

Accepted Manuscript

Automatic Quantitative Analysis of Spontaneous Aphasic Speech

Duc Le, Keli Licata, Emily Mower Provost

PII: S0167-6393(17)30312-6
DOI: [10.1016/j.specom.2018.04.001](https://doi.org/10.1016/j.specom.2018.04.001)
Reference: SPECOM 2552

To appear in: *Speech Communication*

Received date: 10 August 2017
Revised date: 11 March 2018
Accepted date: 4 April 2018



Please cite this article as: Duc Le, Keli Licata, Emily Mower Provost, Automatic Quantitative Analysis of Spontaneous Aphasic Speech, *Speech Communication* (2018), doi: [10.1016/j.specom.2018.04.001](https://doi.org/10.1016/j.specom.2018.04.001)

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Automatic Quantitative Analysis of Spontaneous Aphasic Speech

Duc Le[†], Keli Licata[‡], Emily Mower Provost[†]

University of Michigan, Ann Arbor, MI 48109, USA

[†]Computer Science and Engineering, [‡]University Center for Language and Literacy

{ducle,klicata,emilykmp}@umich.edu

Abstract—Spontaneous speech analysis plays an important role in the study and treatment of aphasia, but can be difficult to perform manually due to the time consuming nature of speech transcription and coding. Techniques in automatic speech recognition and assessment can potentially alleviate this problem by allowing clinicians to quickly process large amount of speech data. However, automatic analysis of spontaneous aphasic speech has been relatively under-explored in the engineering literature, partly due to the limited amount of available data and difficulties associated with aphasic speech processing. In this work, we perform one of the first large-scale quantitative analysis of spontaneous aphasic speech based on automatic speech recognition (ASR) output. We describe our acoustic modeling method that sets a new recognition benchmark on AphasiaBank, a large-scale aphasic speech corpus. We propose a set of clinically-relevant quantitative measures that are shown to be highly robust to automatic transcription errors. Finally, we demonstrate that these measures can be used to accurately predict the revised Western Aphasia Battery (WAB-R) Aphasia Quotient (AQ) without the need for manual transcripts. The results and techniques presented in our work will help advance the state-of-the-art in aphasic speech processing and make ASR-based technology for aphasia treatment more feasible in real-world clinical applications¹.

Index Terms: aphasia, clinical application, quantitative analysis, disordered speech recognition, aphasia quotient estimation

I. INTRODUCTION

Aphasia, a common speech-language disorder typically caused by focal brain damage, currently affects over two million people in the US and more than 180,000 acquire it every year [1]. Aphasia may cause impairments in both expressive and receptive language skills, including speaking, writing, reading, and listening [2]–[4]. Persons with aphasia (PWAs) often face significant communication difficulties, which may lead to feelings of frustration, loneliness, loss of autonomy, and social isolation, among others [5].

Spontaneous speech (e.g., answering an open-ended interview question, retelling a story, describing a picture) plays a prominent role in a PWA’s everyday interaction and is widely regarded in the aphasia literature as one of the most important modalities to analyze [6]–[9]. Example applications of spontaneous speech analysis include aphasia classification [10], treatment planning [7], recovery tracking [11], and diagnosis of residual aphasia post onset [9].

Analysis of spontaneous aphasic speech is typically carried out in clinical settings by Speech-Language Pathologists (SLPs) and often confined to a relatively small amount of speech samples with manually coded transcripts, which can be very time consuming to complete [7]. Furthermore, the analysis itself often requires a SLP’s expert knowledge of aphasia and linguistics. As a result, only the small percentage of PWAs who have frequent interaction with SLPs can access and benefit from spontaneous speech analysis, the results of which carry important implications for a PWA’s everyday interaction and future treatment plans. At the same time, SLPs in many settings have high productivity expectations and limited time outside of direct patient contact, thus restricting them from conducting such analysis regularly.

Techniques in automatic speech processing can potentially help SLPs perform this type of analysis more efficiently, thereby making its results and findings more commonly available to PWAs. However, previous works in the area of aphasic speech processing have two major limitations that prevent the development of fully automated systems capable of analyzing spontaneous aphasic speech. First, they often assume the availability of expertly produced speech transcripts, which are very time consuming to complete manually [12]–[15] and difficult to generate automatically [16], [17]. Second, they typically target speech with known prompts [18]–[22]. This removes the need for unconstrained automatic speech recognition (ASR) and greatly simplifies transcript generation, which can be achieved by variants of forced alignment [18]–[20] or keyword spotting [21], [22]. However, the reliance of this type of system on known prompts makes it inapplicable to spontaneous speech.

It is evident that ASR is a major bottleneck for spontaneous aphasic speech analysis. ASR performance must be sufficiently accurate such that the results and findings are not significantly affected by transcription mismatches. In addition, the features derived from ASR output must be relatively robust to recognition errors. However, the performance of ASR on aphasic speech and robustness of features against transcription errors have been under-explored in the literature. Our work helps bridge this gap by performing one of the first large-scale studies on ASR-based spontaneous aphasic speech analysis.

We present the paper in three sequential components. First, we describe our method for training acoustic models, which leads to significant improvement in aphasic speech recognition accuracy, achieving **37.37%** Word Error Rate (WER) on

¹Our code will be made publicly available upon publication of this paper.

Download English Version:

<https://daneshyari.com/en/article/6960517>

Download Persian Version:

<https://daneshyari.com/article/6960517>

[Daneshyari.com](https://daneshyari.com)