNN-HMM architecture and is trained using domain-matched and speaking-style-matched speech from unimpaired subjects. The assessment of aphasia severity is formulated as either a classification problem or a regression problem.

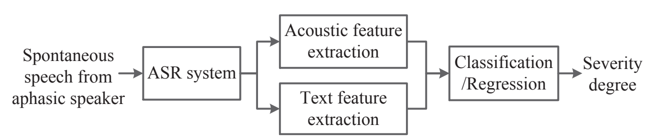


Fig. 1. The proposed approach to automatic speech assessment for aphasic patients.

In Sections 2 and 3, the Cantonese AphasiaBank database and the dedicated ASR system are described respectively. The proposed text features and acoustic features are explained in Section 4. The results of classification and regression experiments with the proposed features are presented and discussed in Section 5.

2. CANTONESE APHASIABANK

Cantonese is a major and influential Chinese dialect spoken by tens of million of people in Hong Kong, Macau, Southern China, and overseas Chinese communities. Cantonese AphasiaBank is a large-scale multi-modal database jointly developed by the University of

Central Florida and the University of Hong Kong, aiming to support both fundamental and clinical research on Cantonese-speaking aphasia population [6]. The corpus contains audio recordings of spontaneous speech from 104 aphasic subjects and 149 unimpaired subjects. All of them are native speakers of Cantonese. The speech recordings were elicited by following the AphasiaBank protocol, with adaptation to local Chinese culture [7][8]. Each speaker was requested to complete 8 narrative tasks, including 4 picture descriptions, 1 procedure description, 2 story telling and 1 personal monologue. Except the personal monologue, the speech produced in each task is expected to be about a specific topic, which is referred as a "story". The audio signals were acquired using a head-worn condenser microphone and a digital recorder. The sampling frequency was 44.1 kHz. The speech data were manually transcribed into a sequence of Chinese characters. Fillers, unintelligible speech and non-speech sounds were represented by special symbols. For the development of Cantonese ASR system, the Chinese characters were converted into Cantonese syllables using a pronunciation lexicon [9].

The aphasic subjects in the database went through a standardized assessment process known as the Cantonese Aphasia Battery (CAB) [10]. The CAB consists of a number of sections measuring fluency, information content, comprehension, repetition and naming abilities of the subject [10]. Each sub-test results in a numerical score, which is given by a trained speech pathologist according to prescribed criteria. The sum of the sub-test scores is named the Aphasia Quotient (AQ). The value of AQ (0 - 100) is regarded as an indication of overall severity of language impairment. Lower AQ value implies higher degree of severity. For the 104 aphasic subjects in the AphasiaBank database, the AQ values range from 11.0 to 99.0.

3. ASR SYSTEM FOR APHASIA ASSESSMENT

Like Mandarin, Cantonese is a monosyllabic and tonal language. Each Chinese character is spoken as a monosyllable carrying a specific tone. Cantonese syllable is described by the *Initial-Final* structure [9]. The *Initial* is typically a consonant, while the *Final* consists of a vowel followed by an optional coda. There are 20 *Initial*s and 53 *Final*s in Cantonese, from which over 600 legitimate *base syllable*s can be formed.

A dedicated Cantonese ASR system is developed to deal with the topical stories in the Cantonese AphasiaBank. The system is built on a pronunciation lexicon that covers 630 *base syllable*s. Both the acoustic models and language models of the ASR system are trained with the speech recordings from 101 unimpaired speakers (about 12.6 hours) in the Cantonese AphasiaBank database. The acoustic models adopt the standard DNN-HMM architecture, which is implemented using the Kaldi Speech Recognition Toolkit [11]. Each of the *Initial*s and *Final*s of Cantonese is represented by a hidden Markov model (HMM) with 3 emission states. A DNN with 6 hidden layers and 1024 neurons per hidden layer is trained to estimate the posterior probabilities of the tri-phone HMM states. The input features are 440-dimensional fMLLR transformed from MFCCs with a contextual window of 11 frames. The language models are syllable bi-grams trained with the orthographic transcriptions of all training speech data. It is noted that the spontaneous speech in the database contains many inserted filler words, lengthened/repeated initial consonants, as well as frequent occurrences of para-verbal sounds like laughing and sighing. In the ASR system, five of the most common non-content sounds are modeled by dedicated HMMs and included as part of the acoustic models.

The ASR system described above is expected to suit the specific domain of the Cantonese AphasiaBank, in terms of both speaking style and spoken content. The system performance is evaluated using the speech data of 7 tasks (except personal monologue) from 17 unimpaired speakers (not overlapping with the training speakers) and 82 aphasic speakers. For the unimpaired speakers, the overall syllable error rate (SER) is 18.24%, which is considered fairly good for spontaneous speech. The SER for individual speakers varies from 7.82% to 36.73%. For the 7 speech tasks, the average SER varies between 16.36% and 22.24%.

The overall SER for aphasic speakers is 48.08%. A significant portion of the recognition errors are due to the occurrences of many unintelligible speech sounds, which are not modeled by the ASR system. These sounds could be recognized as Cantonese syllables and counted as insertion errors. The SER per speaker varies greatly from 14.93% to 96.5%, reflecting the highly diverse types and degree of language impairment. Indeed, the 82 aphasic speakers consist of 52 Anomic, 6 Transcortical sensory, 12 Transcortical motor, 8 Broca's, 1 Isolation, 2 Wernicke's and 1 Global aphasia. Their AQ values range from 27.0 to 99.0. Across the 7 speech tasks, the ASR system shows similar performance, i.e., SER of 42.96% to 45.24%.

4. FEATURE EXTRACTION

4.1. Text Features: Syllable-level Embedding Features

We aim at robust text features that can be derived from the erroneous ASR output. These features must be able to reflect the topic-specific linguistic content of a given story and differentiate the story told by an impaired speaker from those by unimpaired ones. By inspecting the speech transcription of a few selected stories from impaired speakers, we find that the number of topic-specific keywords tends to decrease as the severity of aphasia increases. Subjects and objects are often missing in impaired sentences.

We propose to derive a compact story-level vector representation using the word embedding techniques. First of all, a CBOW model is trained using syllable-level transcriptions of all stories from the unimpaired speakers. The CBOW model adopts the architecture of a three-layer feedforward neural network [5]. It is trained to make prediction of a central word at the output layer, given a set of left and right contextual words at the input layer. While the word representations are discrete (one-hot) at both the input and output layers, a continuous-value vector representation for each word in the vocabulary can be obtained as the corresponding row of the trained weight matrix between the input and hidden layers. Refer to [5] for the detailed explanation of the CBOW model.

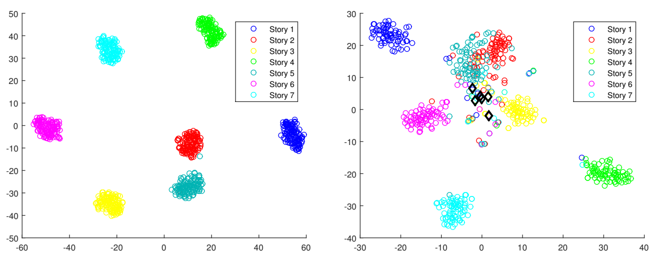
To alleviate the problems of data sparsity and out-of-vocabulary words, syllable-level embedding is applied in the CBOW training. Each *base syllable* of Cantonese is regarded as a word here and the word vectors actually refer to the syllable vectors. For a given story, the story-level vector representation is obtained by taking the average of all syllable vectors in accordance to the transcription of the story. Since the impaired speech in the AphasiaBank is often highly ungrammatical, the CBOW model and the simple averaging method are considered appropriate in that the word order is not taken into account.

The word2vec Toolkit [12] developed by Google is used to implement the CBOW model. The syllable-level transcriptions of speech from 118 unimpaired speakers are used as training data, which contains about $183,000$ syllables and covers 523 unique items. The size of contextual window is set to be 6 (syllables). The dimension of syllable vectors is set to be 50. The size of negative sampling set is set as 5 and the sub-sampling threshold is set as 10^{-3} to improve the accuracy of the learned vectors of rare words [13].

4.1.1. *Text features of impaired speech: from transcription*

We first examine the effectiveness of story vectors in representing the contents of stories, and make a comparison between unimpaired and impaired speech. For each of the 7 stories, we obtain 118 story vectors from unimpaired speakers and 82 story vectors from impaired speakers. The t-distributed stochastic neighbor embedding (t-SNE) algorithm [14] is applied to allow visualization of the 50-dimensional story vectors in a 2-dimensional space, as illustrated in Fig. 2(a) (unimpaired speech) and Fig. 2(b) (impaired speech). Different colors are used to illustrate different story topics.

In the case of unimpaired speech, it is clearly seen that different story topics are well separated from each other. While for impaired speech, a noticeable degree of overlap is seen. Specifically, the story vectors from a patient with Broca's aphasia (AQ: 42.0), marked by black diamonds "$\Diamond$" in the Fig. 2(b), can hardly be separated. The speech of this patient is found to contain mostly function words and few topic-specific content words, such that all stories have similar content.



(a) unimpaired speech (b) impaired speech

Fig. 2. 2-dimensional visualization of story vectors computed from unimpaired speech and impaired speech.

We expect that language impairment leads to significant discrepancy in the story vectors and the degree of such discrepancy is related to severity of the impairment. In this regard, the following two types of text features are proposed to quantify the degree of language impairment of an aphasic speaker:

Inter-story feature: No. of mis-clustered story vectors — As discussed above, the story vectors on different topics tend to be confused with each other if the subject suffers language impairment. The degree of such confusion can be used to indicate the severity of impairment. Given the 7 story vectors from an impaired subject, these vectors are first pooled with the story vectors (7×118) from unimpaired subjects. K-means clustering is applied to divide the pooled data into 7 classes, and the number of mis-clustered story vectors for the impaired subject is counted. Statistical analysis shows that the number of mis-clustered story vectors has strong negative correlation with the AQ value of the subject. Hence we propose to use the number of mis-clustered story vectors as a type of inter-story text feature for assessment of language impairment. The feature value is divided by 7 to be normalized in the range of 0 to 1;

Intra-story feature: similarity w.r.t unimpaired speech — For each of the 7 story topics, a topic vector can be obtained by taking the mean of the respective story vectors from all unimpaired subjects. It is a compact representation of the topic, against which the impaired speech would be compared. Given an impaired subject, we first compute the cosine similarity between each of his/her story vectors and the respective topic vector. The overall similarity is defined as the average of cosine similarity measures over the 7 stories. It is used as an intra-story text feature of assessment of language impairment. Statistical analysis confirms a significant positive correlation between this feature and the subject's AQ value.

4.1.2. *Text features of impaired speech: the effect of ASR*

The text features described above are motivated by the observation on Fig. 2, where the story vectors are computed from manual transcriptions. In practice, the story vectors have to be derived from the error-prone ASR output. As shown in Table 1, we divide the 82 aphasic speakers into two groups based on the SER: below 50% versus over 50%. For each group, we compare the average deviation of text features computed based on the ASR output transcriptions from those based on manual transcriptions. For the inter-story feature, a positive deviation means that the number of mis-clustered story vectors tend to be over-estimated based on the ASR output, hence the degree of impairment would also be over-estimated. For the intra-story feature, a negative deviation indicates over-estimated discrepancy between impaired and unimpaired content, and would lead to the over-estimation of impairment severity.

For the group of low-SER subjects, the average deviation of inter-story feature is 0.02, which is equivalent to about 0.14 mis-clustered story. The average deviation of intra-story feature (cosine similarity) is also very small. For subjects with high SER, the text features computed from ASR output deviate more noticeably. The average deviation of inter-story feature is equivalent to an over-count of 1.27 mis-clustered story (out of 7). In short, poor ASR performance accounts for the over-estimation of impairment severity.

Overall speaking, the proposed text features can effectively quantify the difference of impaired speech from unimpaired one, despite the large variation of ASR performance.

Table 1. Deviation of text features computed from ASR output from those from manual transcriptions. The aphasic subjects are divided into two groups according to SER. The average deviation across all speakers in the group is shown.

SER		SER $\leq$ 50%	SER $>$ 50%
No. of speakers		49	33
Deviation of feature values	Inter-story	0.020	0.182
	Intra-story	0.002	-0.092

4.2. Acoustic Features: Supra-segmental Duration

In addition to the recognized syllable sequence, syllable-level time alignment can be obtained as part of the ASR output. Supra-segmental duration features are derived from the time alignment information given by the ASR decoder. Based on previous studies on acoustical analysis of aphasia [4] [15], we define 13 duration features that are related to fluency, speaking rate, etc. The LASSO regression method [16] [17] is applied to select the most effective features and reduce the feature dimension for the subsequent classification process. The LASSO regression analysis suggests a ranking order for the 13 candidate features. By jointly considering this ranking order and the correlation between each feature and the AQ value, we select the following 5 duration features:

Non-speech-to-speech duration ratio — The duration ratio between non-speech part and speech part, computed over all stories from the subject. The non-speech part covers filler words and silence segments longer than 0.5s. The speech part consists of all syllables recognized by the ASR system;

Average duration of silence segments — The average duration of all silence segments longer than 0.5s, which are computed over all stories from the subject;

Average duration of speech segments — Average duration of all speech segments. Each speech segment is between two silence segments;
Ratio of silence segment count to syllable count — The ratio between the number of silence segments to the number of syllables in all stories from the subject;
Syllable count per second — The average number of syllables per second in all stories. It essentially measures the speaking rate of the subject.

5. EXPERIMENTAL RESULTS AND DISCUSSION

The effectiveness of the proposed features on speech assessment for aphasic subjects is evaluated in two different experiments. The first experiment is a binary classification task, in which the degree of speech impairment is assessed as being severe or not severe. In the second experiment, the possibility of automatically predicting the AQ score is investigated.

5.1. Binary Classification of Aphasia Severity

While the AQ value is an indication of the severity of language impairment, there is no standard cut-off point for differentiating normal and aphasic subjects. In previous studies by speech pathologists, the suggested cut-off values were around 93 to 96 [10][18]. In our experiment, a binary classification task is designed by dividing the 82 aphasic subjects into two groups: High-AQ (AQ $>$ 90) versus Low-AQ (AQ $<$ 90). The main consideration is to have balanced number of subjects in the two groups, so that statistical pattern classification algorithms can be applied. The cut-off value of 90 should not be considered as a clinically meaningful standard.

There are 35 and 47 subjects in High-AQ and Low-AQ groups respectively. Classification experiments are carried out with three classification algorithms, namely binary decision tree (BDT), random forest (RF) and support vector machine (SVM) with polynomial kernel. The leave-one-out cross validation strategy is adopted. In each fold of validation, the feature vector from one of the subjects is reserved as test data and the remaining feature vectors are used for training. The 2-dimensional text feature vector, 5-dimensional acoustic feature vector and their combination are evaluated separately. The classification results, measured in terms of average F1 score, are summarized in Table 2.

It can be seen that the RF classifier performs the best among the three classifiers. Text features are found to be more effective than the acoustic features. These two types of features are complementary to each other. The best performance is an average F1 score of 0.903 with the RF classifier on combined features. With this classifier, the recalls for the Low-AQ and High-AQ groups are 89.4% (42/47), 88.6% (31/35) respectively.

Table 2. Classification results (in terms of average F1 score) on text features, acoustic features and combined features.

	BDT	RF	SVM
Text features only (2)	0.851	0.896	0.841
Acoustic features only (5)	0.792	0.821	0.789
All features (7)	**0.891**	**0.903**	**0.874**

5.2. Automatic Prediction of AQ

Automatic prediction of AQ is done by performing regression on the feature vectors. Two regression models are constructed with the approaches of linear regression (LR) and random forest (RF) respectively. Leave-one-out cross validation strategy is applied to evaluate the effectiveness of text features, acoustic features and the combined features. Table 3 shows the Spearman correlations between the predicted AQ (AQ_p) and the reference AQ (AQ_r) values obtained from the two regression models. It is found that the text features are more useful for prediction of AQ than the acoustic features. The best performance of prediction shows a correlation of 0.839, given by the RF regression model. With this model, 50% (41/82) of the aphasic subjects have the prediction errors $|AQ_p - AQ_r| \leq 5.0$, and 74.4% (61/82) have the prediction errors smaller than 10.0. The results suggest that the proposed text features combined with duration features are fairly reliable in predicting the subject assessment scores.

Table 3. Correlations of predicted AQ with reference AQ values.

	LR	RF
Text features only (2)	0.821	0.820
Acoustic features only (5)	0.651	0.655
All features (7)	0.816	**0.839**

To investigate the possible causes of prediction errors, we analyse two typical aphasic subjects for whom the prediction errors are greater than 10.0. The SER for these two subjects are on the low side, i.e., 25.96% and 29.42%, meaning that the text features computed from their speech are not much affected by the ASR errors (see Table 1). The AQ value of the first subject is found to be underestimated ($AQ_p = 82.7$ vs. $AQ_r = 96.1$). We find that there are frequent repetitions of filler word in his speech, such that the computed intra-story feature is significantly smaller than other subjects with similar subjective AQ. It must be noted that the subject AQ is a composite score measuring multiple aspects of language impairment. Some of the sub-tests, e.g. comprehension, naming, are not relevant to the speech impairments that the proposed features aim to characterize. It is possible that the subject performed very well in most of the sub-tests and thus a high combined score, but was not able to handle the story-telling tasks.

The other selected aphasic subject has an under-estimated impairment, i.e., $AQ_p = 86.2$ versus $AQ_r = 73.2$. The intra-story feature of this subject is relatively good, meaning that the content of his stories are quite relevant to the given topics. However, the syntactic constructions of his spoken sentences are often incomplete and confused. This reveals the major limitation of the CBOW model. Since the word order is not captured by the model, if the subject's speech contains a good number of topic-specific words, the resulted story vectors would not be differentiable from unimpaired speech. Further investigation is needed to explore other text features to characterize syntactic impairment of aphasic patients' speech.

6. CONCLUSIONS

It is believed that state-of-the-art ASR technology can find a more important role in automatic assessment of pathological speech. For impaired speech from aphasic patients, one of the major challenges is due to the spontaneous nature of speech, in which language impairment could be easily confused with unproficiency of speaking. In this paper, we demonstrate the possibility of extracting robust text features and acoustic features from erroneous ASR output, which can be used to differentiate an impaired speaker from unimpaired ones and predict the subjective assessment score. Toward developing clinically usable system for automatic assessment, we need to improve the ASR accuracy and enhance the current design of speech features. Lastly, it is utmost important to enhance the speech database by including more patients with different types of aphasia.

7. REFERENCES

[1] H. Adam, "Dysprosody in aphasia: An acoustic analysis evidence from palestinian arabic," *Journal of Language and Linguistic Studies*, vol. 10, no. 1, pp. 153–162, 2014.

[2] National Aphasia Association, "Aphasia definitions." [Online]. Available: https://www.aphasia.org/aphasia-definitions/

[3] American Speech-Language-Hearing Association, "How is aphasia diagnosed?" [Online]. Available: https://www.asha.org/public/speech/disorders/Aphasia/

[4] T. Lee, A. Kong, V. Chan, and H. Wang, "Analysis of auto-aligned and auto-segmented oral discourse by speakers with aphasia: A preliminary study on the acoustic parameter of duration." *Procedia, social and behavioral sciences*, vol. 94, pp. 71–72, 2013.

[5] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient estimation of word representations in vector space," *arXiv preprint:1301.3781*, 2013.

[6] A. P.-H. Kong and S.-P. Law, "The Cantonese Aphasiabank." [Online]. Available: http://www.speech.hku.hk/caphbank/search/

[7] B. MacWhinney, D. Fromm, M. Forbes, and A. Holland, "Aphasiabank: Methods for studying discourse," *Aphasiology*, vol. 25, no. 11, pp. 1286–1307, 2011.

[8] A. P.-H. Kong, S.-P. Law, C. C.-Y. Kwan, C. Lai, and V. Lam, "A coding system with independent annotations of gesture forms and functions during verbal communication: Development of a database of speech and gesture (dosage)," *Journal of nonverbal behavior*, vol. 39, no. 1, pp. 93–111, 2015.

[9] P. Ching, T. Lee, W. Lo, and H. Meng, "Cantonese speech recognition and synthesis," *Advances in Chinese Spoken Language Processing*, pp. 365–386, 2006.

[10] E. M. Yiu, "Linguistic assessment of Chinese-speaking aphasics: Development of a Cantonese aphasia battery," *Journal of Neurolinguistics*, vol. 7, no. 4, pp. 379–424, 1992.

[11] D. Povey, A. Ghoshal, G. Boulianne, L. Burget, O. Glembek, N. Goel, M. Hannemann, P. Motlicek, Y. Qian, P. Schwarz *et al.*, "The Kaldi speech recognition toolkit," in *IEEE 2011 workshop on automatic speech recognition and understanding*, no. EPFL-CONF-192584, 2011.

[12] Google Inc., "word2vec." [Online]. Available: https://code.google.com/archive/p/word2vec/

[13] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean, "Distributed representations of words and phrases and their compositionality," in *Advances in neural information processing systems*, 2013, pp. 3111–3119.

[14] L. v. d. Maaten and G. Hinton, "Visualizing data using t-SNE," *Journal of Machine Learning Research*, vol. 9, no. Nov, pp. 2579–2605, 2008.

[15] K. Fraser, F. Rudzicz, N. Graham, and E. Rochon, "Automatic speech recognition in the diagnosis of primary progressive aphasia," in *Proceedings of the Fourth Workshop on Speech and Language Processing for Assistive Technologies*, 2013, pp. 47–54.

[16] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 267–288, 1996.

[17] A. Loukina, K. Zechner, L. Chen, and M. Heilman, "Feature selection for automated speech scoring." in *BEA@ NAACL-HLT*, 2015, pp. 12–19.

[18] A. Kertesz, *Aphasia and associated disorders: Taxonomy, localization, and recovery*. Holt Rinehart & Winston, 1979.