

様式 F - 7 - 1

科学研究費助成事業（学術研究助成基金助成金）実施状況報告書（研究実施状況報告書）（平成29年度）

所属研究機関名称		奈良先端科学技術大学院大学	機関番号	14603
研究 代表者	部局	情報科学研究科		
	職	特任准教授		
	氏名	サクリアニ サクティ		

1. 研究種目名 基盤研究(C)(一般) 2. 課題番号 17K00237

3. 研究課題名 ゼロ資源での教師なし音響パターン発見のための研究

4. 補助事業期間 平成29年度～平成31年度

## 5. 研究実績の概要

2020年東京オリンピック・パラリンピックが近づくにつれ、海外からの観光客との言葉の壁はますます深刻な問題となっている。現在の音声認識・音声翻訳技術は、リソースが大きい言語についてはすでに容易に利用できるため、ここでは言語特有の知識も書き起こしデータもないようなゼロ資源の音声処理の問題を対象とする。教師なし音響ユニットモデリングやパターン発見技術は存在するが、実際に言語的および意味的表現との関連まではまだ開発されていないため、本研究では未知言語の音声と意味表現を結びつける手段としてEEG解析に基づく認知知識をゼロ資源モデリングに組み込む手法について提案し、フレームワークを完成させ複数言語での応用を実証する。2017年度は、自然言語処理および認知科学に関する文献調査、ならびに言語および音声の認知についてアフリカ言語（ツォンガ語など）のゼロ資源モデリングおよびEEG解析の設計およびシステム構築、Dirichletプロセスのガウス混合モデルを中心に、音声特徴ベクトルをクラスター化してクラスの動的なセットを行った。各クラスを音響単位とみなすことにより、音声は、クラス後立線系列として表すことができる。この最適化により、サブワードモデリングの品質が大幅に向上することが示された。この研究手法は、ゼロ資源のスピーチチャレンジに参加して最高のパフォーマンスを達成し、コンペティションに優勝したことで示される。また、脳波検査を用いて日本語の文章を判別する実験を行った。この実験では、テンプレートマッチングと分類器を使用して、さまざまな設定でのパフォーマンスを調査した。さらに話者依存についても実験を行った。さらに、シータ、アルファ、ベータ、低ガンマ、およびすべての周波数帯の組み合わせを含む複数の周波数帯で実験を行った結果、複数の組み合わせの周波数帯が最も良い結果を示した。

## 6. キーワード

音声認識 ゼロ資源音声技術 脳波

## 7. 現在までの進捗状況

区分 (2) おおむね順調に進展している。

理由  
当初計画にあったアフリカ言語（ツォンガ語）のゼロリソースモデリングの構築に成功した。さらに、英語、ドイツ語、フランス語、中国語、アフリカ言語を含むゼロリソースの音声チャレンジに勝つことができた。EEG分析に関しては、日本語の分析は予定通り進んだ。ただし、アフリカ系言語のような低資源言語は国内で被験者を見つけるのが難しく課題として残る。

2 版

## 8．今後の研究の推進方策

2018年と2019年に以下の研究活動を継続する。  
 2018年度：ゼロ資源モデリングの構築とEEG 実験の継続、認知知識ソースの解析、およびゼロ資源モデリングへの知識統合のための設計検討  
 2019年度：提案フレームワークの完成、性能検討、複数言語での応用について実証実験。最終的にはソング語から日本語/英語への音声翻訳ができるシステムを開発する。

## 9．次年度使用が生じた理由と使用計画

次年度使用額が無いため、記入しない。

## 10．研究発表（平成29年度の研究成果）

〔雑誌論文〕 計10件（うち査読付論文 10件/うち国際共著 10件/うちオープンアクセス 4件）

1．著者名 Hiroki Watanabe, Hiroki Tanaka, Sakriani Sakti, Satoshi Nakamura	4．巻 Vol.1
2．論文標題 Subject-independent Classification of Japanese Spoken Sentences by Multiple Frequency Bands Phase Pattern of EEG Response during Speech Perception	5．発行年 2017年
3．雑誌名 Proceeding of INTERSPEECH 2017	6．最初と最後の頁 pp. 2431-2435
掲載論文のDOI（デジタルオブジェクト識別子） 10.21437/Interspeech.2017-854	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 該当する

1．著者名 Andros Tjandra, Sakriani Sakti, Satoshi Nakamura	4．巻 Vol.1
2．論文標題 Speech Recognition Features Based On Deep Latent Gaussian Models	5．発行年 2017年
3．雑誌名 Proceeding of IEEE International Workshop on Machine Learning for Signal Processing (MLSP 2017)	6．最初と最後の頁 -
掲載論文のDOI（デジタルオブジェクト識別子） 10.1109/MLSP.2017.8168174	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 該当する

1. 著者名 Andros Tjandra, Sakriani Sakti, Satoshi Nakamura	4. 巻 Vol. 1
2. 論文標題 Local Monotonic Attention Mechanism for End-to-End Speech and Language Processing	5. 発行年 2017年
3. 雑誌名 Proceedings of the The 8th International Joint Conference on Natural Language Processing	6. 最初と最後の頁 pp. 431-440
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 該当する

1. 著者名 Andros Tjandra, Sakriani Sakti, Satoshi Nakamura	4. 巻 なし
2. 論文標題 End-to-End Speech Recognition with Local Monotonic Attention	5. 発行年 2017年
3. 雑誌名 Proceedings of NIPS Workshop on Machine Learning for Audio Signal Processing (ML4Audio)	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 該当する

1. 著者名 Andros Tjandra, Sakriani Sakti, Satoshi Nakamura	4. 巻 Vol. 1
2. 論文標題 Listening while Speaking: Speech Chain by Deep Learning	5. 発行年 2017年
3. 雑誌名 Proceedings of IEEE Automatic Speech Recognition and Understanding (ASRU)	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/ASRU.2017.8268950	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 該当する

1. 著者名 Andros Tjandra, Sakriani Sakti, Satoshi Nakamura	4. 巻 Vol. 1
2. 論文標題 Attention-based Wav2Text with Feature Transfer Learning	5. 発行年 2017年
3. 雑誌名 Proceedings of IEEE Automatic Speech Recognition and Understanding (ASRU)	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/ASRU.2017.8268951	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 該当する

2 版

1. 著者名 Michael Heck, Sakriani Sakti, Satoshi Nakamura	4. 巻 Vol. 1
2. 論文標題 Feature Optimized DPGMM Clustering for Unsupervised Subword Modeling: A Contribution to ZeroSpeech 2017	5. 発行年 2017年
3. 雑誌名 Proceedings of IEEE Automatic Speech Recognition and Understanding (ASRU)	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/ASRU.2017.8269011	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 該当する

1. 著者名 Michael Heck, Sakriani Sakti, Satoshi Nakamura	4. 巻 Vol. E101-D
2. 論文標題 Learning Supervised Feature Transformations on Zero Resources for Improved Acoustic Unit Discovery	5. 発行年 2017年
3. 雑誌名 Transaction on Information and Systems	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) 10.1587/transinf.2017EDP7175	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 該当する

1. 著者名 Hayato Maki, Hiroki Tanaka, Sakriani Sakti, Satoshi Nakamura	4. 巻 Vol. 1
2. 論文標題 Graph Regularized Tensor Factorization for Single-trial EEG Analysis	5. 発行年 2018年
3. 雑誌名 Proceeding of International Conference on Acoustic, Speech, and Signal Processing (ICASSP)	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 該当する

1. 著者名 Hayato Maki, Hiroki Tanaka, Sakriani Sakti, Satoshi Nakamura	4. 巻 Vol. 1
2. 論文標題 Quality Prediction of Synthesized Speech Based on Tensor Structured EEG Signals	5. 発行年 2018年
3. 雑誌名 Transaction of PLOS One	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 該当する

〔学会発表〕 計8件（うち招待講演 0件 / うち国際学会 8件）

1. 発表者名 Hiroki Watanabe
2. 発表標題 Subject-independent Classification of Japanese Spoken Sentences by Multiple Frequency Bands Phase Pattern of EEG Response during Speech Perception
3. 学会等名 INTERSPEECH (国際学会)
4. 発表年 2017年

1. 発表者名 Andros Tjandra, Sakriani Sakti
2. 発表標題 Speech Recognition Features Based On Deep Latent Gaussian Models
3. 学会等名 IEEE International Workshop on Machine Learning for Signal Processing (MLSP 2017) (国際学会)
4. 発表年 2017年

1. 発表者名 Andros Tjandra
2. 発表標題 Local Monotonic Attention Mechanism for End-to-End Speech and Language Processing
3. 学会等名 the International Joint Conference on Natural Language Processing (IJCNLP 2017) (国際学会)
4. 発表年 2017年

1. 発表者名 Andros Tjandra
2. 発表標題 End-to-End Speech Recognition with Local Monotonic Attention
3. 学会等名 NIPS Workshop on Machine Learning for Audio Signal Processing (ML4Audio) (国際学会)
4. 発表年 2017年

2 版

1. 発表者名 Andros Tjandra, Sakriani Sakti, Satoshi Nakamura
2. 発表標題 Listening while Speaking: Speech Chain by Deep Learning
3. 学会等名 IEEE Automatic Speech Recognition and Understanding (ASRU) (国際学会)
4. 発表年 2017年

1. 発表者名 Andros Tjandra, Sakriani Sakti, Satoshi Nakamura
2. 発表標題 Attention-based Wav2Text with Feature Transfer Learning
3. 学会等名 IEEE Automatic Speech Recognition and Understanding (ASRU) (国際学会)
4. 発表年 2017年

1. 発表者名 Michael Heck, Sakriani Sakti
2. 発表標題 Feature Optimized DPGMM Clustering for Unsupervised Subword Modeling: A Contribution to ZeroSpeech 2017
3. 学会等名 IEEE Automatic Speech Recognition and Understanding (ASRU) (国際学会)
4. 発表年 2017年

1. 発表者名 Hayato Maki
2. 発表標題 Graph Regularized Tensor Factorization for Single-trial EEG Analysis
3. 学会等名 International Conference on Acoustic, Speech, and Signal Processing (ICASSP) (国際学会)
4. 発表年 2018年

〔図書〕 計0件

1 1 . 研究成果による産業財産権の出願・取得状況

〔出願〕 計0件

〔取得〕 計1件

産業財産権の名称 国立大学法人 奈良先端科学技術大学院大学	発明者 アンドロスチャンドラ, サクリアニサクティ, 中村哲	権利者 同左
産業財産権の種類、番号 特許、特願2018-1538	取得年 2018年	国内・外国の別 国内

1 2 . 科研費を使用して開催した国際研究集会

計0件

1 3 . 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関			
	University of Indonesia	Institute Technology Bandung	-	-
Indonesia	University of Indonesia	Institute Technology Bandung	-	-
-	-	-	-	-
-	-	-	-	-
-	-	-	-	-
-	-	-	-	-
-	-	-	-	-

1 4 . 備考

-