

令和元年6月3日現在

機関番号：12601

研究種目：基盤研究(A) (一般)

研究期間：2014～2017

課題番号：26240022

研究課題名(和文) 話者を単位とした世界諸英語の自動分類とそれに基づく国際コミュニケーション支援

研究課題名(英文) Individual-based and automatic clustering of World Englishes and its application to assist international communication

研究代表者

峯松 信明 (Minematsu, Nobuaki)

東京大学・大学院工学系研究科(工学部)・教授

研究者番号：90273333

交付決定額(研究期間全体)：(直接経費) 29,980,000円

研究成果の概要(和文)：世界諸英語の発音多様性に着眼し、1)話者を単位とした世界諸英語の発音分類を行うべく、任意の二話者の発音距離推定技術を構築し、2)特定の学習者を中心に据えて、諸英語を俯瞰できる諸英語発音ブラウザを開発した。しかし、発音の相違が常にコミュニケーション上問題となる訳ではない。そこで、外国語訛りとコミュニケーションエラーに着眼し、3)日本人英語を米国人がどこで聞き誤るのかを予測する技術を構築し、4)聞き取りを誤る他の要因として環境要因にも着眼し、日本人の英語聴取能力の頑健性向上を検討した。5)日本語学習者に対しても、了解性の高い発音を効率的に習得できるように、韻律読み上げチュートを開発した。

研究成果の学術的意義や社会的意義

本プロジェクトは、英語発音の多様性に着眼し、学習者自らを世界諸英語の中心に据えて、世界諸英語を俯瞰するブラウザを構築し、発音の多様性への意識を向上させる機会を提供した。このような試みはこれまでに例がなく、国内外の関連する学会にて、7回の招待講演、基調講演を行い、学術的に大きく貢献した。また、発音の崩れによる了解性の低下を避けるべく、日本語学習者向けに構築した韻律読み上げチュートは、既に14ヶ国語に翻訳され、世界中の教育機関で使われるに至っている。このシステムは、本プロジェクト開始前から開発を進めていたが、2019年4月時点で、40カ国、140回の講習会を重ねており、社会的意義は極めて高い。

研究成果の概要(英文)：This project primarily focused on pronunciation variation in World Englishes and built 1) a technique of predicting the pronunciation distance between any two English speakers for the purpose of drawing a pronunciation map, and 2) a bird's-eye browser of WE pronunciations with a specific speaker located at the origin of the map. However, it is a fact that two speakers with different pronunciations can not rarely communicate smoothly with each other. Then, we proposed 3) a technique of predicting when native American listeners misunderstand English spoken with Japanese accents. Misunderstanding can happen not only due to accents but to other environmental factors. So, we analyzed 4) what kind of environmental factors easily reduce listening performance of Japanese listeners, and we proposed a method to improve robustness of their listening. Finally, we developed 5) prosodic reading tutor for Japanese learners to improve naturalness and intelligibility of their spoken Japanese.

研究分野：音声工学・科学

キーワード：世界諸英語 発音自動分類 音声の構造的表象 機械学習 発音の了解性 聴取の頑健性 分節的・韻律的特徴 韻律読み上げチュート

## 様式 C-19、F-19-1、Z-19、CK-19 (共通)

### 1. 研究開始当初の背景

国際社会において英語が使われる場合、その多くは非母語話者による英語であり、当然、母語訛りを伴う。発音以外にも統語、語彙選択、綴り、談話戦略など様々な面で多様化している(世界諸英語)が、本研究では発音の多様性(母語訛り、地方訛り)に着眼する。一般的な英語教材において、世界諸英語を聞く教材は存在するが、世界諸英語の全体像を提示するには至っておらず、学習者自身を世界諸英語の中に位置付け、諸英語を俯瞰する教材はない。本研究では世界中の英語発音を対象としてそれらを分類し、学習者自身の英語を中心に世界諸英語発音を俯瞰できるブラウザを構築し、それを効果的に教育応用することを目的として計画された。

### 2. 研究の目的

本研究では当初、世界中の話者から英語音声を集めている海外の研究者と協力を図りつつ、1)話者を単位とした世界諸英語の発音分類技術の構築、2)インターネット上の英語コンテンツを利用した、世界諸英語ブラウザ(発音地図化)の構築、3)構築したシステムの教育利用、4)構築したシステムのビジネス利用、を研究目的とした。下記に示すように、まず、1)、2)について技術的実装を行なった。諸英語分類や、諸英語ブラウザは、ある話者の英語発音と、別の話者の英語発音の「発音距離」を推定することが基本的な技術的課題となる。これらの研究成果の発表を通して、発音距離を推定すること以上に、ある話者の発音は、ある聴取者にどの程度聞き取り易いのか、難しいのか、intelligibility に基づく話者間距離を求めることのほうが、より本質的な問題であることに気づいた。つまり音響的に観測される発音に基づく二話者距離ではなく、相手の発音が聞き取り易いのか難しいのか、という認知的距離の推定の方がより本質的な課題である。後者の方が対象としての難度が上がるため、a)日本人英語を(日本人と会話したことがない)米国人にどの程度聞き取ってもらえるのか、発音了解性の推定を検討することとした。更には、b)外国語訛り以外の要因(例えば外部雑音や無線音声などの音響歪み)によってどの程度聞き取り難くなるのか、について検討し、聴取の頑健性の向上を試みた。昨今、移民問題に伴う日本語教育支援の必要性が叫ばれているが、c)我々母語話者が聞き取り易い日本語発音を学ぶための韻律教材についても検討・開発した。

### 3. 研究の方法

#### 1)話者を単位とした世界諸英語の発音分類技術の構築

N 個の要素からなる集合を分類(クラスタリング)する場合、任意の二要素間の距離を求める必要がある(距離行列)。二話者間の発音距離を定義する場合、声と声の距離を求めてしまうと、男女差、年齢差の影響を直接的に受ける。そこで、同一文を多数の英語利用者によって読み上げさせた音声コーパス(Speech Accent Archive、全音声サンプルは IPA 音声記号による書き起こし付き)を用い、二話者間の IPA 書き起こし間距離を発音距離と定義し、この発音距離を二話者の音声信号のみから推定するタスクを設定した。性別、年齢などに起因する音響変動の抑制には音声の構造的表象を用い、これら音声特徴量から発音距離を推定する予測器としては、各種の回帰モデルを実験的に検討した。

#### 2)世界諸英語ブラウザ(発音地図化)の構築

多数の英語話者から、話者を単位とした発音距離行列を取得できれば、彼らを発音クラスタリングできるが、その結果を教育的に、かつ、効果的に可視化する必要がある。ここでは、ある特定の学習者を中心として(原点に配置し)、その学習者の周辺に、その他の母語話者、非母語話者の英語話者を配置することを考える。この場合、原点に配置された話者の性別と同性、異性、更には、学習者の年齢なども考慮し、効果的な可視化を検討した。

a)日本人の英語は(日本人と会ったことがない)米国人にどのように聞こえているのか  
二話者(あるいは、話者と聴取者)の距離を、両者の発音ではなく、一方の発音が他方にとってどのくらい聞き取り易いのか、を検討するため、日本人の英語は米語話者にどのように聞こえて入りうのかを推定する技術開発を行なった。ここでは先行研究で収集した、日本人の文読み上げ音声に対する米国人の聴取&復唱コーパス(聴取後即座に聞こえた通りに復唱させ、復唱音声を書き起こしたコーパス)を利用する。このコーパスを用いて、日本人英語音声の各単語に対して正しく復唱された率を計算し、それを各単語音声に対する了解性として定義した。さらに、その単語音声の分節的特徴、韻律的特徴、更には意図された単語の音素並びの特徴、語としての特徴(N-gram スコアなど)、統語解析結果から得られる特徴を用い、機械学習(回帰モデル)を用いて、了解性を推定するタスクを設定し、実験的に検討した。

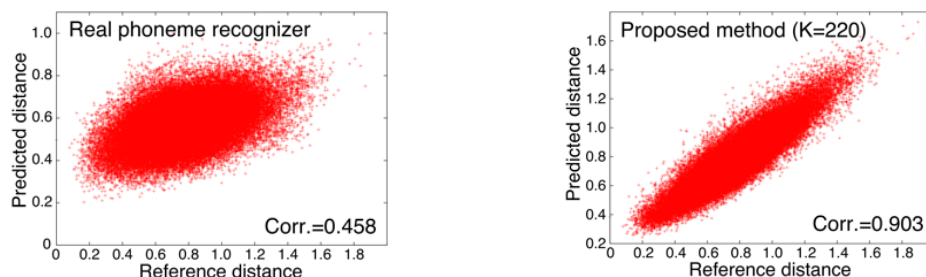
b)聴取能力の向上を目的とした音響歪みが付与された音声資料の積極的利用とその効果  
聞き取りは、話者の外国語訛りだけではなく、会話環境に起因する音響要因も、聞き取りを妨げる。母語話者が母語話者に語る場合に、聞き取りを妨げない音響歪みでも、学習者にとっては聞き取りが極めて難しい場合がある。ここでは、無線歪みがそれに該当することを示し、無線歪みに基づく聴取教材によって聞き取りの頑健性が効果的に改善するか否かを検討した。

c)日本語学習者が(母語話者が聞き取り易い)日本語発音を学ぶための韻律教材の開発・普及

学習対象言語を日本語とした場合、韻律（特にフレーズング）の制御を適切に導入すれば、初級者であっても聞き取り易い日本語となることが知られている。日本語音声合成の韻律予測モジュールを活用し、任意文の読み上げを想定し、それを聞き取り易く読み上げるための韻律制御を可視化し、その通りに読み上げた合成音声も提供する、韻律読み上げチュータを開発した。任意のテキストの読み上げに対して、必要となる韻律制御を個々の文に自動付与するインフラ提供は日本語教育史上初であり、本インフラの普及活動も尽力した。

#### 4. 研究成果

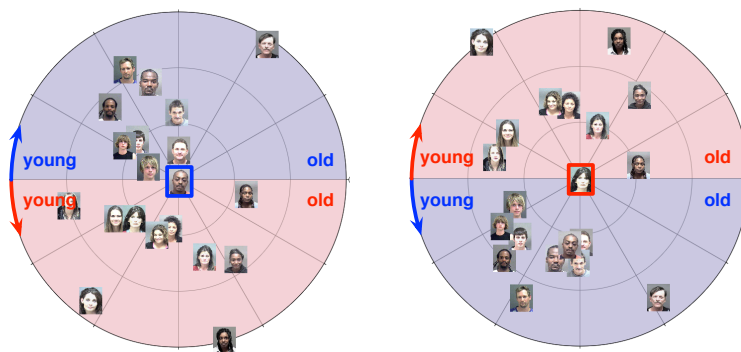
1) 話者を単位とした世界諸英語の発音分類(発音地図構築)技術の構築  
 同一テキストを読み上げた二話者間の発音距離を、二話者のスペクトログラムの距離として計算すれば、年齢や性別の違いに大きく依存する形で距離が計測される。年齢や性別による音響変動を抑えるために、音声の構造的表象に基づいて特徴抽出し、それに基づく多次元特徴を回帰モデル (SVR) へ入力し、IPA 書き起こし間距離を予測した。比較対象として、二話者の音声を連続音素認識し、認識結果の音素列間の距離でもって予測する手法と比較した。その結果、提案手法では IPA 書き起こし予測が相関 0.903 で実現できたのに対し、連続音素認識では、0.458 と提案手法が非常に高い精度で予測できることを実験的に示した。



連続音素認識結果による予測精度と、提案手法による予測精度

#### 2) 世界諸英語ブラウザーの構築

距離行列化された話者群を、その中の特定の一人を中心に他の話者の発音を俯瞰する可視化手法を検討した。提案した可視化を下記に示す。原点に特定話者を置き、それ以外の話者を半径 R の円の中に配置する。原点話者とそれ以外の話者の距離が発音距離を意味し、上半円は同性話者、下半円は異性話者を配置し、各話者の年齢を使って、配置場所の角度を調整した。評価実験の結果、自らを中心に配置しない従来の可視化手法に比べて高い教育的妥当性を得た。



提案した特定話者から世界諸英語を俯瞰するための可視化手法

a) 日本人の英語は（日本人と会ったことがない）米国人にどのように聞こえているのか  
 日本人英語音声を米国人に聴取&復唱させ、その復唱音声を書き取らせたコーパスを用い、日本人英語音声の各単語がどの程度聞き取ってもらえるのか、その率を、各種音声特徴、言語特徴を用いて回帰モデルを通して予測した。ここでは、10%未満を very unintelligible, 10-30%を rather unintelligible として定義し、前者の単語、後者の単語を同定するタスクとした。結果を下記に示す。適切なモデルを採択することで、7割から8割を正解することができた。

rather unintelligible, very unintelligible な単語の同定率

	very unintelligible word				rather unintelligible word			
	CART	AdaB	RF	ERT	CART	AdaB	RF	ERT
SET 1	65.44	66.80	67.38	<u>67.54</u>	70.45	<u>73.50</u>	72.90	73.13
SET 1+3	68.01	68.13	<u>69.22</u>	68.91	77.59	77.41	78.63	<u>78.94</u>
SET 1+2	71.48	71.21	71.97	<u>72.10</u>	83.21	83.97	<u>84.06</u>	83.89
SET 1+2+3	71.66	71.68	72.59	<u>72.74</u>	84.11	84.66	<u>84.78</u>	84.70

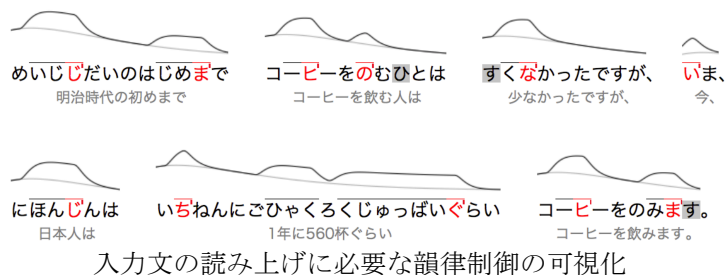
b) 聴取能力の向上を目的とした音響歪みが付与された音声資料の積極的利用とその効果  
 音声変形技術を用いることで、単一音源を巨人や小人の(ような)声、ホールでこだました声、  
 背景雑音が重畳した声、携帯電話を通じた声、航空無線のような歪んだ声、更には各歪みを定  
 量的に操作した上で組み合わせることも容易である。航空無線音声は学習者の聞き取りを大き  
 く妨げるが、母語話者は高い精度で聞き取れることが実験的に示された。航空無線音声を使っ  
 た「Robust Listening Drill」を構成し、その効果を検証した。下記は誤り削減率である。上  
 級者は無線音声のみならず、巨人・小人声や、無雑音音声での正解率も大きく改善した。

Pre test → Post test の誤り削減率

Larger ERR values than 40 are shown in bold.

Part	TOEIC	N	Orig.	GF	ATC	GF+ATC
A	400-600	15	9.9	35.6	2.3	-11.5
	600-800	32	-17.2	22.7	3.7	9.4
	800-990	8	<b>80.1</b>	<b>80.1</b>	<b>55.5</b>	23.9
B	400-600	15	33.4	8.8	13.9	0.0
	600-800	32	-11.3	6.2	5.1	6.8
	800-990	8	<b>42.9</b>	<b>58.4</b>	<b>40.0</b>	4.5

c) 日本語学習者が(母語話者が聞き取り易い)日本語発音を学ぶための韻律教材の開発・普及  
 任意の日本語文に対して、聞き取り易い音声で読み上げるための韻律読み上げチュータを開発  
 し、その効果を実験的に検討した。入力文は下記に示すような形で、その韻律は可視化され、  
 また、その通りに読み上げた合成音声で提示される。実験の結果、合成音声提示よりも、可視  
 化提示の方が(中国人日本語学習者には)効果的であることが示された。本システムは、日本  
 語教育史上初の、任意のテキストに対して共通語で読み上げるための韻律制御を可視化するシ  
 ステムとして教育現場に導入され、世界中で利用されるに至っている。2019年4月現在、40  
 カ国で140回以上の講習会を実施しており、日本語音声教育へ大きく貢献した。本システムの  
 開発に対して、2016年に電子情報通信学会より、ISS論文賞を受賞することができた。



入力文の読み上げに必要な韻律制御の可視化

## 5. 主な発表論文等

[雑誌論文] (査読プロセスを経て掲載された論文, 計 17 件)

- ① 峯松信明, 音声分析・合成・認識技術を用いた多様な 外国語教育支援, 日本音響学会誌, 査読有, 74, 9, 2018, pp.525-530  
DOI:10.20697/jasj.74.9\_525
- ② N. Minematsu, I. Nakamura, M. Suzuki, H. Hirano, C. Nakagawa, N. Nakamura, Y. Tagawa, K. Hirose, and H. Hashimoto, Development and evaluation of online infrastructure to aid teaching and learning of Japanese prosody, Trans. IEICE, 査読有, E100-D, 4, 2017, pp. 662-669  
DOI:10.1587/transinf.2016AWI0007
- ③ Y. Yamauchi, J. Yue, K. Ito & N. Minematsu, Investigation of teacher-selected sentences and machine-suggested sentences in terms of correlation between human ratings and GOP-based machine scores, Proc. SLaTE, 査読有, 2017  
[http://www.slate2017.org/papers/SLaTE\\_2017\\_paper\\_43.pdf](http://www.slate2017.org/papers/SLaTE_2017_paper_43.pdf)
- ④ 山内豊, 峯松信明, 川村明美, 西川恵, 加藤集平, 語彙サイズ, 認識速度, 処理の自動化を測定するオンラインL2語彙テストの開発, 外国語教育メディア学会, 査読有, 関東支部研究紀要, 1, 1, 2017, pp.1-24
- ⑤ F. Shiozawa, D. Saito, and N. Minematsu, Improved prediction of the accent gap between speakers of English for individual-based clustering of World Englishes, Proc. IEEE SLT, 査読有, 2016, pp.129-135  
DOI:10.1109/SLT.2016.7846255
- ⑥ N. Minematsu, H. Hirano, N. Nakamura, K. Oikawa, Improvement of naturalness of learners' spoken Japanese by practicing with the Web-based prosodic reading tutor, Suzuki-kun, Proc. ISCA Speech Prosody, 査読有, 2016, pp.252-256  
DOI:10.21437/SpeechProsody.2016-53
- ⑦ Y. Sato, Y. Kashiwagi, N. Minematsu, D. Saito, K. Hirose, Noise-robust and stress-free visualization of pronunciation diversity of World Englishes using a learner's

- self-centered viewpoint, Proc. OCOCODSA, 査読有, 2015, pp.1-5  
DOI:10.1109/ICSDA.2015.7357855
- ⑧ T. Pongkittiphan, N. Minematsu, T. Makino, D. Saito, K. Hirose, Automatic prediction of intelligibility of English words spoken with Japanese accents, Proc. ISCA SLaTE, 査読有, 2015, pp.19-22  
<https://www.slate2015.org/files/submissions/Pongkittiphan15-APO.pdf>
- ⑨ T. Shi, S. Kasahara, T. Pongkittiphan, N. Minematsu, D. Saito, K. Hirose, A measure of phonetic similarity to quantify pronunciation variations by using ASR technology, Proc. ICPHS, 査読有, 2015, CD-ROM  
<https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/Papers/ICPHS0432.pdf>
- ⑩ 峯松信明, 日本語音声・テキストコーパス情報処理に基づくオンライン韻律教育インフラの構築, 音声研究, 査読有, 19, 1, 2015, pp.18-31  
DOI:10.24467/onseikenkyu.19.1\_18
- ⑪ M. Suzuki, G. Kurata, M. Nishimura, N. Minematsu, Discriminative re-ranking for automatic recognition by leveraging invariant structures, Speech Communication, 査読有, 72, 2015, pp.208-217  
DOI:10.1016/j.specom.2015.06.007
- ⑫ 川村明美, 山内豊, Husky Kay, 英語シャドーイングの自動評価システムの総合的熟達度評価への応用に関する実証的研究, 東京国際大学言語コミュニケーション学部論叢, 査読有, 11, 2015, pp.125-135
- ⑬ Y. Kawase, N. Minematsu, D. Saito, K. Hirose, Visualization of pronunciation diversity of World Englishes from a speaker's self-centered viewpoint, Proc. O-COCOSDA, 査読有, 2014, pp.149-153  
DOI:10.1109/ICSDA.2014.7051437
- ⑭ T. Pongkittiphan, N. Minematsu, T. Makino, K. Hirose, Improvement of intelligibility prediction of spoken words in Japanese accented English using phonetic pronunciation distance and word confusability, Proc. O-COCOSDA, 査読有, 2014, pp.276-281
- ⑮ N. Minematsu, S. Kasahara, T. Makino, D. Saito, K. Hirose, Speaker-basis accent clustering using invariant structure analysis and the speech accent archive, Proc. Odyssey, 査読有, 2014, pp.158-165  
<https://www.superlectures.com/odyssey2014/speaker-basis-accent-clustering-using-invariant-structure-analysis-and-the-speech-accent-archive>
- ⑯ S. Kasahara, S. Kitahara, N. Minematsu, H.-P. Shen, T. Makino, D. Saito, K. Hirose, Improved and robust prediction of pronunciation distance for individual-basis clustering of World Englishes pronunciation, Proc. IEEE ICASSP, 査読有, 2014, pp.3240-3244  
DOI:10.1109/ICASSP.2014.6854194
- ⑰ S. Kasahara, N. Minematsu, H.-P. Shen, D. Saito, K. Hirose, Structure-based prediction of English pronunciation distances and its analytical investigation, Proc. IEEE ICIST, 査読有, 2014, pp.331-335  
DOI:10.1109/ICIST.2014.6920396
- [学会発表] (基調講演, 招待講演, チュートリアル講演を中心に示す。計 17 件)
- ① 牧野武彦, 英語辞書の発音表記から最近のアメリカ発音を知る, 日本音声学会 2018 年度音声学入門講座 (チュートリアル講義)
- ② T. Makino, Prosodic patterns in English Read by Japanese Phonetic Corpus: An interim report, 10th Pronunciation in Second Language Learning and Teaching conference, 2018
- ③ N. Minematsu, How speech technologies can support foreign language learning. Keynote speech, L2 Japanese Education with Technology, 2017
- ④ 峯松信明, 音声言語情報処理を用いた外国語学習支援, 基調講演, CASTEL/J, 2017
- ⑤ 峯松信明, 音声言語情報処理を用いた計算機援用学習支援, 基調講演, 日本英語音声学会 関東支部第 16 回研究大会, 2017
- ⑥ T. Makino, Intelligibility of Japanese-accented pronunciation of English: A phonetic analysis based on English Read by Japanese database, 9th Pronunciation in Second Language Learning and Teaching conference, 2017
- ⑦ N. Minematsu, Speaker-based clustering of world Englishes pronunciation, Invited talk at ASA/ASJ Joint meeting, 2016
- ⑧ N. Minematsu, OJAD: web-based prosodic reding tutor of Japanese, Keynote talk of ISAPh2016, 2016
- ⑨ N. Minematsu, Perceptual and structural analysis of pronunciation diversity of World Englishes, Invited talk of EJHIB2015, 2015
- ⑩ N. Minematsu, OJAD and its practical use for teaching/learning Japanese prosody,

- Invited talk at ASSTA Research Event Program, 2015
- ⑪ Y. Yamauchi & N. Minematsu, Development of Automatic Evaluation Systems of Reading Aloud and Shadowing Performances, Invited talk at ASSTA Workshop, 2015
  - ⑫ T. Makino, Transcription and Description of Interlanguage Prosody in Japanese Speakers' English, Accents 2015: 9th International Conference on Native and Non-native Accents of English, 2015
  - ⑬ 山内豊, 学習者の言語処理の自動化－英語総合力を伸ばす新しいシャドーイング指導－, 招待講演, 英語教育総合学会, 2015
  - ⑭ 山内豊, 言語処理の自動化を目指す音読・シャドーイング, 招待講演, 関西英語教育学, 2015
  - ⑮ N. Minematsu, Perceptual and structural analysis of pronunciation diversity of World Englishes, Keynote speech of O-COCOSDA, 2014
  - ⑯ N. Minematsu, How to survive the diversity of pronunciation of World Englishes? -- Experimental facts and technical solutions --, Keynote speech of the 19th conference of Pan-Pacific Association of Applied Linguistics, 2014
  - ⑰ N. Minematsu, Invariant structure and its application to speech processing, Plenary talk of International Conference on Information Science and Technology, 2014

[図書]

[産業財産権]

[その他]

## 6. 研究組織

### (1) 研究分担者

研究分担者氏名：牧野 武彦  
ローマ字氏名：(MAKINO, Takehiko)  
所属研究機関名：中央大学  
部局名：経済学部  
職名：教授  
研究者番号 (8桁)：00269482

研究分担者氏名：山内 豊  
ローマ字氏名：(YAMAUCHI, Yutaka)  
所属研究機関名：創価大学 (2018年より)  
部局名：教育学部  
職名：教授  
研究者番号 (8桁)：30306245

研究分担者氏名：齋藤 大輔  
ローマ字氏名：(SAITO, Daisuke)  
所属研究機関名：東京大学  
部局名：大学院工学系研究科  
職名：講師  
研究者番号 (8桁)：40615150

### (2) 研究協力者

※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。