

先端科学技術研究科 修士論文要旨

所属研究室 (主指導教員)	自然言語処理学 (渡辺 太郎 (教授))		
学籍番号	2111402	提出日	令和 5年 7月 24日
学生氏名	井手 佑翼		
論文題目	Predicting Japanese Lexical Complexity for Non-Native Readers		
要旨			
<p>Lexical complexity prediction (LCP) is the task of predicting how difficult words in a text are on a continuous scale. It plays a vital role in simplifying or annotating complex words to assist readers. To study lexical complexity in Japanese, we construct the first Japanese LCP dataset. Our dataset provides separate complexity scores for Chinese/Korean annotators and others to address the readers' L1-specific needs. In the analysis, we show that the L1 background affects complexity and deserves special consideration. In the baseline experiment, we demonstrate the effectiveness of our BERT-based system for Japanese LCP.</p>			