

51st Academy of Aphasia Proceedings

Analysis of auto-aligned and auto-segmented oral discourse by speakers with aphasia: A preliminary study on the acoustic parameter of duration

Lee T.^{a,*}, Kong A.^b, Chan V.^a, Wang H.^a

^a Department of Electronic Engineering, The Chinese University of Hong Kong, Hong Kong SAR

^b Department of Communication Sciences and Disorders, University of Central Florida, Orlando, FL, USA

Introduction

The use of forced alignment for automatic acoustic-phonetic segmentation of oral discourse produced by Cantonese speakers with aphasia was first reported by Lee et al. (2012). Specifically, using the HTK (Hidden Markov Model Toolkit), alignment and segmentation of initial consonants and rime of syllables in samples of read speech and spontaneous oral production was conducted automatically. This method of acoustic sample processing allowed subsequent measurement of various acoustic parameters of the speech samples.

This study aims to investigate the statistical variation of speech duration in Cantonese oral discourse by speakers with aphasia. It is focused on the duration of continuous speech chunks and the duration ratio of speech chunk and inter-chunk pauses. The observations on aphasic speech are compared with those of normal speech.

Methods

Language samples and their corresponding audio files were collected from 17 Cantonese-speaking individuals with fluent aphasia, according to the results of the Cantonese version of the Western Aphasia Battery (Yiu, 1992). All participants were interviewed following a standard protocol from the Cantonese AphasiaBank database (Kong et al., 2009), which included narrative tasks of personal monologue, picture and sequential description, and story-telling.

For each speaker, automatic alignment was done on all nine recordings for processing. Each audio recording was accompanied by an orthographic transcription, in which the spoken content was represented primarily in the form of Chinese characters. The alignment process involved a set of Cantonese acoustic models, which were developed for speaker-independent large vocabulary continuous speech recognition (Ching et al. 2006). The alignment results were represented in the “TextGrid” format, which could be retrieved, viewed and manipulated using the Praat software. The alignments were given at sub-syllable level, syllable level and word level.

Given that aphasic utterances tended to contain frequent occurrences of non-speech or unintelligible speech segments, these segments were skipped in the automatic alignment process. This was done by reducing the amplitudes of all signal samples in the segments to zero, and treating them as silence. In this way, the remaining part of the utterance could be time-aligned accurately and

* Corresponding author.

E-mail address: tanlee@ee.cuhk.edu.hk.

efficiently.

Results

Speakers with aphasia generally had poorer voice quality as compared with normal speakers. However, their speech remained highly intelligible in most cases. The Cantonese acoustic models trained with normal speech were found adequate for automatic alignment purpose.

The major observations on duration characteristics, based on a preliminary analysis carried out for the recordings from one normal speaker and one speaker with aphasia, include the following:

1. Normal speech contained 50% more words than aphasic speech;
2. The number of continuous speech chunks in aphasic speech was nearly three times of that in normal speech;
3. The average duration of speech chunk in aphasic speech was 1.73 second, versus 2.79 second in normal speech;
4. The average number of words in a speech chunk was 5.8 for aphasic speech, versus 10.5 for normal speech;
5. The average duration of inter-chunk pauses in aphasic speech was 0.82 second, versus 0.62 second in normal speech.

References

Kong, A.P.H., Law, S.P., & Lee, A.S.Y. (2009). *The construction of a corpus of Cantonese-aphasic-discourse: a preliminary report*. Poster presented at the American Speech-Language-Hearing-Association Convention, New Orleans, LA, USA.

Lee, A., Kong, A.P.H., & Law, S. (2012). Using forced alignment for automatic acoustic-phonetic segmentation of aphasic discourse. *Procedia Social and Behavioral Sciences*, 61, 92-93.

Yiu, E.M.L. (1992) Linguistic assessment of Chinese-speaking aphasics: Development of a Cantonese aphasic battery. *Journal of Neurolinguistics*, 7, 379- 424.

Ching, P.C., Lee, T., Lo, W.K., & Meng, H. (2006). Cantonese speech recognition and synthesis. In C.H. Lee, H. Li, L.S. Lee, R. Wang, & Q. Huo (Eds). *Advances in Chinese Spoken Language Processing* (pp. 365-386). Hackensack, NJ: World Scientific.