

## 雑談内容を用いた留学生の 日本語能力推定手法に関する研究

CHEN XINNAN

近年、Siri に代表する音声アシスタントサービスや SNS 上のユーザ発話に対して自動で応答をするサービス (Bot) など、スマートフォンや PC を用いて人間とコンピュータによる対話 (雑談) を行う機会が増加しており、システムとユーザが対話を行う雑談システムに関する研究が多く行われている。留学生にとっても、日本語の練習に利用するなどの活用が考えられるが、これまでのシステムは相手の日本語を考慮しないため、留学生が利用することは難しい。本研究では、留学生の実際の会話を分析することで、雑談の内容から日本語能力を測り、相手の日本語能力に合わせた応答を行う雑談システムを構築することを目指す。

そのために、本研究では日本語能力検定試験の文法表と語彙表をシステムに導入することで、留学生の日本語能力を測ることができるかどうかを調べることを目的とする。WeChat や LINE を用いてオンラインで実験実施者と対話してもらう。対話は日本語で 30 分くらい行う。また、対象者の日本語能力検定試験の級、検定試験のスコアを教えてもらい、そのスコアを対話内容から推定する機械学習モデルを構築する。この機械学習モデルの特徴として、二つがある。一つは会話文の単語と文法を日本語検定試験の文法表と語彙表で参照しながら、5次元と10次元のベクトルを作成した。もう一つはスコアである。今回被験者の中に N1 能力者は 14 人、N2 能力者は 3 人、日本人は 3 人 (スコアを N1 180 点で設定した) がいた。参加者の会話文を用いて語彙ベクトルと文法ベクトルを作成した。これに基づいて 5次元と10次元のベクトルを作成する。

分析方法としては回帰分析の線形回帰と K-近傍回帰で平方平均 2 乗誤差 (RMSE: Root Mean Squared Error) を求めて、一番小さい誤差が機械学習しない場合の誤差より小さい場合、留学生の会話文からスコアを予測するのは有効である。実験の結果、留学生の検定試験のスコアから機械学習しない予測をした場合、N1 の被験者を対象として平方平均 2 乗誤差は 18.8 であった。留学生の会話文を分析して構築された機械学習モデルの方の場合、一番低いのは線形回帰で 5次元ベクトルの平方平均 2乗誤差は 14.0 であった。これにより留学生の会話文から日本語能力を測る方が、スコアだけによる予測より誤差が小さい結果を得られた。

しかしながら、留学生の日本語能力を測る際に、単語と文法の抽出に誤りがあることから、抽出手法を改善することが課題である。また、会話文中の不自然な表現を特徴量として抽出することも課題である。そして、実験に用いたデータ数が少ないため、結果の信頼性を高めることも課題となる。

(指導教員 若林 啓)