

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)	知能コミュニケーション (中村 哲 (教授))		
学籍番号	2111230	提出日	令和 4年 7月 22日
学生氏名	古川 慧		
論文題目	Applying Syntax-Prosody Mapping Hypothesis, Prosodic Well-Formedness Constraints, and Boundary-Driven Theory to Neural Sequence-to-Sequence Speech Synthesis		
要旨			
<p>End-to-end text-to-speech synthesis (TTS), which generates speech sounds directly from strings of texts or phonemes, has improved the quality of speech synthesis over the conventional TTS. However, most previous studies have been evaluated based on subjective naturalness and have not objectively examined whether they can reproduce pitch patterns of phonological phenomena such as downstep, rhythmic boost, and initial lowering that reflect syntactic structures in Japanese. These phenomena can be linguistically explained by phonological constraints and the syntax-prosody mapping hypothesis (SPMH), which assumes projections from syntactic structures to phonological hierarchy. However, the mechanism of downstep is still open to question, and modifications to the current phonological theory are needed to provide a unified explanation for the several phenomena including unaccented downstep. We propose a new phonological theory called boundary-driven theory, which can explain downstep, unaccented downstep, rhythmic boost, initial lowering, and other phonological phenomena in a unified manner. To synthesize linguistic phenomena involving syntactic or phonological constraints, we propose a model using phonological symbols based on the SPMH, prosodic well-formedness constraints, and boundary-driven theory. Experimental results showed that the proposed method synthesized similar pitch patterns to those reported in linguistics experiments for the phenomena of initial lowering and rhythmic boost. The proposed model efficiently synthesizes phonological phenomena in the test data that were not explicitly included in the training data.</p>			