

A Study on Estimation Method of  
L2 Learner ' s Second Language Ability  
by using Features in Conversation  
会話データからの特徴量を用いて  
第二言語学習者の言語能力推定手法に関する研究

Student No.: 201821642

Name: CHEN XINNAN

The research theme of this paper is to support second language (L2) learners to improve their second language ability by utilizing chat system. The main problem of existing chat systems is that it is not possible to chat with learners to adapt their second language level. In order to add a function that adapts the learner's second language level to an existing chat system, we need to measure second language level by examining conversation data with the learner. Therefore, to extract learner's second language capability, we propose a method to predict the language examination score of learners from chat context. This research first investigates whether the number of utterances, number of sentences, word tokens and word types per utterance of chat context are correlated with second language examination score. Second, we build a predicting model to see the relationship between the chat context and second language examination score. As feature values of regression model for predicting the language examination score, we use features of chat time, sentence time, word token and word type. Also, the unnatural sentence structure is considered as a feature. For evaluation, we use the root mean square error to check the results of prediction model. We apply the proposed model to Japanese and English chat and compare the results. We also collect public opinions using crowdsourcing platforms about the same conversation dataset to compare our prediction model's performance with subjective analysis.

Linguistic features show no clear correlation with language examination score whereas error feature shows very weak correlation with score. Also results show human cannot accurately judge a person's language ability by just looking at the simple conversation sentences. Both objective and subjective analysis results reveal that language ability cannot be predicted by simple conversation data. So future studies should focus on following some guidelines in data collection process and also some other feature vectors.

Academic Advisors: Principal: Tetsuji SATOH

Secondary: Kei WAKABAYASHI