

## INTONATION EVALUATION OF ENGLISH UTTERANCES USING SYNTHESIZED SPEECH FOR COMPUTER-ASSISTED LANGUAGE LEARNING

AKINORI ITO, TOMOAKI KONNO, MASASHI ITO, SHOZO MAKINO  
Graduate School of Engineering  
Tohoku University  
6-6-05 Aramaki aza Aoba, Sendai, 980-8579, Japan  
{ aito; tomoaki; itojin; makino }@makino.ecei.tohoku.ac.jp

MOTOYUKI SUZUKI  
Institute of Technology and Science  
University of Tokushima  
2-1, Minamijosanjima-cho, Tokushima, 770-8506, Japan  
moto@m.ieice.org

Received January 2009; revised July 2009

**ABSTRACT.** *In this paper, we will describe a system for evaluating the intonation of English utterances made by Japanese native speakers using synthesized speech for the rapid development of a computer-assisted language learning (CALL) system. To evaluate the intonation of learners' utterances, reference utterances are needed, for which native English speakers' utterances should be used. However, it is costly to gather native speakers' utterances for all sentences in the system. Therefore, we examined an intonation evaluation method using synthesized speech generated by text-to-speech systems instead of real speech. The intonation evaluation system calculates scores between a learner's utterance and corresponding utterances by the teachers. We first compared the reliability of intonation evaluation using native and synthesized utterances, and found that the reliability of evaluation using synthesized utterances could be improved by using the weighted Mahalanobis distance for calculating the evaluation score. Next, we investigated a method of combining multiple scores of different teachers. In addition, we incorporated a feature for evaluating rhythm into intonation evaluation. As a result, the correlation between scores by human evaluators and the system was improved. Furthermore, we analyzed the tendency of intonation evaluations made by the system through limiting the evaluation utterances to find out for degradation of the system's performance.*

**Keywords:** Computer assisted language learning, CALL, Intonation evaluation, English, Synthesized speech

1. **Introduction.** In recent years, English proficiency has become increasingly important for people who are not native English speakers, and Computer-Assisted Language Learning (CALL) systems have become a popular way of learning foreign languages. In addition to developing skills for reading, writing and listening skills, some systems offer tools for learners to develop their speaking skills, such as through pronunciation and conversation exercises [1]. CALL systems can save time and money when learning a foreign language compared with taking classes. In fact, some systems have been actually used in universities as teaching materials in English classrooms [2].

Most systems that include the development of speaking skills evaluate learners' pronunciations from phoneme to phoneme or from word to word. Ehsani and Knodt [3] reviewed voice-interactive CALL systems that focused on pronunciation training.