JAIST Repository
https://dspace.jaist.ac.jp/

| | |
|---|---|
| Title | 情動による感情音声発話のメカニズムとモデル化に関する研究 |
| Author(s) | 党，建武 |
| Citation | 科学研究費助成事業研究成果報告書: 1-6 |
| Issue Date | 2019-05-30 |
| Type | Research Paper |
| Text version | publisher |
| URL | http://hdl.handle.net/10119/16041 |
| Rights | |
| Description | 基盤研究(C)（一般），研究期間：2016～2018，課題番号：16K00297，研究者番号：80334796，研究分野： 言語科学、音声生成 |

JAIST
JAPAN
ADVANCED INSTITUTE OF
SCIENCE AND TECHNOLOGY

Japan Advanced Institute of Science and Technology

様　式　Ｃ－１９、Ｆ－１９－１、Ｚ－１９　（共通）

# 科学研究費助成事業　　研究成果報告書

令和　　元 年　　５ 月 ３０ 日現在

機関番号：１３３０２
研究種目：基盤研究(C)（一般）
研究期間：2016～2018
課題番号：１６Ｋ００２９７

研究課題名（和文）情動による感情音声発話のメカニズムとモデル化に関する研究

研究課題名（英文）A study on production mechanism and modeling of emotional speech

研究代表者
党　建武（Dang, Jianwu）

北陸先端科学技術大学院大学・先端科学技術研究科・教授

研究者番号：８０３３４７９６

交付決定額（研究期間全体）：（直接経費）　　3,600,000円

研究成果の概要（和文）：本研究では、発話計画の脳活動、発話運動及び行動学的な観測から感情音声の発話メカニズムの究明とそのモデル化を行った。発話運動分析により、悲しみの発話では舌背の位置が低く後ろ寄りになり、幸せの発話時に舌背が前寄りになることがわかった。眼球運動と脳電図を用いて連続朗読の脳活動を全時間帯で計測した。再構成した発話計画、語意処理と文法処理の脳ネットワークを分析したところ、発話計画ネットワークは単語毎に活動しており、語意処理ネットワークが朗読の前半でより活発し、文法処理ネットワークが後半ではより活発していることを明らかにした。それは、後半になると文脈情報が豊かになり語意処理しやすくなることを示唆した。

研究成果の学術的意義や社会的意義
これまでの音声に関する脳研究のほとんどは、音声資料は孤立の言葉を使用するか、または特定な時間区間のみに注目している。もう一つの問題は、音声生成過程には発話運動による筋電の影響で脳電図の分析が困難である。本研究では、眼球運動と脳電図を用いて連続文章の朗読における脳活動を全時間帯で計測することによって音声生成における脳ネットワークの動的特性を分析する。本研究は非侵襲的な手法により脳活動を連続的全時間帯で計測するのは音声生成町域で初めての挑戦である。実験ディザインとデータ分析への貢献は、眼球運動のデータを脳活動の時間的な指標として使用し、脳ネットワーク活動と時間との対応関係を明確にした。

研究成果の概要（英文）：In this study, we investigated production mechanism of emotional speech and its modeling based on articulatory movements, neurological activities and behavioral observations. Articulation analysis showed that the position of the tongue dorsum moves lower and backward for sad speech, while the dorsum moves forward for happiness. We used eye movement and EEG to observe the brain activations in the full time scale during continuously oral reading. Analysis of the brain networks reconstructed based on EEG showed that the speech planning network acts for each word, and the semantic processing network is more active in the first half of reading, while the grammar processing network becomes more active in the second half. This suggests that the semantic processing gets easier since the context information becomes rich in the second half of the continuous reading.

研究分野：言語科学、音声生成

キーワード：音声生成　感情音声　脳ネットワーク　神経学的モデル　音声生成モデル

**様　式　C－１９、F－１９－１、Z－１９、CK－１９（共通）**

１．研究開始当初の背景

音声情報は、言語情報と個人性・感情などの非言語情報を含んでいる。特に、感情などの非言語情報はコミュニケーションに欠かせない役割を果たしており、感情を表現しない音声は日常会話にほとんど見られない。人間の感情の生成、変化及び知覚は、自分の心情と周囲の環境との相互作用によるものである。それゆえ、本研究では、人間感情の生成と知覚の中枢機構とした脳のメカニズムのみならず、音声による感情の表現と認識、またソーシャルメディアによる感情の表現や心情の表白などの影響を総合的に考えて、言語音声を介して、人間の感情生成と知覚に関するメカニズムを考察することにした。

音声感情の生成と知覚に関する研究が音声信号処理の観点から数多く行われてきた。それらの研究では、主に音声の音響特性に着目しているが、感情生成と感情知覚における人間のメカニズムを十分に考慮に入れていなかった。如何に人間のメカニズムを音声の感情生成と知覚の研究に取り入れるかが本研究の目標の一つである。

インターネットの普及とともに、ソーシャルメディアは人間が気分を晴らし、感情を表し、心情を伝える重要な場となっている。うつ病は感情精神病、感情障害、気分障害の病気といわれ、大きな社会問題につながっている。ソーシャルメディアを通して、うつ病患者を助け、またはうつ病潜在患者を早期発見する研究もある。本研究ではソーシャルメディアの情報によりうつ病患者の支援やうつ病潜在患者の早期発見の方法も検討した。

近年、fMRI や脳磁場計測装置(MEG)及び脳電図（EEG）などの非侵襲脳機能計測装置を用いる脳科学研究では、言葉の意味の理解と感情情報の処理が統合するタイミングを正確に測定できるようになってきた。NICT の井原らは MEG を用いて音声処理の過程における感情的な情報は、脳の左右両半球の前頭部で統合されることを観測し、感情の情報が言葉の解釈に影響を与える仕組み、すなわち、言語情報と感情情報の統合に関する脳のメカニズムを提唱した。しかし、これまでの音声に関する脳研究のほとんどは、音声資料は孤立の言葉を使用するか、または特定な時間区間のみに注目している。もう一つの問題は、音声生成過程には発話運動による筋電の影響で脳電図の分析が困難である。そのため、音声に関連する脳研究のほとんどは聴覚の側面から行っている。近年、EEG 信号処理技術も長足の進歩があり、特に UCSD が提供した無料ソフトウェアパッケージ（EEGLAB）により、発話運動に伴う脳電信号の分析が可能となった。

スパイキング・ニューラルネットは従来のニューラルネットと異なり、脳神経細胞が発する活動電位（スパイク）の時間系列（パルス）としてニューラルネット上で再現したことである。生物の脳の神経回路網を工学的に再現した（人工的）ニューラルネットの次世代モデルであるといわれた。本研究では、スパイキング・ニューラルネットを用いて、発話の神経学的モデルへの応用を試みることにした。

２．研究の目的

音声情報は、言語情報と個人性・感情などの非言語情報を含んでいる。特に、感情などの非言語情報はコミュニケーションに欠かせない役割を果たしており、感情を表現しない音声は日常会話にほとんど見られない。人間の感情の生成、変化及び知覚は、自分の心情と周囲の環境との相互作用による結果である。それゆえ、本研究では、人間感情の生成と知覚の中枢機構とした脳のメカニズムのみならず、音声による感情の表現と知覚、またソーシャルメディアによる気分晴らし、心情の表白などの影響を総合的に考えて、言語音声を介して、人間の感情生成と知覚に関するメカニズムを考察することを目的とした。

３．研究の方法

本研究の目的を達成するため、下記の研究方法を用いた。

（1） 感情音声発話と知覚のメカニズムとその応用に関する研究では、①我々は音声信号だけでなく磁気センサーにより感情音声の発話運動も記録し、音声の感情表現は音響特徴量と調音動作との関連を分析した。②人間が音声を知覚する際、ボトムアップの方式を利用し階層の異なる特徴量を統合して音声を知覚する。そのメカニズムに基づいた音響特徴量・素朴的特徴量・感情特徴量の三層モデルを介して、人間の知覚メカニズムを音声感情の推定と認識及び音質変換に応用した。③人間の感情音声知覚の知見を現在流行っているディープニューラルネットワークに取り入れて感情の認識率を向上させる手法を研究した。④ソーシャルメディアでの気分晴らし、心情の表白などの情報を追跡して、うつ病患者の支援またはうつ病潜在患者の早期発見の方法を検討した。

（2） 音声生成と知覚過程における神経学的な動的特性に関する研究では、128 チャンネル脳電図を用いて全時間帯で脳活動の動的な活動を分析する。①動詞、動詞寄り動名詞、名詞寄り動名詞、名詞の 4 種類音声資料を用い、名詞と動詞の神経分離のメカニズムを究明する。② 眼球運動と脳電図を用いて、連続朗読音声における音声生成発話メカニズムを研究する。そこで、EEGLAB を用いた脳信号源の再構成方法により、脳活動の時間系列を求め、fMRI に基づいた脳機能的サブネットワークにより脳活動の時間系列をフィルタリングして、脳ネットワークの動的特性を求めた。③ 調音に唇の丸めると丸めない同義語セットを被験者に聴取させ、EEG 信号を用い脳信号源の再構成により、音声知覚の運動理論の神経学的メカニズムを検証した。

（3） 発話運動の生理学的・神経学的モデルに関する研究では、①MRI 動画、声道内音圧と音声信号を用いて、形態学的、機械的、および音響的測定により、非鼻音化音声、特に有声破裂音の生成メカニズムを調べて、音声生成における軟口蓋を 2 層モデルを提案した。②

声門音源と声道パラメータを同時に推定し、合成による分析法を用い感情音声への影響要因を考察した。③　生物の脳の神経回路網を工学的に再現できるスパイキング・ニューラルネットを用いて、単一のニューロンはマルチカテゴリー分類に対応することができる効率的なマルチスパイク学習則を研究した。

４．研究成果

(1) 感情音声発話と知覚のメカニズムとその応用

① 感情音声発話メカニズムの研究では、2 人の被験者に平静、怒り、悲しみ及び幸せの感情を持つように発話させ、音声を収録する同時に磁気センサーにより調音運動も記録した。調音分析により、幸せと怒りの発話は一般的に平静と悲しみのそれよりピーチが高く、口がより開いている。悲しみ発話の場合、舌背の位置が低く、後ろ寄り、それに対して幸せの発話時に舌背が前寄りへ移行することがわかった。音響分析により、第一と第二共振周波数からなる(F1-F2)平面では、平静音声に比べ、感情音声の分散が大きくなっている。聴取実験を実施したところ、聴者は表現された感情をほぼ正しく認識できた。調音と感情との関係について、声の高さが高いほど、大きさが大きいほど、感情の知覚はしやすくなる傾向があることが分かった（研究発表[21]参照)。

② 人間のメカニズムを考慮に入れた音声感情の推定と音質変換：声を知覚する時に、人間がまず音声の高さ・大きさ・明るさ・太さなどの素朴的な特徴量を知覚し、それらの量に基づいて総合的に音声の感情を推測するとされる。人間のこのようなメカニズムを考慮に入れて、音響特徴量・素朴的特徴量・感情特徴量の三層モデルを構築し、音声感情の推定や認識及び音質変換などに応用した。それらの成果[2, 3, 7]は学術論文として発表した。

③ 人間のメカニズムを考慮に入れた音声感情の認識：現在、音声感情認識の研究の大半は深層学習による方法に基づいて行われている。そこで、スペクトログラムを用い、深層ニューラルネットワークの構造を変更しながら研究を行っており、人間の感情認識の知識はほとんど使われていない。本研究では、人間が音声感情知覚によく使っている F0、ストレスと持続時間などの低レベル特徴量 (low-level descriptor, LLDs)をスペクトログラムに結合することで、感情の誤認識率は 30%減少された。これらの研究成果[4, 6]は国際学会で発表された。

④ うつ病は感情精神病、感情障害、気分障害の病気といわれ、大きな社会問題になっている。本研究では、うつ病患者とその潜在患者のソーシャルメディアで発表したものを追跡することにより、うつ病患者の自殺などの極端行為の防止及びうつ病の潜在患者の早期発見の方法を提案した。その成果[1]を学術論文として発表した。

(2) 音声生成と知覚過程における神経学的な動的特性に関する研究

① 知覚過程の語意解読の神経学的メカニズムの研究：この部分では、128 チャンネル脳電図を用いて、名詞と動詞の神経解離およびのワードと疑似ワードの知覚の二つの研究を行った。名詞と動詞の神経解離について、名詞、名詞寄り動名詞、動詞寄り動名詞と動詞の 4 種類音声資料を用い、名詞と動詞の間の神経解離の決定要因は文法分類または意味的な違いによるものであるかを調べた。その結果、文法的分類より意味的な違いが名詞と動詞の神経分離の要因とであることを明らかにした。この成果[9]は Neuroscience で発表された。また、真の言葉と疑似言葉の知覚について、脳電図により全時間帯で観測し、脳の信号源を再構成して、分析した。その結果より、言葉認知過程を 6 つの区間に区切ることができて、真の言葉と疑似言葉の処理脳ネットワークはそれぞれの接続パターンや領野間の情報フローは異なっていることが分かった。この成果[11, 17]は国際学会で発表された。

② 眼球運動と脳電図を用いた連続発話の神経学的メカニズム：　本研究では、従来研究と異なり、孤立単語でなく連続文章を朗読しながら、眼球運動、音声信号と脳電図（EEG）を収録した。被験者がある単語を注目してからこの単語の発話までの時間は発話計画の潜時時間とし、脳の発話計画の動的特性を全時間帯で解析する。行動学の分析により、連続朗読する際、文章の後部に行くほど発話計画の潜時時間は短くなり、文尾単語の潜時時間は文頭単語のそれより 40%短くなっている。同じ文章の単語の語順をランダムにした無意味文に対して、発話計画の潜時時間は単語の位置に依存しないことが判明し、発話の潜時時間は個々の単語のみならず、文章の理解もかかわっていることを明らかにした。これらの成果[12]は国際学会で発表した。さらに、fMRI に基づいた視覚、発話計画、語意処理と文法処理の脳ネットワークを用いて、EEG により再構成した脳ネットワークを最適化した。その結果、発話計画脳ネットワークは単語ごとにモーターコマンドを作り出し、朗読の前半で語意処理ネットワークがより活発しているが、後半になると文法処理のネットワークがより活発していることが分かった。潜時時間の変化も取り入れて考えると、連続文章の後半になると、文章理解のため文法処理ネットが活発しており、それに対して文脈情報が豊かになって語意処理ネットワーの活動は軽減されたことを示唆した。その成果[5]は国際学会に発表された。

③ 音声生成と音声知覚と相互作用に関する神経学的メカニズム：音声知覚の運動理論によると、音声生成と音声知覚過程には同じ音声ジェスチャーを使っているとされている。そうなら、音声知覚時にその音声を生成する発話器官に関連する脳領野が反応すべきである。本研究では、唇の丸めると丸めない同義語セットを用いて被験者に聴取させ、EEG 信号を用い脳信号源の再構成により、唇の脳領野に活動があるかを調べた。その結果は、音声知覚の運動理論の神経学的メカニズムを実証できた。その内容[10]を国際学会で発表した。

(3) 発話運動の生理学的・神経学的モデルに関する研究

① 音声生成の声道モデルに関する研究：これまで主に口腔側の声道のモデル化を中心として生理学的発話モデルを構築した。鼻腔の影響を取り入れるため、本研究では、形態学的、機械的、および音響的測定を使用して、非鼻音化音声の生成における軟口蓋の機能を調べました。MRI 映画を分析したところ、日本人話者が有声破裂子音を発するときに、声門上腔に有意な拡大は見られず、口蓋の厚さは母音によって異なることがわかった。これらの結果に基づいて、非鼻音化音声の生成における 2 層モデルを提案した。音声生成の声道モデルに関する成果[15, 20]は学術論文誌と国際学会論文で発表された。

② 音声生成の音源モデル：既知の感情音声の支配的な要因（例えば、F0、強度、および持続時間）の影響を除去した後で、声門音源および声道の特徴量が音声感情の知覚への寄与はまだ不明である。本研究では、声門音源と声道パラメータを同時に推定し、合成による分析法を用い感情音声への影響要因を考察した。その結果は、F0、強度および持続時間の要因を抑えた場合、声道情報は感情の valence と arousal の知覚に寄与することが分かった。音声生成の音源モデルに関する成果は研究発表の[7,19]がある。

③ 音声生成に関する神経学的モデル: スパイキング・ニューラルネットは、私たち生物の脳の神経回路網（生物的ニューラルネット）を工学的に再現した（人工的）ニューラルネットの次世代モデルであるといわれている。本研究では、レートとタイミングの両方に基づくスパイクパターンを分類するため、ニューロンの入力パターンに応じて異なる数の出力スパイクを発火するように訓練することができ、単一のニューロンはマルチカテゴリー分類に対応することができる効率的なマルチスパイク学習則を提案した。その結果は、提案法が他のベンチマークより効率は高く、実装は簡単であることを示した[8, 13]。将来、音声生成と音声知覚過程をこのスパイキング・ニューラルネットワークにより模擬する予定です。

研究発表


[1] Z. Peng, Q. Hu, J. Dang (2019) Multi-kernel SVM based depression recognition using social media data, International Journal of Machine Learning and Cybernetics, Vol. 10, pp 43–57.

[2] Xingfeng LI, Masato AKAGI (2019) “Improving multilingual speech emotion recognition by combining acoustic features in a three-layer model," Speech Communication, 110, pp.1-12.

[3] Y. Xue, Y. Hamada, M. Akagi (2018) “Voice conversion for emotional speech: Rule-based synthesis with degree of emotion controllable in dimensional space”, Speech Comm., 102, pp. 54-67

[4] L. Guo, L. Wang, J. Dang, L. Zhang, H. Guan (2018). “ A feature fusion method based on extreme learning machine for speech emotion recognition”, ICASSP, April, 15-20, 2018 Canada

[5] B. Zhao, J. Huang, J. Dang, G. Zhang (2018). Revealing Spatiotemporal Brain Dynamics of Speech Production Based on EEG and Eye Movement. INTERSPEECH, Sep. 2-6, 2018, India.

[6] L. Zhang, L. Wang, J. Dang, L. Guo and H. Guan (2018) Convolutional Neural Network with Spectrogram and Perceptual Features for Speech Emotion Recognition, ICONIP, 2018, Cambodia

[7] Y. Li, J. Li and M. Akagi, (2018) “Contributions of the glottal source and vocal tract cues to emotional vowel perception in the valence-arousal space,” J. Acoust. Soc. Am., 144, pp. 908-916.

[8] Qiang Yu, Longbiao Wang and Jianwu Dang (2018) Efficient Multi-Spike Learning with Tempotron-like LTP and PSD-like LTD, ICONIP, 13-16, Dec, 2018, Siem Reap, Cambodia

[9] Zhao B, Dang J, Zhang G (2017). EEG Source Reconstruction Evidence for the Noun-Verb Neural Dissociation along Semantic Dimensions[J]. Neuroscience, 2017, 359.

[10] Zhao Bin; Gaoyan Zhang; Jianwu Dang (2017) A Neuro-Experimental Evidence for the Motor Theory of Speech Perception., InterSpeech, 2017, Sweden

[11] Y. Si; J. Dang; G. Zhang (2017), Global monitoring of dynamic functional interactions in the brain during Chinese verbs perception, ISSP, 2017, China

[12] Jinfeng Huang; Di Zhou; Jianwu Dang (2017), Estimation of Speech-planning mechanism based on eye movement, ISSP2017, China

[13] Yu Q, Wang L, Dang J. (2017), Neuronal Classifier for both Rate and Timing-Based Spike Patterns[C] ICONIP. Springer, Cham, 2017:759-766.

[14] Jianwu Dang, Jianguo Wei, Kiyoshi Honda, and Takayoshi Nakai (2016). “A study on transvelar coupling for non-nasalized sounds,” J. Acoust. Soc. Am. 139 (1), pp.）441–454

[15] J. Wang, J. Zhang, K. Honda, J. Wei, J. Dang (2016), “Audio-visual speech recognition integrating 3D lip information obtained from the Kinect” Multimedia Systems 22:315–323

[16] Bin Zhao, Jianwu Dang, Gaoyan Zhang (2016), “Investigation of Noun-Verb Dissociation Based on EEG Source Reconstruction”, APSIPA, 2016, Korea

[17] Yuke Si, Jianwu Dang, Gaoyan Zhang (2016), “Investigation of the Spatiotemporal Dynamics of the Brain during Perceiving Words”, ISCLP 2016, China

[18] Bin Zhao, Jianwu Dang, Gaoyan Zhang (2016), “EEG Evidence for a Three-Phase Recurrent Process during Spoken Word Processing”, ISCSLP 2016, China

[19] S. Fan, K. Honda, Jianwu Dang, H. Feng (2016), “Effects of Subglottal-Coupling and Interdental-Space on Formant Trajectories during Front-to-Back Vowel Transitions in Chinese,” INTERSPEECH 2016, USA

[20] J. Dang, S. Wang, M. Unoki (2016), “Investigations into vowel and consonant structures in

articulatory and auditory spaces using Lapalacian eigenmaps," ICASSP 2016, China.
[21] Erickson, D., Zhu, C., Kawahara, S., Suemitsu (2016), A. "Articulation, acoustics and perception of Mandarin Chinese emotional speech", Open Linguistics, 620-635, 2016

５．主な発表論文等

〔**雑誌論文**〕（**計** 17 **件**）

（1） Zhichao Peng, Qinghua Hu, Jianwu Dang (2019) "Multi-kernel SVM based depression recognition using social media data", International Journal of Machine Learning and Cybernetics, Vol. 10, pp. 43–57. 査読有

（2） Xingfeng LI, Masato AKAGI (2019) "Improving multilingual speech emotion recognition by combining acoustic features in a three-layer model," Speech Comm., 110, pp.1-12. 査読有

（3） Y. Xue, Y. Hamada, M. Akagi (2018) "Voice conversion for emotional speech: Rule-based synthesis with degree of emotion controllable in dimensional space", Speech Comm., 102, pp. 54-67 査読有

（4） W. Lu, Z, Chen, L. Li, X. Cao, J. Wei, N. Xiong, J. Li and J. Dang (2018). " Watermarking Based on Compressive Sensing for Digital Speech Detection and Recovery, " Sensors , 18(7), 2390; doi:10.3390/s18072390 査読有

（5） Z. Oo, L. Wang, K. Phapatanaburi, M. Iwahashi, S. Nakagawa, J. Dang (2018) "Phase and reverberation aware DNN for distant-talking speech enhancement", Multimedia Tools Appl, 77, 18865–18880. 査読有

（6） C. Cai, K. Ogawa, T. Kochiyama, H. Tanaka, H. Imamizu, (2018) "Temporal recalibration of motor and visual potentials in lag adaptation in voluntary movement", NeuroImage, 172, pp.654-662 査読有

（7） Zhao B, J. Dang, Zhang G (2017). EEG Source Reconstruction Evidence for the Noun-Verb Neural Dissociation along Semantic Dimensions[J]. Neuroscience, 2017, 359. 183-195. 査読有

（8） Lu W, Zhang J, Zhao X, J. Wang, J. Dang (2017). Multimodal sensory fusion for soccer robot self-localization based on long short-term memory recurrent neural network[J]. Journal of Ambient Intelligence & Humanized Computing, 2017, 8(11): 885–893. 査読有

（9） Wang J, Zhang J, Lu W, Wei J, J. Dang. Automatic speech recognition with robot noise[J]. Qinghua Daxue Xuebao/journal of Tsinghua University, 2017, 57(2):153-157. 査読有

（10） Wang G, J. Dang, Kong J. Modeling of the tongue tip in Standard Chinese using MRI[J]. Journal of Tsinghua University, 2017, 57(2):158-163. 査読有

（11） 田中宏和, (2017) "運動制御と感覚処理の最適理論", 日本ロボット学会誌, 25, pp.500-505 査読有

（12） 赤木正人(2017)"コンピュータが歌うということ" 日本音響学会誌, 73, pp.210-211. 査読有

（13） Kawahara, S., Erickson, D., Suemitsu, A.(2017) "The phonetics of jaw displacement in Japanese vowels", Acoustical Science and Technology, 38, pp.99-107 査読有

（14） J. Dang, Jianguo Wei, Kiyoshi Honda, and Takayoshi Nakai (2016). "A study on transvelar coupling for non-nasalized sounds," J. Acoust. Soc. Am. 139 (1), pp.441–454, 査読有

（15） J. Wang, J. Zhang, K. Honda, J. Wei, J. Dang (2016), "Audio-visual speech recognition integrating 3D lip information obtained from the Kinect" Multimedia Systems 22:315–323 査読有

（16） Erickson, D., Zhu, C. Kawahara, S. Suemitsu (2016), A. "Articulation, acoustics and perception of Mandarin Chinese emotional speech", Open Linguistics, 620-635, 2016 査読有

（17） 田中宏和, "脳を理解するとはどういうことかーある計算論的神経科学者の頭の中", BRAIN and NERVE-神経研究の進歩, pp.1379-1384, 2016 査読有

〔**学会発表**〕（**計** 23 **件**）

（1） L. Guo, L. Wang, J. Dang, L. Zhang, H. Guan (2018). " A feature fusion method based on extreme learning machine for speech emotion recognition", ICASSP, 2018

（2） Bin Zhao, Jinfeng Huang, J. Dang, Gaoyan Zhang, Revealing Spatiotemporal Brain Dynamics of Speech Production Based on EEG and Eye Movement. INTERSPEECH, 2018.

（3） D. Li, L. Wang, J. Dang, Meng Liu, Zeyan Oo, Seiichi Nakagawa, Haotian Guan, Xiangang Li, Multiple Phase Information Combination for Replay Attacks Detection. INTERSPEECH, 2018.

（4） L. Guo, L. Wang, J. Dang, L. Zhang, H. Guan, X. Li. Speech Emotion Recognition by Combining Amplitude and Phase Information Using Convolutional Neural Network. INTERSPEECH, 2018.

（5） L. Zhang, L. Wang, J. Dang, Lili Guo, and Qiang Yu, Gender-Aware CNN-BLSTM for Speech Emotion Recognition, ICANN, 2018

（6） L. Zhang, L. Wang, J. Dang, Lili Guo and Haotian Guan, "Convolutional Neural Network with Spectrogram and Perceptual Features for Speech Emotion Recognition", ICONIP, 2018

（7） Qiang Yu, Longbiao Wang and J. Dang, Efficient Multi-Spike Learning with Tempotron-like LTP and PSD-like LTD, ICONIP, 2018

（8） Di Zhou, Jinfeng Huang and J. Dang, Investigation of the Comprehension Process during Silent Reading based on Eye Movements, ISCSLP, 2018

（9） Meng Liu, Longbiao Wang, Zeyan Oo, J. Dang, Replay Attacks Detection Using Phase and Magnitude Features with Various Frequency Resolutions, ISCSLP, 2018
（10） Zhao Bin; Gaoyan Zhang; J. Dang, A Neuro-Experimental Evidence for the Motor Theory of Speech Perception., InterSpeech2017
（11） Yu Q, Wang L, J. Dang, Neuronal Classifier for both Rate and Timing-Based Spike Patterns[C]// International Conference on Neural Information Processing. Springer, Cham, 2017:759-766.
（12） Zhao C, Wang L, J. Dang, Phonemic Restoration Based on the Movement Continuity of Articulation[J].International Conference on Neural Information Processing, 2017.
（13） Y. Si; J. Dang, G. Zhang (2017), Global monitoring of dynamic functional interactions in the brain during Chinese verbs perception, International Seminar on Speech Production (ISSP), 2017
（14） Jinfeng Huang; Di Zhou; J. Dang, Estimation of Speech-planning mechanism based on eye movement, ISSP2017, 2017
（15） M. Yuan; L. Wang; H. Wang; J. Dang; R. Yu, Articulatory Features for Robust Intelligibility Assessment of Dysarthria Speech, International Seminar on Speech Production (ISSP), 2017
（16） Zhichao Peng, Zhi Zhu, Masashi Unoki, J. Dang, and M. Akagi (2017), “Speech Emotion Recognition Using MPCRNN based on Gammatone auditory Filterbank”, APSIPA2017.
（17） Li, Y., Sakakibara, K-I., Morikawa, D., and Akagi, M “Commonalities of glottal sources and vocal tract shapes among speakers in emotional speech”, ISSP, 2017.
（18） Bin Zhao, J. Dang, Gaoyan Zhang, “Investigation of Noun-Verb Dissociation Based on EEG Source Reconstruction”, APSIPA, 2016
（19） Xue, Y., Hamada, Y., and Akagi, M. “Voice Conversion to Emotional Speech based on Three-layered Model in Dimensional Approach and Parameterization of Dynamic Features in Prosody”, APSIPA, 2016
（20） Yuke Si, J. Dang, Gaoyan Zhang, “Investigation of the Spatiotemporal Dynamics of the Brain during Perceiving Words”, ISCSCLP 2016
（21） Bin Zhao, J. Dang, Gaoyan Zhang, “EEG Evidence for a Three-Phase Recurrent Process during Spoken Word Processing”, ISCSLP 2016
（22） S. Fan, K. Honda, J. Dang, H. Feng, “Effects of Subglottal-Coupling and Interdental-Space on Formant Trajectories during Front-to-Back Vowel Transitions in Chinese,” INTERSPEECH 2016
（23） J. Dang, S. Wang, M. Unoki, “Investigations into vowel and consonant structures in articulatory and auditory spaces using Lapalacian eigenmaps,” ICASSP 2016.

**〔図書〕（計 1 件）**

(1) Qiang Fang, Jianwu Dang, Pascal Perrier, Jianguo Wei, Longbiao Wang, Nan Yan (Eds.), Springer, “Studies on Speech Production”, 2018, ISBN 978-3-030-00125-4, pp.236

**〔産業財産権〕**

**○出願状況（計 0 件）**

**○取得状況（計 0 件）**

**〔その他〕**

https://www.jaist.ac.jp/profiles/info.php?profile_id=277&textfield=&andor=&obj1=&obj2=&obj3=&obj4=&obj5=

６．研究組織

(1)研究分担者

研究分担者氏名：赤木　正人
ローマ字氏名：アカギ　マサト
所属研究機関名：北陸先端科学技術大学院大学
部局名：先端科学技術研究科
職名：教授
研究者番号（8 桁）：２０２４２５７１

研究分担者氏名：田中　宏和
ローマ字氏名：タナカ　ヒロカズ
所属研究機関名：北陸先端科学技術大学院大学
部局名：先端科学技術研究科
職名：准教授
研究者番号（8 桁）：００３３２３２０

研究分担者氏名：　末光　厚夫
ローマ字氏名：スエミツ　アツオ
所属研究機関名：札幌保健医療大学
部局名：保健医療学部
職名：教授
研究者番号（8 桁）：２０４２２１９９

**※科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属されます。**