

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	2111408	提出日	令和 6年 1月 18日
学生氏名	的川 雄飛		
論文題目	多言語固有表現に由来する国際音声記号 (IPA) 言語モデルの構築		
要旨			
<p>固有表現 (NE) のリンキングやアライメントにおいて、音声的類似性を用いた手法が効果を発揮することがあるが、そこで用いられる音声的特徴には言語のカバレッジや音声情報の反映の程度という点で課題がある。このようなNE関連タスクへの応用を踏まえると、あらゆる言語の音声を記述する体系である国際音声記号 (IPA) を活用し、多言語NEの表記から変換されたIPA系列を入力とした言語モデルが有用である可能性があるが、こうしたモデルは従来の研究で構築されてこなかった。本研究では、多言語NEとそれに対応するIPA系列を格納したデータセットをルールベースのIPA変換により作成したのち、そのうちのIPA系列を入力したモデルを構築することで、このような言語モデリングを実現した。本研究で構築した言語モデルは、従来よりも多数の言語に対応した、テキストの表層のみにとどまらない音声的特徴の利用という形でNEタスクに応用することが期待できる。</p>			