



doi:10.1016/j.csl.2008.03.001

 [Cite or Link Using DOI](#)

Copyright © 2008 Elsevier Ltd All rights reserved.

Automatic pronunciation scoring of words and sentences independent from the non-native's first language

 Tobias Cincarek^{a,1},  
 Rainer Gruhn^{a,1}, Christian Hacker^b, Elmar Nöth^b and Satoshi Nakamura^a
^aATR Spoken Language Translation Research Labs,
 2-2-2 Hikaridai, Keihanna Science City, 619-0288 Japan







^bInstitute for Pattern Recognition, Friedrich-Alexander University Erlangen-Nuremberg, Germany

Received 6 October 2006; revised 8 January 2008; accepted 5 March 2008. Available online 12 March 2008.

Abstract

This paper describes an approach for automatic scoring of pronunciation quality for non-native speech. It is applicable regardless of the foreign language student's mother tongue. Sentences and words are considered as scoring units. Additionally, mispronunciation and phoneme confusion statistics for the target language phoneme set are derived from human annotations and word level scoring results using a Markov chain model of mispronunciation detection. The proposed methods can be employed for building a part of the scoring module of a system for computer assisted pronunciation training (CAPT). Methods from pattern and speech recognition are applied to develop appropriate feature sets for sentence and word level scoring. Besides features well-known from and approved in previous research, e.g. phoneme accuracy, posterior score, duration score and recognition accuracy,

▶ Article Toolbox




-  E-mail Article
-  Add to my Quick Links
-  Add to 
-  Permissions & Reprints
-  Cited By in Scopus (0)

Related Articles in ScienceDirect

- [Generating non-native pronunciation variants for lexico...
Speech Communication](#)
- [Other Challenges: Non-native Speech, Dialects, Accents,...
Multilingual Speech Processing](#)
- [On using units trained on foreign data for improved mul...
Speech Communication](#)
- [Articulatory-feature-based confidence measures
Computer Speech & Language](#)
- [Native-language sensitivities: evolution in the first y...
Trends in Cognitive Sciences](#)

 ▶ [View More Related Articles](#)
[View Record in Scopus](#)

The research collaboration tool

-  No user tags yet
-  This article has not yet been bookmarked
-  Not yet shared with any groups

 Be the first to add this article in 

new features such as high-level phoneme confidence measures are identified. The proposed method is evaluated with native English speech, non-native English speech from German, French, Japanese, Indonesian and Chinese adults and non-native speech from German school children. The speech data are annotated with tags for mispronounced words and sentence level ratings by native English teachers. Experimental results show, that the reliability of automatic sentence level scoring by the system is almost as high as the average human evaluator. Furthermore, a good performance for detecting mispronounced words is achieved. In a validation experiment, it could also be verified, that the system gives the highest pronunciation quality scores to 90% of native speakers' utterances. Automatic error diagnosis based on a automatically derived phoneme mispronunciation statistic showed reasonable results for five non-native speaker groups. The statistics can be exploited in order to provide the non-native feedback on mispronounced phonemes.

Keywords: Non-native speech; Pronunciation assessment; Sentence scoring; Word scoring; Mispronunciation detection; Phoneme mispronunciation statistic

Article Outline

1. [Introduction](#)
 2. [Data and labels](#)
 - 2.1. [ATR SLT non-native database](#)
 - 2.2. [PF-STAR non-native database](#)
 3. [Analysis of the human evaluation](#)
 - 3.1. [ATR data](#)
 - 3.2. [PF-STAR data](#)
 4. [Sentence scoring](#)
 - 4.1. [Sentence level pronunciation features](#)
 - 4.2. [Scoring method](#)
 - 4.3. [Score adjustment](#)
 5. [Word classification](#)
 - 5.1. [Word level pronunciation features](#)
 - 5.2. [Classification method](#)
 6. [Experimental setup](#)
 7. [Experimental results](#)
 - 7.1. [ATR data](#)
 - 7.1.1. [Sentence scoring results](#)
 - 7.1.2. [Word classification results](#)
 - 7.2. [PF-STAR data](#)
 - 7.2.1. [Sentence scoring](#)
 - 7.2.2. [Word classification](#)
 8. [Automatic error diagnosis](#)
 - 8.1. [Word mispronunciation model](#)
 - 8.2. [Phoneme mispronunciation statistics](#)
 - 8.2.1. [ATR data](#)
 - 8.2.2. [PF-STAR data](#)
 9. [Conclusion](#)
- [Acknowledgements](#)
- [References](#)



Corresponding author. Present address: Yahoo Japan Research,
Yahoo Japan Corporation, Midtown Tower, 9-7-1 Akasaka, Minato-ku,
107-6211 Japan.

¹ Presently, the author is with Harman/Becker and University of Ulm,
Germany.

Sponsored Links

[Full-Text Online Journals](#)

Full-text journals for academic
research at Questia Online Library.
www.Questia.com/Journals

[International MBA Online](#)

The University of Liverpool offers
International MBAs online. Get info
www.Online-Education.net

[Pre Publish Peer Review](#)

Paper format styles, pre submission
peer reviews - Get Ready to Publish
www.Peer-Review-Ready.com

Computer Speech & Language

Volume 23, Issue 1, January 2009, Pages 65-88

[Home](#) [Browse](#) [My Settings](#) [Alerts](#) [Help](#)



[About ScienceDirect](#) | [Contact Us](#) | [Information for Advertisers](#) | [Terms & Conditions](#) | [Privacy Policy](#)

Copyright © 2009 Elsevier B.V. All rights reserved. ScienceDirect® is a registered trademark of Elsevier B.V.