

## 科学研究費助成事業 研究成果報告書

令和 2 年 6 月 21 日現在

機関番号：14603

研究種目：基盤研究(C) (一般)

研究期間：2017～2019

課題番号：17K00237

研究課題名(和文) ゼロ資源での教師なし音響パターン発見のための研究

研究課題名(英文) Research for unsupervised acoustic pattern discovery with zero resources

研究代表者

サクリアニ サクティ (Sakti, Sakriani)

奈良先端科学技術大学院大学・先端科学技術研究科・特任准教授

研究者番号：30625083

交付決定額(研究期間全体)：(直接経費) 3,500,000円

研究成果の概要(和文)：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した。未知言語を理解するため、人間の脳が言語を処理する過程の分析・調査を行った。さらに、深層学習に基づく閉ループスピーチチェーンモデルを開発して、機械が話している間、聞く方法を学習できるようにした。これは、人間の音声認識と生産行動を統合した最初のディープラーニングモデルである。

研究成果の学術的意義や社会的意義

アフリカ言語(ツォンガ語)とインドネシア言語のゼロリソースモデリングの構築に成功した。また、2017年と2019年の世界ゼロ資源スピーチチャレンジに参加し、提案手法で上位結果を得ることができた。さらに、深層学習に基づく閉ループスピーチチェーンモデルを開発して、機械が話している間、聞く方法を学習できるようにした。2019年では世界言語言語コンソーシアムのためにユネスコとも協力した。この研究の結果は、トップ会議(ASRU、Interspeech、ICASSP)とトップジャーナル(IEEE / ACM TASLP)で公開された。さらに、スピーチチェーンモデルの特許も取得した。

研究成果の概要(英文)：With the Tokyo Olympics and Paralympics approaching, language barriers between tourists are becoming critical problems to overcome. Current speech recognition and speech translation have been readily available, but only for several languages where large resources are available. Here, we addressed zero-resource speech problem where language specific knowledge and collection of transcribed data are not available. In order to understand the unknown language, we analyzed and investigated the process by which the human brain processes language. In addition, we have developed a closed-loop speech chain model based on deep learning so that we can learn how to listen while the machine is speaking. This is the first deep learning model that integrates human speech recognition and production behavior.

研究分野：情報学

キーワード：音声認識 ゼロ資源音声技術 脳波 音声翻訳

## 1. 研究開始当初の背景

21世紀に入り、異なる国同士の人が情報を交換する機会は飛躍的に増えている。特に、日本では、海外からの渡航者に対し、観光、移住、留学など、その対象範囲は広がる一方である。こうした変化により、異なる言語を話す人々のインタラクションを助けるような創意工夫が必要とされている。また、東京オリンピック・パラリンピックの開催が近づくにあたり、海外からの観光客との言葉の壁も克服すべき深刻な問題となっている。話し言葉の翻訳は、自分の言葉で話し、お互いに理解ができるようになる画期的な技術のひとつであり、音声翻訳技術は言葉の壁を越え、文化間ギャップを埋めるために極めて重要である。

音声翻訳システムの開発にあたって重要な要素となるのが、発話音声を文字テキストに変換する自動音声認識 (ASR) である。最近では、深層学習アルゴリズムがさまざまな分野で最先端の成果を多く生み出しており、ニューラルネットワークを ASR に利用する動きも復活している。事実、言語によっては、ASR 技術を気軽に使えるところまで開発が進んでいる。しかし、世界の人口 35 億人が話す 6000 を超える言語をカバーするにはまだ程遠い。

現代の最先端の ASR システムでは大容量の音声データ、音声書き起こし・アノテーション、語彙単位や音韻セットといった言語の事前知識など大規模なデータ資源が必要とされ、このような資源は、さらに音声音響特性を意味のある単語にマッピングするモデルの学習に必要である。そのため、音声認識システムの開発はこれまで資源が多く存在する言語に偏る傾向があった。こうした「言語格差」の問題を解消する 1 つの方法として、低資源言語への対応が考えられる。

最近では、言語の専門性にあまり頼らなくてよい、教師なしのデータ駆動型システムの開発に力を入れているプロジェクトも多い。ゼロ資源のモデリングは、書き起こしがなく音響データだけを頼りに言語の音声構造要素を学習する完全な教師なし手法である。つまり、言語によっては、書き起こしやアノテーション、その他の事前知識がなく、音声データだけが存在する場合がある。ゼロ資源モデリングには、教師なし音響ユニットモデリングと話し言葉の検知が含まれる。未知言語から音響特性を分類する手法としては、Dirichlet process Gaussian mixture model (DPGMM) などのベイズモデルがある。

しかしながら、未知言語の場合、音響ユニット/パターンと実際の言語的、意味的ユニットの関係を見出すにはまた問題が残っている。これは、母国語の音声を認知し発話するシステムを気づく人間の子供の発育過程に似ている。言語表現に基づく人間の脳活動の定量分析は、脳神経科学の分野で積極的に行われている。これまで、EEG (electroencephalography) と呼ばれる脳波計が、言語理解における認知プロセスの研究に欠かせないシステムとして利用されていた。事実、発話音声は、韻律 (デルタバンド)、シラブル (シータバンド)、音素 (ガンマバンド) など、神経律動要素の構造と極めて似ているという研究発表もある。

## 2. 研究の目的

音響ユニット/パターンと実際の言語的、意味的ユニットの関係は、母国語の音声を認知し発話するシステムに気づく人間の子供の発育過程に似ている。本研究では、音声と認知の両方の観点で研究を進める。ただし、言語特有知識と書き起こしテキストデータが大量に存在するような言語理解の音声認識ではなく、そうしたデータが存在しない未知言語のゼロ資源モデリングに焦点をあてる。未知言語を理解するため、EEG (electroencephalography) の研究に基づき人間の脳が言語を処理する過程の分析・調査を行う。さらに、そうした研究結果に基づき、未知言語の音声と意味表現を結びつける手がかりとして、認知に関する知識をゼロ資源の音声モデリングに適用して、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。

## 3. 研究の方法

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した。

### (1) ゼロ資源音声モデリング

世界には多くの言語がそれぞれ異なる特性を持ち存在する。本研究では、母国語、既知言語、未知言語の言語処理に限定して研究を行う。特に、母国語/既知言語としては英語/日本語を、未知言語としてアフリカ (ツォンガ語など) およびアジア (インドネシア語など) の言語の 1 つを対象とする。

(2) 言語理解および音声理解のための EEG 解析

EEG の結果に基づき、言語的および意味的要素などゼロ資源モデリングに必要な認知知識ソースをすべて抽出/数式化する。

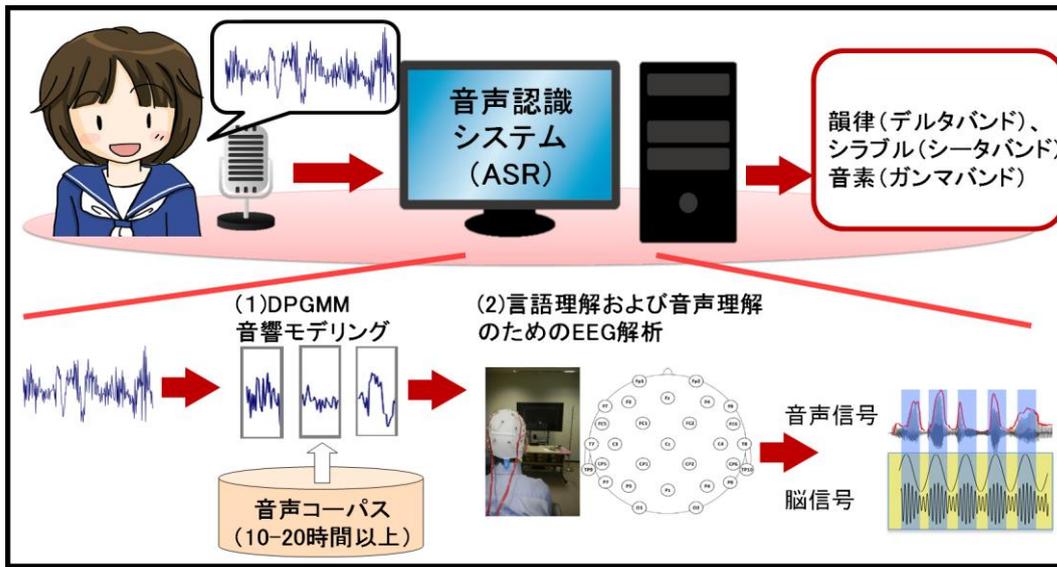


図 1 : EEG 手法および認知アプローチ

(3) 音声モデリングへの認知知識の統合を実施する。

認知知識ソースのゼロ資源モデリングへの統合多様な知識ソースでモデリングを行う困難に対応するため、深層学習手法を用いる。最近、end-to-end タイプのニューラルネットワークベースの音声認識システムが優れた成果を出している。可能性のあるアーキテクチャーとして、アテンションベースのDNNが考えられる。これは、エンコーダ・デコーダのアーキテクチャーを使用しており、エンコーダおよびデコーダとして機能する2つのディープニューラルネットワークモデルを組み合わせたものである。エンコーダ・デコーダモデルは、まず、音声信号を中間表現にエンコードし、それをさらに意味表現にデコードする。また、出力の各要素に対し、様々な入力に注意を払うことを自動学習するアテンションメカニズムも用いた。さらに、脳活動の認知知識に加え、脳活動を中間表現にエンコードするエンコーダも構築した。この中間表現は、その後、アテンションメカニズムを介して音声信号と結合される。このようにして、音声をベースにして、知識モデルに使えるトップダウン情報の統合までを行う手法の開発を目指した。

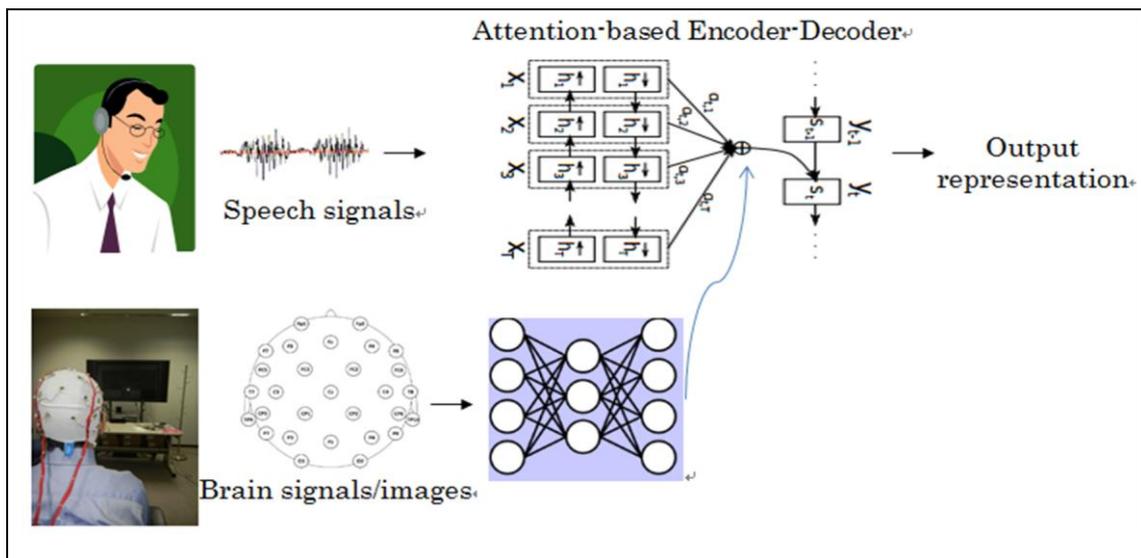


図2 : 本研究が提案するフレームワークのフロー

#### 4. 研究成果

2017 年度では、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、2017 年のゼロリソースの音声チャレンジに勝つことができた。インドネシアの大学との連携もまだ進行中であるが、期待された研究成果は得られなかった。さらに、脳解析研究について、脳波検査を用いて文章を判別する実験を行った。一方で、深層学習に基づく閉ループスピーチチェーンモデルを開発して、機械が話している間、聞く方法を学習できるようにした。これは、人間の音声認識と生産行動を統合した最初のディープラーニングモデルであり、子どもが話しながら聞いて新しい言語を学ぶことができるという特性の1つである。子供が言語を学ぶようにコンピューターが言語を学ぶことが可能になることを目指して開発した。

2018 年度では、インドネシア言語のゼロリソースモデリングの構築に成功した。今回は、Dirichlet プロセスのガウス混合モデルを利用する代わりに、ディープラーニングに基づいてシステムを構築した。このシステムでは、(1) サブワード単位を発見すること、(2) 音声を合成することの両方とも教師なしで行うことができた。また、2019 年 3 月の世界ゼロ資源スピーチチャレンジに参加し、提案手法で上位結果を得ることができた。さらに、脳解析研究について、Speech-Imagination 中の EEG 振動と相手の音声包絡線との間の同期を明らかにするための研究を行った。具体的には、(1) Speech-Imagination 中の EEG ベースの回帰音声エンベロープが顕在音声エンベロープと相関するかどうか、および (2) Imagined-EEG が参加した異なるエンベロープで音声刺激を分類できるかどうかを調べた。これらの結果は、Speech-Imagination 中の EEG と明白な対応物のエンベロープとの間の同期を示している。

2019 年度では引き続き 2020 年 3 月の World Zero Resource Speech Challenge に参加し、システムのパフォーマンスを向上させることができた。また、テキストを書き起こさずに、未知の言語用の教師なし音声変換を作成し、IEEE ASRU で公開した。また、世界言語言語コンソーシアムのためにユネスコとも協力して、すべての言語、すべての人々、すべての国の言語テクノロジーをサポートすることを目指している。このプロジェクトは、「国連国際先住民言語の 10 年」プロジェクトの下で継続される予定である（2022 年から 2023 年）。

## 5. 主な発表論文等

〔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 7件）

1. 著者名 Hayato Maki, Sakriani Sakti, Hiroki Tanaka, Satoshi Nakamura	4. 巻 13
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 PloS One	6. 最初と最後の頁 pp. 1-13
掲載論文のDOI (デジタルオブジェクト識別子) 10.1371/journal.pone.0193521	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 Michael Heck, Sakriani Sakti, Satoshi Nakamura	4. 巻 26
2. 論文標題 Dirichlet Process Mixture of Mixtures Model for Unsupervised Subword Modeling	5. 発行年 2018年
3. 雑誌名 IEEE/ACM Transactions on Audio, Speech, and Language Processing	6. 最初と最後の頁 pp. 2027 - 2042
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/TASLP.2018.2852500	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 Hiroki Watanabe, Hiroki Tanaka, Sakriani Sakti, Satoshi Nakamura	4. 巻 Volume E102.D, issue 2
2. 論文標題 Neural Oscillation-Based Classification of Japanese Spoken Sentences During Speech Perception	5. 発行年 2019年
3. 雑誌名 IEICE Transactions on Information and Systems	6. 最初と最後の頁 pp. 383-391
掲載論文のDOI (デジタルオブジェクト識別子) 10.1587/transinf.2018EDP7293	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 Hiroki Tanaka, Hiroki Watanabe, Hayato Maki, Sakti Sakriani, Satoshi Nakamura	4. 巻 13
2. 論文標題 Electroencephalogram-Based Single Trial Detection of Language Expectation Violations in Listening to Speech	5. 発行年 2019年
3. 雑誌名 Frontiers in Computational Neuroscience	6. 最初と最後の頁 なし
掲載論文のDOI (デジタルオブジェクト識別子) 10.3389/fncom.2019.00015	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 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 IEICE 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 Andros Tjandra, Sakriani Sakti, Satoshi Nakamura	4. 巻 Vol. 7
2. 論文標題 End-to-End Speech Recognition Sequence Training with Reinforcement Learning	5. 発行年 2019年
3. 雑誌名 IEEE Access	6. 最初と最後の頁 pp. 79758-79769
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/ACCESS.2019.2922617	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 Johanes Effendi, Sakriani Sakti, Katsuhito Sudoh, Satoshi Nakamura	4. 巻 Vol. E103-D No.03
2. 論文標題 Leveraging Neural Caption Translation with Visually Grounded Paraphrase Augmentation	5. 発行年 2020年
3. 雑誌名 IEICE Transactions on Information and System	6. 最初と最後の頁 pp. 674-683
掲載論文のDOI (デジタルオブジェクト識別子) 10.1587/transinf.2019EDP7065	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 Volume E103.D No, 2
2. 論文標題 Recurrent Neural Network Compression based on Low-Rank Tensor Representation	5. 発行年 2020年
3. 雑誌名 IEICE Transactions on Information and System	6. 最初と最後の頁 pp. 435-449
掲載論文のDOI (デジタルオブジェクト識別子) 10.1587/transinf.2019EDP7040	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. 28
2. 論文標題 Machine Speech Chain	5. 発行年 2020年
3. 雑誌名 Journal IEEE Transactions on Audio, Speech, and Language Processing (TASLP)	6. 最初と最後の頁 pp. 976 - 989
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/TASLP.2020.2977776	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 Hiroki Watanabe, Hiroki Tanaka, Sakriani Sakti, Satoshi Nakamura	4. 巻 Vol. 153
2. 論文標題 Synchronization between overt speech envelope and EEG oscillations during imagined speech	5. 発行年 2020年
3. 雑誌名 Neuroscience Research	6. 最初と最後の頁 pp. 48-55
掲載論文のDOI (デジタルオブジェクト識別子) 10.1016/j.neures.2019.04.004	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

〔学会発表〕 計44件 (うち招待講演 0件 / うち国際学会 34件)

1. 発表者名 Andros Tjandra, Sakriani Sakti, Satoshi Nakamura
2. 発表標題 Sequence-to-Sequence ASR Optimization via Reinforcement Learning
3. 学会等名 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (国際学会)
4. 発表年 2018年

1. 発表者名 Hayato Maki, Hiroki Tanaka, Sakriani Sakti, Satoshi Nakamura
2. 発表標題 Graph regularized tensor factorization for single-trial EEG analysis
3. 学会等名 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (国際学会)
4. 発表年 2018年

1. 発表者名 Sashi Novitasari, Quoc-Truong Do, Sakriani Sakti, Dessi Lestari, Satoshi Nakamura
2. 発表標題 Construction of English-French Multimodal Affective Conversational Corpus from Drama TV Series
3. 学会等名 LREC 2018 (国際学会)
4. 発表年 2018年

1. 発表者名 Sashi Novitasari, Quoc-Truong Do, Sakriani Sakti, Dessi Lestari, Satoshi Nakamura
2. 発表標題 Multi-modal Muli-task Deep Learning for Speaker and Emotion Recognition of TV-series Data
3. 学会等名 Oriental COCOSDA 2018 (国際学会)
4. 発表年 2018年

1. 発表者名 Sahoko Nakayama, Takatomo Kano, Quoc-Truong Do, Sakriani Sakti, Satoshi Nakamura
2. 発表標題 Japanese-English Code-Switching Speech Data Construction
3. 学会等名 Oriental COCOSDA 2018 (国際学会)
4. 発表年 2018年

1. 発表者名 Hiroki Tanaka, Hiroki Watanabe, Hayato Maki, Sakriani Sakti, Satoshi Nakamura
2. 発表標題 Single-trial Detection of Semantic Anomalies from EEG during Listening to Spoken Sentences
3. 学会等名 International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC 2018) (国際学会)
4. 発表年 2018年

1. 発表者名 Takuma Mori, Andros Tjandra, Sakriani Sakti, Satoshi Nakamura
2. 発表標題 Compressing End-to-End ASR Networks by Tensor-Train Decomposition
3. 学会等名 Interspeech 2018 (国際学会)
4. 発表年 2018年

1. 発表者名 Bin Wu, Sakriani Sakti, Satoshi Nakamura
2. 発表標題 Optimizing DPGMM Clustering in Zero Resource Setting Based on Functional Load
3. 学会等名 SLTU 2018 (国際学会)
4. 発表年 2018年

1. 発表者名 Tomoya Yanagita, Sakriani Sakti, Satoshi Nakamura
2. 発表標題 Incremental TTS for Japanese Language
3. 学会等名 Interspeech 2018 (国際学会)
4. 発表年 2018年

1. 発表者名 Andros Tjandra, Sakriani Sakti, Satoshi Nakamura
2. 発表標題 Machine Speech Chain with One-shot Speaker Adaptation
3. 学会等名 Interspeech 2018 (国際学会)
4. 発表年 2018年

1 . 発表者名 Sahoko Nakayama, Andros Tjandra, Sakriani Sakti, Satoshi Nakamura
2 . 発表標題 Speech Chain for Semi-Supervised Learning of Japanese-English Code-Switching ASR and TTS
3 . 学会等名 IEEE SLT ( 国際学会 )
4 . 発表年 2018年

1 . 発表者名 Andros Tjandra, Sakriani Sakti, Satoshi Nakamura
2 . 発表標題 Multi-scale Alignment and Contextual History for Attention Mechanism in Sequence-to-Sequence Model
3 . 学会等名 IEEE SLT ( 国際学会 )
4 . 発表年 2018年

1 . 発表者名 Quoc-Truong Do, Sakriani Sakti, Satoshi Nakamura
2 . 発表標題 Toward Multi-features Emphases Speech Translation: Assessment of Human Emphases Production and Perception with Speech and Text Clues
3 . 学会等名 IEEE SLT ( 国際学会 )
4 . 発表年 2018年

1 . 発表者名 Kaho Osamura, Takatomo Kano, Sakriani Sakti, Satoshi Nakamura
2 . 発表標題 Using Spoken Word Posterior Features in Neural Machine Translation
3 . 学会等名 IWSLT 2018 ( 国際学会 )
4 . 発表年 2018年

1. 発表者名 Johanes Effendi, Sakriani Sakti, Katsuhito Sudoh, Satoshi Nakamura
2. 発表標題 Multi-paraphrase Augmentation to Leverage Neural Caption Translation
3. 学会等名 IWSLT 2018 (国際学会)
4. 発表年 2018年

1. 発表者名 Andros Tjandra, Sakriani Sakti, Satoshi Nakamura
2. 発表標題 Machine Speech Chain with Deep Learning
3. 学会等名 日本音響学会2018年秋季研究発表会
4. 発表年 2018年

1. 発表者名 Nurul Lubis, Michael Heck, Sakriani Sakti, Koichiro Yoshino, Satoshi Nakamura
2. 発表標題 Multimodal Database of Negative Emotion Recovery in Dyadic Interactions: Construction and Analysis
3. 学会等名 日本音響学会2018年秋季研究発表会
4. 発表年 2018年

1. 発表者名 中山佐保子, ドクオック チュオン, サクティ サクリアニ, 中村哲
2. 発表標題 日英コードスイッチング音声データの構築
3. 学会等名 日本音響学会2018年秋季研究発表会
4. 発表年 2018年

1. 発表者名 Johanes Effendi, Sakriani Sakti, Satoshi Nakamura
2. 発表標題 Visual Description Paraphrase Corpus Creation with Various Elementary Operations
3. 学会等名 日本音響学会2018年秋季研究発表会
4. 発表年 2018年

1. 発表者名 叶 高朋, Sakriani Sakti, 中村 哲
2. 発表標題 カリキュラムラーニングを用いた音声翻訳の学習戦略の提案
3. 学会等名 言語処理学会 第25回年次大会 (NLP2019)
4. 発表年 2019年

1. 発表者名 中山 佐保子, Andros Tjandra, Sakriani Sakti, 中村 哲
2. 発表標題 Machine Speech Chainに基づく半教師あり学習を用いた日英コードスイッチング音声の認識
3. 学会等名 言語処理学会 第25回年次大会 (NLP2019)
4. 発表年 2019年

1. 発表者名 Nurul Lubis, Sakriani Sakti, Koichiro Yoshino and Satoshi Nakamura
2. 発表標題 Affect-sensitive Dialogue Response Generation for Positive Emotion Elicitation
3. 学会等名 言語処理学会 第25回年次大会 (NLP2019)
4. 発表年 2019年

1. 発表者名 Johanes Effendi, Sakriani Sakti, Katsuhito Sudoh and Satoshi Nakamura
2. 発表標題 Enhancing Neural Machine Translation with Image-based Paraphrase Augmentation
3. 学会等名 言語処理学会 第25回年次大会 (NLP2019)
4. 発表年 2019年

1. 発表者名 Sashi Novitasari, Quoc Truong Do, Sakriani Sakti, Dessi Lestari and Satoshi Nakamura
2. 発表標題 Speaker and Emotion Recognition of TV-Series Data Using Multimodal and Multitask Deep Learning
3. 学会等名 言語処理学会 第25回年次大会 (NLP2019)
4. 発表年 2019年

1. 発表者名 Andros Tjandra, Sakriani Sakti and Satoshi Nakamura
2. 発表標題 Unifying Speech Recognition and Generation with Machine Speech Chain
3. 学会等名 言語処理学会 第25回年次大会 (NLP2019)
4. 発表年 2019年

1. 発表者名 Hiroki Watanabe, Hiroki Tanaka, Sakriani Sakti, Satoshi Nakamura
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, Satoshi Nakamura
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, Sakriani Sakti, Satoshi Nakamura
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, Sakriani Sakti, Satoshi Nakamura
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

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, Hiroki Tanaka, Sakriani Sakti, Satoshi Nakamura
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

1. 発表者名 Takatomo Kano, Sakriani Sakti, Satoshi Nakamura
2. 発表標題 Neural Machine Translation with Acoustic Embedding
3. 学会等名 IEEE Automatic Speech Recognition and Understanding (ASRU) (国際学会)
4. 発表年 2019年

1 . 発表者名 Sahoko Nakayama, Andros Tjandra, Sakriani Sakti, Satoshi Nakamura
2 . 発表標題 Zero-shot Code-switching ASR and TTS with Multilingual Machine Speech Chain
3 . 学会等名 IEEE Automatic Speech Recognition and Understanding (ASRU) ( 国際学会 )
4 . 発表年 2019年

1 . 発表者名 Johanes Effendi, Andros Tjandra, Sakriani Sakti, Satoshi Nakamura
2 . 発表標題 Listening while Speaking: Improving ASR through Multimodal Chain
3 . 学会等名 IEEE Automatic Speech Recognition and Understanding (ASRU) ( 国際学会 )
4 . 発表年 2019年

1 . 発表者名 Andros Tjandra, Sakriani Sakti, Satoshi Nakamura
2 . 発表標題 Speech-to-speech Translation between Untranscribed Unknown Languages
3 . 学会等名 IEEE Automatic Speech Recognition and Understanding (ASRU) ( 国際学会 )
4 . 発表年 2019年

1 . 発表者名 Andros Tjandra, Berrak Sisman, Mingyang Zhang, Sakriani Sakti , Haizhou Li and Satoshi Nakamura
2 . 発表標題 VQVAE Unsupervised Unit Discovery and Multi-scale Code2Spec Inverter for Zerospeech Challenge 2019
3 . 学会等名 INTERSPEECH ( 国際学会 )
4 . 発表年 2019年

1 . 発表者名 Sashi Novitasari, Andros Tjandra, Sakriani Sakti and Satoshi Nakamura
2 . 発表標題 Sequence-to-Sequence Learning via Attention Transfer for Incremental Speech Recognition
3 . 学会等名 INTERSPEECH ( 国際学会 )
4 . 発表年 2019年

1 . 発表者名 Ivan Halim Parmonangan, Hiroki Tanaka, Sakriani Sakti , Shinnosuke Takamichi and Satoshi Nakamura
2 . 発表標題 Speech Quality Evaluation of Synthesized Japanese Speech using EEG
3 . 学会等名 INTERSPEECH ( 国際学会 )
4 . 発表年 2019年

1 . 発表者名 Ewan Dunbar, Robin Algayres, Julien Karadayi, Mathieu Bernard, Juan Benjumea, Xuan-Nga Cao, Lucie Miskic, Charlotte Dugrain, Lucas Ondel, Alan W Black, Laurent Besacier, Sakriani Sakti and Emmanuel Dupoux
2 . 発表標題 The Zero Resource Speech Challenge 2019: TTS without T
3 . 学会等名 INTERSPEECH ( 国際学会 )
4 . 発表年 2019年

1 . 発表者名 Holy Lovenia, Hiroki Tanaka, Sakriani Sakti, Ayu Purwarianti, Satoshi Nakamura
2 . 発表標題 Speech Artifact Removal from EEG Recordings of Spoken Word Production with Tensor Decomposition
3 . 学会等名 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) ( 国際学会 )
4 . 発表年 2019年

1. 発表者名 Andros Tjandra, Sakriani Sakti, and Satoshi Nakamura
2. 発表標題 End-to-end feedback loss in speech chain framework via straight-through estimator
3. 学会等名 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (国際学会)
4. 発表年 2019年

1. 発表者名 Marco Vetter, Sakriani Sakti, and Satoshi Nakamura
2. 発表標題 Cross-lingual speech-based ToBI label generation using bidirectional LSTM
3. 学会等名 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (国際学会)
4. 発表年 2019年

〔図書〕 計0件

〔出願〕 計1件

産業財産権の名称 スピーチチェーン装置、コンピュータプログラムおよびDNN音声認識・合成相互学習方法	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 奈良先端科学技術大学院大学
産業財産権の種類、番号 特許、特願2018-1538	出願年 2018年	国内・外国の別 国内

〔取得〕 計0件

〔その他〕

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究分担者	中村 哲  (Nakamura Satoshi)  (30263429)	奈良先端科学技術大学院大学・データ駆動型サイエンス創造センター・教授   (14603)	