

Automatic sentence stress feedback for non-native English learners

Gary Geunbae Lee ^a, Ho-Young Lee ^{b,*}, Jieun Song ^c, Byeongchang Kim ^d,
Sechun Kang ^e, Jinsik Lee ^f, Hyosung Hwang ^b

^a Department of Computer Science and Engineering, Pohang University of Science and Technology, 77 Cheongam-ro, Nam-gu, Pohang, Gyeongbuk, South Korea

^b Department of Linguistics, Seoul National University, Daehak-dong, Gwanak-gu, Seoul, South Korea

^c Speech Hearing and Phonetic Sciences, University College London, 2 Wakefield St., London WC1N 1PF, UK

^d School of Computer and Information Communication Engineering, Catholic University of Daegu, Hayang-ro, Gyeongsan, Gyeongbuk, South Korea

^e Software R&D Center, Samsung Electronics, Maetan-dong, Yeongtong-gu, Suwon, Gyeonggi, South Korea

^f DMC R&D Center, Samsung Electronics, Maetan-dong, Yeongtong-gu, Suwon, Gyeonggi, South Korea

Received 19 March 2014; received in revised form 11 April 2016; accepted 13 April 2016

Available online 6 June 2016

Abstract

This paper proposes a sentence stress feedback system in which sentence stress prediction, detection, and feedback provision models are combined. This system provides non-native learners with feedback on sentence stress errors so that they can improve their English rhythm and fluency in a self-study setting. The sentence stress feedback system was devised to predict and detect the sentence stress of any practice sentence. The accuracy of the prediction and detection models was 96.6% and 84.1%, respectively. The stress feedback provision model offers positive or negative stress feedback for each spoken word by comparing the probability of the predicted stress pattern with that of the detected stress pattern. In an experiment that evaluated the educational effect of the proposed system incorporated in our CALL system, significant improvements in accentedness and rhythm were seen with the students who trained with our system but not with those in the control group.

© 2016 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Keywords: Sentence stress; Sentence stress feedback system; Stress prediction model; Stress detection model; Stress feedback provision model; CALL

1. Introduction

In English, every word has one or more lexical stresses¹ depending on the structure of the word and the number of syllables, but not all word stresses are phonetically realized in utterance. Content words, which deliver major semantic information and therefore require listeners' attention, normally receive stress on the utterance level whereas function words do not (cf. [Kingdon, 1958](#), for a detailed discussion about content and function words). Stress imposed on the utterance level has been traditionally called 'sentence stress' ([Gimson, 1980](#); [Jones, 1972](#)).

* Corresponding author at: Department of Linguistics, Seoul National University, Daehak-dong, Gwanak-gu, Seoul, South Korea. Tel: +82 10 2744 0584; fax: +82 2 882 2451.

E-mail address: hylee@snu.ac.kr (H.Y. Lee).

¹ Each word has one primary stress.

The major function of sentence stress is to highlight semantically important words and to form the rhythmic pattern of the utterance. It has been known that sentence stress occurs at regular intervals and unstressed syllables between consecutive stressed syllables are reduced, causing the impression of ‘stress-timed rhythm’. Stress-timed rhythm has been traditionally distinguished from ‘syllable-timed rhythm’ in which syllables are pronounced with similar duration without vowel reduction, as in French and Italian (Abercrombie, 1967; Lloyd James, 1940; Pike, 1945). However, the idea that the units of rhythm (i.e., feet in stress-timed rhythm and syllables in syllable-timed rhythm) are of equal duration has been negated in several empirical studies (e.g., Dauer, 1983; Roach, 1982; Wenk and Wiolland, 1982) and for many languages, the rhythmic classification is not as clear-cut as previously believed. Nonetheless, the rhythm classes are generally regarded as reflecting the different rhythmic characteristics.

Sentence stress is distinguished from pitch accent that carries pitch prominence caused by an important intonation event (Bolinger, 1958; Pierrehumbert, 1980) as well as rhythmic prominence caused by sentence stress. While all pitch accents are imposed on stressed syllables, some sentence stresses do not involve pitch prominence and only affect the rhythmic pattern of a sentence. Thus pitch accent can be considered to be ranked higher than sentence stress in the prosodic hierarchy.

It has been recognized in some previous research that prosody plays an equal or greater role than segments in the judgment of comprehensibility and/or accentedness of non-native speech (e.g. Anderson-Hsieh et al., 1992; Boula de Mareüil and Vieru-Dimulescu, 2006; Hahn, 2004; Tajima et al., 1997; Wennerstrom, 2000). As for Korean learners of English, great difficulties have been observed in rhythm and fluency. Low proficiency learners tend to place sentence stress on most of the words in a sentence, even on function words (Lee, 2011). They tend to use strong vowels even in unstressed syllables, giving the impression of syllable-timed rhythm to native English listeners.

Previous research shows that teaching only the pronunciation of segments does not significantly improve comprehensibility in non-native spontaneous speech (Elliott, 1997), whereas prosody teaching does (Derwing and Rossiter, 2003; Derwing et al., 1998). Hence this paper aims to propose an automatic sentence stress feedback system designed to provide non-native English learners, especially Koreans, with feedback on their sentence stress errors. It is hoped that this system will help non-native learners correct their errors and thereby improve their English rhythm and fluency, ultimately resulting in an increase in their oral proficiency.

Automatic prosody scoring systems have been proposed in previous literature to evaluate the English prosody (i.e., stress, rhythm and/or intonation) of both native and non-native speakers (Cheng, 2011; Hönig et al., 2010; Ito et al., 2009; Liscombe, 2007; Maier et al., 2009; Mostow and Duong, 2009; Suzuki et al., 2008; Tepperman et al., 2010; Yamashita et al., 2005). Several automated speech assessment systems with a prosody evaluation component have also been developed (Chandel et al., 2007; Chen and Zechner, 2011; Teixeira et al., 2000; Zechner et al., 2011). These automatic prosody or speech assessment systems are useful for stimulating and encouraging English learners, but their educational effect is somewhat limited because corrective feedback is not offered and learners are left without knowing what to correct. Hence, there have been a few studies concerning systems that offer automated feedback on prosody or speech (Bang et al., 2013; Sitaram et al., 2011).

Imoto et al. (2002) proposed a sentence stress detection model that provides diagnostic information to learners by comparing native speakers’ reference speech with Japanese learners’ speech. This study focused only on stress detection. Our proposed system, however, focuses on the integration of sentence stress prediction, detection, and error feedback technologies to help non-native learners effectively improve their English rhythm and fluency.

While several previous studies including Imoto et al. (2002) used native speech as a reference for direct comparison with non-native speech, this study uses automatically predicted sentence stresses generated by our sentence stress prediction model as a reference for sentence stress detection. This allows the proposed system to evaluate and process any given utterance.

Frequent and repetitive practice is necessary for non-native learners to improve their English rhythm and fluency. Since traditional face-to-face language learning opportunities are costly due to time and space barriers, CALL (Computer-Assisted Language Learning), which overcomes these barriers and offers non-native learners easy access to computer-based practice programs, has received much attention since Levy (1997) and Witt and Young (1997). Lee et al. (2011) showed that a CALL system based on dialog management technologies (cf. Lee et al., 2010) significantly improved learners’ spoken language skills. To evaluate our sentence stress feedback system, we set up a CALL system with which our sentence stress prediction, detection, and feedback provision models were incorporated.

The remainder of this paper is structured as follows: Section 2 describes the materials used for the proposed system. Section 3 proposes and describes the sentence stress prediction, detection, and feedback provision models. Section 4

Download English Version:

<https://daneshyari.com/en/article/6951512>

Download Persian Version:

<https://daneshyari.com/article/6951512>

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