

令和 3 年 5 月 25 日現在

機関番号：11301

研究種目：基盤研究(A) (一般)

研究期間：2017～2020

課題番号：17H00823

研究課題名(和文) 深層学習に基づくマルチモーダル対話型英会話学習システムの研究開発

研究課題名(英文) Research and development of multi-modal interactive English learning system based on deep learning

研究代表者

伊藤 彰則 (ITO, Akinori)

東北大学・工学研究科・教授

研究者番号：70232428

交付決定額(研究期間全体)：(直接経費) 22,500,000円

研究成果の概要(和文)：深層学習に基づく英会話学習システムのための要素技術を開発し、英会話の練習ができるCALLシステムを作成した。まず、深層学習に基づく音声・表情・ジェスチャに対するインターフェースの高度化・高精度化として、日本人英語音声を高精度に認識する技術を確立した。次に、深層学習に基づく英語発音評価・英会話シミュレーション技術を確立させるため、表情とジェスチャが英語の習熟度評価に与える影響を調べた。また、対話音声に対して高精度に発音評価を行う手法を確立した。最後に基盤要素技術を統合して音声対話型英会話学習システムを作成した。

研究成果の学術的意義や社会的意義

CALLシステムは手軽な英語学習の手段として使われているが、人間の教師に対して行うような柔軟な対話による英会話練習を実現したシステムは存在せず、人間の教師と同等の教育効果が得られる英会話学習システムの登場が望まれていた。

今回開発したシステムは、人間同士のような柔軟な対話を行うことができ、また事前にシナリオがない音声対話に対して高精度な発音評価ができるようになった。自由な対話をもとに学習ができる英会話学習システムはこれまで存在しないため、学術的に意義があると同時に、社会的にも価値が大きいと考えられる。

研究成果の概要(英文)：We developed technologies for an English conversation learning system based on deep learning and created a CALL system for practicing English conversation: (1) We established technology for recognizing English speech spoken by Japanese with high accuracy to improve the accuracy of interfaces for speech, facial expressions, and gestures based on deep learning. (2) To establish English pronunciation evaluation and English conversation simulation technology based on deep learning, we investigated the effects of facial expressions and gestures on English proficiency evaluation. In addition, we established a method to evaluate pronunciation with high accuracy for interactive speech. (3) We integrated the technologies to create a spoken dialogue English conversation learning system.

研究分野：音声言語情報処理

キーワード：外国語教育システム 音声対話システム 英語発音評価 英語韻律評価 雑談対話システム 音声感情認識

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、研究者個人に帰属します。

1. 研究開始当初の背景

ビジネス, 学術研究, 公的活動など, あらゆる分野での英語の重要性は論を待たない. 国の教育政策においても, 読み書きだけでなくコミュニケーションスキル教育に重点が置かれており, 従来の「読む」「書く」能力に加えて「聞く」「話す」能力がより重要視されている. 英会話練習の機会は多様化し, 英会話教室に通う従来の方法だけでなく, ビデオ会議を用いて海外の講師と安価に会話練習を行うオンライン英会話や, コンピュータを教師の代わりとして学習を行う **CALL (Computer Assisted Language Learning)** システムの利用などが広がってきている. 初期の **CALL** システムは主にマルチメディア教材を用いたりリスニング練習用のシステムであったが, 近年音声認識の技術を用いてスピーキング練習を行うための **CALL** システムが開発され, 実用化も行われている.

発音練習は重要であるが, 実際の言語運用能力を向上させるためには, 実際に英語を使ってみる練習を反復して行うことが欠かせない. そのためには, 単に発音を評価するだけでなく, 総合的に英語を運用しながら, 同時に発音・文法・語彙などを学習していくアプローチが必要である. 教師との **1 対 1** の会話を通じた学習法ではこのような効果が期待されるが, このような学習を可能とした **CALL** システムはいまだ実現されていない. そのため我々は, 音声認識・合成・対話技術を援用・発展させる形で学習者とシステムが実際に音声で対話を行う学習システムを実現するための技術を開発してきた.

一方, 深層学習を用いたパターン認識・理解・合成技術が近年飛躍的に発展し, 画像認識・音声認識・音声合成・自然言語処理などの分野で大規模なパラダイムシフトが起こっている. 音声認識・音声合成分野では, 従来用いられていた手法はほぼ深層学習を用いた方法に置き換わり, それに伴って性能が向上している. そこで, 音声対話型 **CALL** システムの分野においても, 深層学習を導入することで, より高精度かつ高度なシステムの実現が強く期待される.

2. 研究の目的

本研究課題では, これまで研究代表者らが培ってきた個別の要素技術を深層学習に基づき発展・融合させ, 従来よりも高度かつ高性能な音声対話型 **CALL** システムのプロトタイプを作成し, その有効性を大規模実験により実証することを目的とし, 課題全体を以下の具体的な **3** つの項目に分割し, それぞれについて研究・開発を行うことを当初の研究目的とした.

(1) 深層学習に基づく音声・表情・ジェスチャに対するインターフェースの高度化・高精度化

- 音響モデル・言語モデルの改良に基づく英語 **L2** 話者の音声認識性能の向上
- マルチモーダル深層学習に基づくジェスチャ・表情・感情認識

(2) 深層学習に基づく英語発音評価・英会話シミュレーション技術の確立

- 深層学習による自動発音評価の精度向上
- マルチモーダル情報を利用した対話シミュレーションのための応答生成

(3) 基盤要素技術の統合

- エージェントを利用した対話型マルチモーダル **CALL** システムの開発

3. 研究の方法

初年度はまず大項目(1)のシステムとの英会話に必要なインターフェースの高精度化・高度化のためのデータベースの準備を行い、深層学習に基づく理論の構築、実装および評価実験を行った。2年目以降は(1)に引き続き取り組むとともに、学習者に習熟度についての適切なフィードバックと応答生成を行うため大項目(2)の発音評価および英会話シミュレーション技術について研究を進めた。これらの基盤要素技術を確立した後、3, 4年目には大項目(3)にある対話型マルチモーダル CALL システムを開発した。

4. 研究成果

上記の研究項目に分けて、研究成果について述べる。

(1) 深層学習に基づく音声・表情・ジェスチャに対するインターフェースの高度化・高精度化

日本語話者の英語音声を高精度に認識するための音響モデルの検討

英会話のための音声対話システムを作るためには、日本語母語話者の英語音声を高精度に認識する音声認識システムが必要不可欠である。従来の音声認識システムでは、大語彙連続音声認識の音響モデルとして、**GMM(Gaussian Mixture Model)**と**HMM(Hidden Markov Model)**を組み合わせた**GMM-HMM**モデルを採用していた。一方、近年ディープラーニングの音声認識への貢献が大きいため、音響モデルを**GMM**から**DNN(Deep Neural Network)**に置き換えることで、**GMM-HMM**の結果と比較して認識精度を向上させることができる。そこで、非ネイティブ音声認識タスクにおいて、**DNN-HMM**に基づいた音響モデルを開発し、その性能を**GMM-HMM**の性能と比較した。認識実験の結果、従来の**GMM-HMM**を用いた認識手法と比較して、大幅な認識精度(単語誤り率: **WER**, 文誤り率: **SER**)の向上が得られた。

クラウド型汎用音声認識 API とドメイン依存音声認識器の出力統合による認識精度向上

近年の音声認識は、大規模な音声・言語資源を用いて構築された高精度な音声認識器を利用するのが普通である。**Google** や **IBM**, **Amazon** などの企業はクラウドサービスによって音声認識 API を提供しており、現在ではだれでも容易に汎用的な音声認識器を使えるようになっている。一方で、今回対象にしているような特定の話題に関する音声認識では、ドメイン依存の音声認識器により高い性能が得られる。しかしながら、特定のドメインにおける言語・音声資源を大量に収集することは現実的ではなく、高精度なドメイン依存の音声認識器を構築することは容易ではない。そこで本研究では、複数のクラウド型汎用音声認識 API と、少量のデータから構築されたドメイン依存の音声認識システムを組み合わせることで、対象ドメインの音声認識精度を向上させる手法を検討した。

我々が提案する音声認識システムは特定ドメインを対象とするドメイン依存の音声認識器と複数のオープンドメイン音声認識器、及び **BLSTM** による単語選択器から構成される。ドメイン依存の音声認識器は対象ドメインで収集した言語・音声資源を用いて学習される。また、オープンドメイン音声認識器としてはクラウドサービスによって提供される汎用音声認識 API を利用した。認識実験の結果から、特定ドメインの評価データにおいて提案手法が最もよい精度を示した。

マルチモーダル情報を利用した英会話能力の評価

現状のほとんどの CALL システムは、言語情報と音声情報のみを対象としている。しかし実際の英会話においては、英語の運用能力だけでなく、会話に伴う表情やジェスチャな

どを含めた総合的なコミュニケーション能力が求められている。しかしながら、表情やジェスチャなどの非言語的情報が英会話の流暢さの評価にどの程度影響するかは明らかではない。そこで、マルチモーダル対話データを収集し、会話の流暢さ及び自然性と表情、ジェスチャとの関連性を分析し、機械学習による表情・ジェスチャの自然性の自動推定手法の検討を行い、マルチモーダル情報を利用した英会話能力への自動推定の検討を行い、その効果を検証した。

そのため、まずマルチモーダル英会話データベースの作成を行った。次に、集めたデータを英会話能力の自動評価に利用できるように、マルチモーダルデータベースを定められた基準に応じて評価した。その結果から、特に視線や頭部方向の上下の動きの平均値が主観評価と高い相関を持つことがわかった。

大規模感情音声データベースを用いた対話音声の感情認識

英会話のための対話システムの自然性を向上するためには、単に音声で対話をするだけでなく、マルチモーダル情報を利用して感情に合わせたリアクションを行うことが重要である。そこで、大規模感情データベース **JTES** および感情を含む対話音声を用いた音声感情認識の実験を行った。認識結果から、対話音声での音声感情認識精度が低いことがわかった。これは、読み上げと対話という発話スタイルの違いと、収録機材などの環境の違いの2つに由来すると考えられる。次に、音声の中の感情知覚がどのように行われるかを調べるため、**JTES** 音声を実験参加者が聞いて評価する再評価を行った。3名のラベラーが **JTES** の全発話に対して主観評価を行った結果、話者の意図する感情と聞き手のラベルは **70%**程度しか一致しなかった。また、聞き手が感じる「感情の強度」と、自動推定による感情推定の再現率の関係を調べたところ、感情強度が高いほど再現率も高い傾向がみられた。

(2) 深層学習に基づく英語発音評価・英会話シミュレーション技術の確立

深層学習による自動発音評価の精度向上

対話音声を対象とした英語音声の発音の自動評価手法を開発した。従来の音声の発音評価では、対象となる音声の発話内容が既知であることが前提となっていた。しかし、英会話練習においては、必ずしも事前に設定された内容を発話するとは限らない。提案手法は、英語ネイティブ音声から学習した音響モデルと、日本人英語音声から学習した音響モデルの2つを使って音声認識結果を2つ出力し、その違いの情報を使って発音の良さを評価する。この方法を使えば、従来のように読み上げ文を最初に用意して読ませるシステムだけでなく、事前のシナリオがない対話音声に対しても高い精度で発音評価が可能になる。

読み上げ音声である **ERJ**、シナリオを事前に用意して暗記した対話を行う **TMNED**、および完全に自由な対話である **TEC** の3つのコーパスに対して提案法による発音評価実験を行った。3つのコーパスに対して推定した発音評価値と主観評価スコアとの相関係数は、それぞれ **0.83**, **0.80**, **0.80**であった。

深層学習によるコーパス生成に基づく発話からの文法誤り検出

対話においては、学習者は多くの文法誤りを含んだ発話をする。それらの発話に含まれる文法的・語彙的誤りを指摘することを目標とする。これはキーボード入力であれば容易であるが、音声認識によって文法的・語彙的に誤りのある発話をそのまま認識することは難しい。そこで本研究では、日本人学習者による文法誤りを含んだ音声を「文法誤りを含んだ状態のまま」認識することを目標とした。そのため、従来は誤りのない英語からルールによって誤りを付加した文を大量に生成し、そこから言語モデルを学習していた。提案法では、誤りのない英語文を

日本人の犯しがちな誤りを含む英語文に変換するニューラルネットワークを学習し、これによって言語モデルを学習した。この結果から、機械学習を使って誤り文を生成して言語モデルを学習することには効果があること、ルールによって生成した文をさらに混ぜても精度が向上しないことなどが示された。

文の発音難易度の推定

本研究では評価に影響を与える要因の中で文の発音難易度に注目し、発音難易度がスコアに与える影響、また文の発音難易度を表す特徴量の検討を行った。特に本研究では、文の発音難易度を表す特徴量として、文章の調音的特性を表す特徴量と、文を読み上げる際の話者の心理的特性を考慮した特徴量の検討を行った。

まず、ある話者がある文を読んだ時の主観評価値をもとに、それを話者スコアと文の発音難易度のスコアの和で近似する手法を開発し、これを用いて文の発音難易度の目標値を算出した。次に、文から発音難易度を自動推定するため、特徴量の検討を行った。その結果、文の長さ、音素の種類、単語親密度、調音特徴量、**Surprisal**などの特徴量が有効であることが分かった。

提案する特徴量を重回帰分析することで文の発音難易度を推定した。相関係数は **0.48** となり、文の特徴のみから文の発音しやすさを推定することが初めて可能になった。

深層学習に基づく英語韻律の自動評価

音声認識技術を用いた CALL システムに関しては発音評価と韻律評価研究が行われている。特に発音評価においては多くの研究がなされている一方、発話の韻律も英語のコミュニケーションにおいて重要な役割を果たすが、発音評価に比べると韻律評価に関する研究は十分になされてこなかった。本研究では、鈴木らが提案した単語重要度係数の考え方に基づき、深層学習を導入した韻律評価手法を検討する。提案手法は Attention-based bidirectional LSTM をモデルとし、入力特徴量として言語情報と韻律情報を用いる。その結果、従来法である単語重要度係数を用いた評価と比較して、より高い相関を得ることができた。

対話シミュレーションのための親しみ感を考慮した応答生成

英会話学習に音声対話システムを用いる場合、対話システムを継続的に使用してもらるように、より質の高い対話システムを提供することも重要な要因となる。本研究では、ユーザと自然な対話を行うシステムの質を高めるための対話システムの構築を目的とし、段階的に話し方を変化させることで、ユーザとシステムの親和性向上を促す手法を検討した。

(3)基盤要素技術の統合

任意の音声の発音評価ができるシステムを利用し、音声対話型 **CALL** システムを開発した。プラットフォームとして **Python** ベースの **Flask** を利用し、利用者の音声入力をもとに、シナリオベースの英会話を行うシステムを開発した。学習者の発話に応じてシナリオに合った返事をするほか、これまで開発した発音評価システムを利用して、発音の良さをフィードバックする機能を持っている。

以上、深層学習に基づく技術を総合して英会話学習を行うシステムの開発を行った。世界的に見ても、実際に自由な英会話を行いながら英語学習ができるシステムは存在しない。そのため、本研究の成果は、世界的に見ても大きいインパクトがあると考えられる。

5. 主な発表論文等

〔雑誌論文〕 計24件（うち査読付論文 24件 / うち国際共著 0件 / うちオープンアクセス 6件）

1. 著者名 Jiang Fu, Yuya Chiba, Takashi Nose, Akinori Ito	4. 巻 110
2. 論文標題 Evaluation of English Speech Recognition for Japanese Learners Using DNN-Based Acoustic Models	5. 発行年 2018年
3. 雑誌名 Recent Advances in Intelligent Information Hiding and Multimedia Signal Processing, IH-MSP 2018. Smart Innovation, Systems and Technologies	6. 最初と最後の頁 93-100
掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/978-3-030-03748-2_11	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 -
1. 著者名 Mai Yamanaka, Yuya Chiba, Takashi Nose, Akinori Ito	4. 巻 110
2. 論文標題 A Study on a Spoken Dialogue System with Cooperative Emotional Speech Synthesis Using Acoustic and Linguistic Information	5. 発行年 2018年
3. 雑誌名 Recent Advances in Intelligent Information Hiding and Multimedia Signal Processing, IH-MSP 2018. Smart Innovation, Systems and Technologies	6. 最初と最後の頁 101-108
掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/978-3-030-03748-2_12	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 -
1. 著者名 Takashi Kimura, Takashi Nose, Shinji Hirooka, Yuya Chiba, Akinori Ito	4. 巻 110
2. 論文標題 Comparison of Speech Recognition Performance Between Kaldi and Google Cloud Speech API	5. 発行年 2018年
3. 雑誌名 Recent Advances in Intelligent Information Hiding and Multimedia Signal Processing, IH-MSP 2018. Smart Innovation, Systems and Technologies	6. 最初と最後の頁 109-115
掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/978-3-030-03748-2_13	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 -
1. 著者名 Shinya Hanabusa, Takashi Nose, Akinori Ito	4. 巻 110
2. 論文標題 Segmental Pitch Control Using Speech Input Based on Differential Contexts and Features for Customizable Neural Speech Synthesis	5. 発行年 2018年
3. 雑誌名 Recent Advances in Intelligent Information Hiding and Multimedia Signal Processing, IH-MSP 2018. Smart Innovation, Systems and Technologies	6. 最初と最後の頁 124-131
掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/978-3-030-03748-2_15	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 -

1. 著者名 Sou Miyamoto, Takashi Nose, Kazuyuki Hiroshiba, Yuri Odagiri, Akinori Ito	4. 巻 110
2. 論文標題 Two-Stage Sequence-to-Sequence Neural Voice Conversion with Low-to-High Definition Spectrogram Mapping	5. 発行年 2018年
3. 雑誌名 Recent Advances in Intelligent Information Hiding and Multimedia Signal Processing, IH-MSP 2018. Smart Innovation, Systems and Technologies	6. 最初と最後の頁 132-139
掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/978-3-030-03748-2_16	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 -

1. 著者名 Toru Ishikawa, Takashi Nose, Akinori Ito	4. 巻 110
2. 論文標題 DNN-Based Talking Movie Generation with Face Direction Consideration	5. 発行年 2018年
3. 雑誌名 Recent Advances in Intelligent Information Hiding and Multimedia Signal Processing, IH-MSP 2018. Smart Innovation, Systems and Technologies	6. 最初と最後の頁 157-164
掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/978-3-030-03748-2_19	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 -

1. 著者名 Akinori Ito	4. 巻 110
2. 論文標題 Muting Machine Speech Using Audio Watermarking	5. 発行年 2018年
3. 雑誌名 Recent Advances in Intelligent Information Hiding and Multimedia Signal Processing, IH-MSP 2018. Smart Innovation, Systems and Technologies	6. 最初と最後の頁 74-81
掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/978-3-030-03748-2_9	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 -

1. 著者名 Akinori Ito	4. 巻 110
2. 論文標題 Leveraging a Small Corpus by Different Frame Shifts for Training of a Speech Recognizer	5. 発行年 2018年
3. 雑誌名 Recent Advances in Intelligent Information Hiding and Multimedia Signal Processing, IH-MSP 2018. Smart Innovation, Systems and Technologies	6. 最初と最後の頁 82-89
掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/978-3-030-03748-2_10	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 -

1. 著者名 Shunsuke Tada, Yuya Chiba, Takashi Nose, Akinori Ito	4. 巻 -
2. 論文標題 Effect of Mutual Self-Disclosure in Spoken Dialog System on User Impression	5. 発行年 2018年
3. 雑誌名 Proceedings of APSIPA Annual Summit and Conference	6. 最初と最後の頁 806-810
掲載論文のDOI (デジタルオブジェクト識別子) 10.23919/APSIPA.2018.8659630	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 -

1. 著者名 Haoran Wu, Yuya Chiba, Takashi Nose, Akinori Ito	4. 巻 -
2. 論文標題 Analyzing Effect of Physical Expression on English Proficiency for Multimodal Computer-Assisted Language Learning	5. 発行年 2018年
3. 雑誌名 Proceedings of Interspeech	6. 最初と最後の頁 1746-1750
掲載論文のDOI (デジタルオブジェクト識別子) 10.21437/Interspeech.2018-1425	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 -

1. 著者名 Yuya Chiba, Takashi Nose, Taketo Kase, Mai Yamanaka, Akinori Ito	4. 巻 -
2. 論文標題 An Analysis of the Effect of Emotional Speech Synthesis on Non-Task-Oriented Dialogue System	5. 発行年 2018年
3. 雑誌名 Proceedings of the 19th Annual SIGdial Meeting on Discourse and Dialogue	6. 最初と最後の頁 371-375
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 -

1. 著者名 Yukiko Kageyama, Yuya Chiba, Takashi Nose, Akinori Ito	4. 巻 -
2. 論文標題 Improving User Impression in Spoken Dialog System with Gradual Speech Form Control	5. 発行年 2018年
3. 雑誌名 Proceedings of the 19th Annual SIGdial Meeting on Discourse and Dialogue	6. 最初と最後の頁 235-240
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 -

1. 著者名 Chiba Yuya, Nose Takashi, Ito Akinori	4. 巻 40
2. 論文標題 Multi-condition training for noise-robust speech emotion recognition	5. 発行年 2019年
3. 雑誌名 Acoustical Science and Technology	6. 最初と最後の頁 406 ~ 409
掲載論文のDOI (デジタルオブジェクト識別子) 10.1250/ast.40.406	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 -

1. 著者名 Yusuke Yamada, Yuya Chiba, Takashi Nose, Akinori Ito	4. 巻 11
2. 論文標題 Effect of Training Data Selection for Speech Recognition of Emotional Speech	5. 発行年 2021年
3. 雑誌名 International Journal of Machine Learning and Computing	6. 最初と最後の頁 362-366
掲載論文のDOI (デジタルオブジェクト識別子) 10.18178/ijmlc.2021.11.5.1062	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 -

1. 著者名 Fu Jiang, Chiba Yuya, Nose Takashi, Ito Akinori	4. 巻 41
2. 論文標題 Language modeling in speech recognition for grammatical error detection based on neural machine translation	5. 発行年 2020年
3. 雑誌名 Acoustical Science and Technology	6. 最初と最後の頁 788 ~ 791
掲載論文のDOI (デジタルオブジェクト識別子) 10.1250/ast.41.788	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 -

1. 著者名 Fu Jiang, Chiba Yuya, Nose Takashi, Ito Akinori	4. 巻 116
2. 論文標題 Automatic assessment of English proficiency for Japanese learners without reference sentences based on deep neural network acoustic models	5. 発行年 2020年
3. 雑誌名 Speech Communication	6. 最初と最後の頁 86 ~ 97
掲載論文のDOI (デジタルオブジェクト識別子) 10.1016/j.specom.2019.12.002	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 -

1. 著者名 Kanagaki Aoi, Tanaka Masaya, Nose Takashi, Shimizu Ryohei, Ito Akira, Ito Akinori	4. 巻 -
2. 論文標題 CycleGAN-Based High-Quality Non-Parallel Voice Conversion with Spectrogram and WaveRNN	5. 発行年 2020年
3. 雑誌名 Proceedings of 9th IEEE Global Conference on Consumer Electronics	6. 最初と最後の頁 456-457
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/GCCE50665.2020.9291952	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 -

1. 著者名 Hayasaka Takuma, Nose Takashi, Ito Akinori	4. 巻 -
2. 論文標題 A Study on Minimum Spectral Error Analysis of Speech	5. 発行年 2020年
3. 雑誌名 Proceedings of 9th IEEE Global Conference on Consumer Electronics	6. 最初と最後の頁 462-463
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/GCCE50665.2020.9291840	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 -

1. 著者名 Yahagi Ryota, Chiba Yuya, Nose Takashi, Ito Akinori	4. 巻 -
2. 論文標題 Incremental Response Generation Using Prefix-to-Prefix Model for Dialogue System	5. 発行年 2020年
3. 雑誌名 Proceedings of 9th IEEE Global Conference on Consumer Electronics	6. 最初と最後の頁 482-483
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/GCCE50665.2020.9291883	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 -

1. 著者名 Naijo Satsuki, Chiba Yuya, Nose Takashi, Ito Akinori	4. 巻 -
2. 論文標題 Analysis and Estimation of Sentence Speakability for English Pronunciation Evaluation	5. 発行年 2020年
3. 雑誌名 Proceedings of 9th IEEE Global Conference on Consumer Electronics	6. 最初と最後の頁 449-451
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/GCCE50665.2020.9292072	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 -

1. 著者名 Fujimaki Daisuke, Nose Takashi, Ito Akinori	4. 巻 -
2. 論文標題 Integration of Accent Sandhi and Prosodic Features Estimation for Japanese Text-to-Speech Synthesis	5. 発行年 2020年
3. 雑誌名 Proceedings of 9th IEEE Global Conference on Consumer Electronics	6. 最初と最後の頁 484-485
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/GCCE50665.2020.9291906	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 -

1. 著者名 Yamazaki Yoshihiro, Chiba Yuya, Nose Takashi, Ito Akinori	4. 巻 -
2. 論文標題 Filler Prediction Based on Bidirectional LSTM for Generation of Natural Response of Spoken Dialog	5. 発行年 2020年
3. 雑誌名 Proceedings of 9th IEEE Global Conference on Consumer Electronics	6. 最初と最後の頁 486-487
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/GCCE50665.2020.9291867	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 -

1. 著者名 Shota Ueda, Yuya Chiba, Takashi Nose, Akinori Ito	4. 巻 -
2. 論文標題 A Study on Automatic English Prosody Evaluation Using Attention-based Bidirectional LSTM	5. 発行年 2020年
3. 雑誌名 Proceedings of 2020 RISP International Workshop on Nonlinear Circuits, Communications and Signal Processing	6. 最初と最後の頁 295-298
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 -

1. 著者名 Mai Yamanaka, Takashi Nose, Yuya Chiba, Akinori Ito	4. 巻 -
2. 論文標題 Labeling and Analysis of Perceived Emotion for Japanese Large-Scale Emotional Speech Database JTES	5. 発行年 2020年
3. 雑誌名 Proceedings of 2020 RISP International Workshop on Nonlinear Circuits, Communications and Signal Processing	6. 最初と最後の頁 230-233
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 -

〔学会発表〕 計37件（うち招待講演 1件 / うち国際学会 4件）

1. 発表者名 Haoran Wu, Yuya Chiba, Takashi Nose, Akinori Ito
2. 発表標題 Analyzing Effect of Multimodal Information Processing on English Proficiency
3. 学会等名 International Workshop on Emerging ICT (国際学会)
4. 発表年 2019年

1. 発表者名 Jiang Fu, Yuya Chiba, Takashi Nose, Akinori Ito
2. 発表標題 Automatic English Proficiency Assessment for Japanese Learners without Reference Transcriptions
3. 学会等名 日本音響学会2019年春季研究発表会
4. 発表年 2019年

1. 発表者名 山中麻衣, 千葉祐弥, 能勢隆, 伊藤彰則
2. 発表標題 話者・環境適応と波形正規化を用いた音声感情認識の精度改善
3. 学会等名 日本音響学会2019年春季研究発表会
4. 発表年 2019年

1. 発表者名 多田駿介, 千葉祐弥, 能勢隆, 伊藤彰則
2. 発表標題 相互自己開示に基づく対話システムにおける傾聴的応答生成の効果の検証
3. 学会等名 日本音響学会2019年春季研究発表会
4. 発表年 2019年

1. 発表者名 欧陽江丹, 能勢隆, 伊藤彰則
2. 発表標題 WaveNet ボコーダにおける few-shot 話者適応の検討
3. 学会等名 日本音響学会2019年春季研究発表会
4. 発表年 2019年

1. 発表者名 呉浩然, 千葉祐弥, 能勢隆, 伊藤彰則
2. 発表標題 日本人英会話における自然性自動評価のためのマルチモーダル情報の分析
3. 学会等名 日本音響学会2019年春季研究発表会
4. 発表年 2019年

1. 発表者名 山中麻衣, 能勢隆, 千葉祐弥, 伊藤彰則
2. 発表標題 感情音声データベース JTES の主観ラベル付与に向けた予備的検討
3. 学会等名 日本音響学会2019年春季研究発表会
4. 発表年 2019年

1. 発表者名 花房伸哉, 能勢隆, 伊藤彰則
2. 発表標題 音声とGUIによりピッチ制御が可能なDNN音声合成システムの検討
3. 学会等名 日本音響学会2019年春季研究発表会
4. 発表年 2019年

1. 発表者名 木村崇, 能勢隆, 廣岡伸治, 千葉祐弥, 伊藤彰則
2. 発表標題 日本語音声を対象としたKaldiとGoogle Cloud Speech APIの認識性能比較
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 山中麻衣, 千葉祐弥, 能勢隆, 伊藤彰則
2. 発表標題 音響情報と言語情報を用いた協調的発話感情付与に基づく音声対話システムの検討
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 多田駿介, 千葉祐弥, 能勢隆, 伊藤彰則
2. 発表標題 相互自己開示によりユーザの印象を向上させる音声対話システムの構築と評価
3. 学会等名 第123回SLP研究会
4. 発表年 2018年

1. 発表者名 蔭山友紀子, 千葉祐弥, 能勢隆, 伊藤彰則
2. 発表標題 対話エージェントの段階的口調変化の有無による印象の比較と検討
3. 学会等名 日本音響学会2018年春季研究発表会
4. 発表年 2018年

1. 発表者名 多田駿介, 千葉祐弥, 能勢隆, 伊藤彰則
2. 発表標題 相互自己開示によりユーザの印象を向上させる対話システムの検討
3. 学会等名 日本音響学会2018年春季研究発表会
4. 発表年 2018年

1. 発表者名 呉浩然, 千葉祐弥, 能勢隆, 伊藤彰則
2. 発表標題 英会話学習システムのためのマルチモーダル対話データベースの収集、評価と分析
3. 学会等名 日本音響学会2018年春季研究発表会
4. 発表年 2018年

1. 発表者名 Yukiko Kageyama, Yuya Chiba, Takashi Nose, Akinori Ito
2. 発表標題 Changing Impression to the Spoken Dialog System by Gradual Control of Expressions
3. 学会等名 International Workshop on Emerging ICT (国際学会)
4. 発表年 2017年

1. 発表者名 Akinori Ito
2. 発表標題 Human-Machine Meta-Communication
3. 学会等名 International Conference on Intelligent Information Technology (招待講演) (国際学会)
4. 発表年 2018年

1. 発表者名 Yoshihiro Yamazaki, Yuya Chiba, Takashi Nose, Akinori Ito
2. 発表標題 Construction of Multimodal Chat-talk Corpus Considering the Closeness in Dyads and Analysis of Dialog Act
3. 学会等名 International Workshop on Emerging ICT (国際学会)
4. 発表年 2019年

1. 発表者名 山中麻衣, 能勢 隆, 千葉祐弥, 伊藤彰則
2. 発表標題 クラウドソーシングを用いた感情音声データベースJTESへの主観ラベルの付与と分析
3. 学会等名 第21回音声言語シンポジウム・音声研究会
4. 発表年 2019年

1. 発表者名 山崎善啓, 千葉祐弥, 能勢隆, 伊藤彰則
2. 発表標題 親密性を考慮した対話システムのためのマルチモーダル雑談コーパスの構築と分析
3. 学会等名 第87回言語・音声理解と対話処理研究会 (第10回対話システムシンポジウム)
4. 発表年 2019年

1. 発表者名 藤巻大輔, 能勢隆, 伊藤彰則
2. 発表標題 日本語テキスト音声合成におけるアクセント結合・韻律特徴量推定の統合に関する検討
3. 学会等名 第377回音響工学研究会
4. 発表年 2019年

1. 発表者名 早坂琢真, 能勢隆, 伊藤彰則
2. 発表標題 高品質ボコーダのためのスペクトル誤差最小分析法の提案
3. 学会等名 第377回音響工学研究会
4. 発表年 2019年

1. 発表者名 山崎善啓, 千葉祐弥, 能勢隆, 伊藤彰則
2. 発表標題 親密性を考慮したマルチモーダル雑談コーパスにおける発話意図及び韻律・表情に関する分析
3. 学会等名 第377回音響工学研究会
4. 発表年 2019年

1. 発表者名 山崎善啓, 千葉祐弥, 能勢隆, 伊藤彰則
2. 発表標題 話者間の親密性を考慮したマルチモーダル雑談コーパスの構築と分析
3. 学会等名 日本音響学会2019年秋季研究発表会
4. 発表年 2019年

1. 発表者名 付江, 千葉祐弥, 能勢隆, 伊藤彰則
2. 発表標題 Automatic Generation of Text with Errors using Neural Machine Translation for Grammatical Error Detection
3. 学会等名 日本音響学会2019年秋季研究発表会
4. 発表年 2019年

1. 発表者名 植田翔太, 千葉祐弥, 能勢隆, 伊藤彰則
2. 発表標題 Attention-based LSTMを用いた英語韻律の自動評価手法の提案
3. 学会等名 電気関係学会東北支部連合大会
4. 発表年 2019年

1. 発表者名 付江, 千葉祐弥, 能勢隆, 伊藤彰則
2. 発表標題 Automatic English Proficiency Assessment for Japanese Learners without Reference Transcriptions
3. 学会等名 日本音響学会2019年春季研究発表会
4. 発表年 2019年

1. 発表者名 Satsuki Naijo, Akinori Ito, Takashi Nose
2. 発表標題 Extraction and Estimation of Sentence Speakability for English Pronunciation Evaluation
3. 学会等名 International Workshop on Emerging ICT
4. 発表年 2020年

1. 発表者名 早坂琢真, 能勢隆, 伊藤彰則
2. 発表標題 時間・周波数連続性を考慮した畳み込みニューラルネットに基づく位相復元の検討
3. 学会等名 令和3年東北地区若手研究者研究発表会
4. 発表年 2021年

1. 発表者名 山崎善啓, 千葉祐弥, 能勢隆, 伊藤彰則
2. 発表標題 言語・韻律情報の同時モデル化に基づく音声応答生成の検討
3. 学会等名 第90回言語・音声理解と対話処理研究会 (第11回対話システムシンポジウム)
4. 発表年 2020年

1. 発表者名 内城沙月, 伊藤彰則, 能勢隆
2. 発表標題 英語発音評価における文の発音難易度推定のための特徴量の検討
3. 学会等名 日本音響学会2020年秋季研究発表会
4. 発表年 2020年

1. 発表者名 山崎善啓, 千葉祐弥, 能勢隆, 伊藤彰則
2. 発表標題 SMOCコーパス: 対話システムのための大規模マルチモーダル対面雑談コーパス
3. 学会等名 日本音響学会2020年秋季研究発表会
4. 発表年 2020年

1. 発表者名 山崎善啓, 千葉祐弥, 能勢隆, 伊藤彰則
2. 発表標題 雑談コーパスを用いた双方向LSTMに基づくフィラー予測の検討
3. 学会等名 日本音響学会2020年秋季研究発表会
4. 発表年 2020年

1. 発表者名 藤巻大輔, 能勢隆, 伊藤彰則
2. 発表標題 日本語テキスト音声合成のためのアクセント情報を考慮した単語埋め込み手法の検討
3. 学会等名 日本音響学会2020年秋季研究発表会
4. 発表年 2020年

1. 発表者名 植田 翔太, 千葉 祐弥, 能勢 隆, 伊藤 彰則
2. 発表標題 転移学習とAttention-based LSTMに基づく英語自動韻律評価
3. 学会等名 日本音響学会2020年春季研究発表会
4. 発表年 2020年

1. 発表者名 木村 崇, 千葉 祐弥, 能勢 隆, 伊藤 彰則
2. 発表標題 クラウド型汎用音声認識APIとドメイン依存音声認識器の出力統合による認識精度向上
3. 学会等名 日本音響学会2020年春季研究発表会
4. 発表年 2020年

1. 発表者名 山中 麻衣, 能勢 隆, 千葉 祐弥, 伊藤 彰則
2. 発表標題 大規模感情音声データベースJTESへの聞き手ラベルの付与と音声感情認識での利用
3. 学会等名 日本音響学会2020年春季研究発表会
4. 発表年 2020年

1. 発表者名 内城沙月, 千葉祐弥, 能勢隆, 伊藤彰則
2. 発表標題 日本人英語評価における話者の発音能力と文の読みやすさの関係性の分析
3. 学会等名 日本音響学会2020年春季研究発表会
4. 発表年 2020年

〔図書〕 計0件

〔産業財産権〕

〔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	能勢 隆 (NOSE Takashi) (90550591)	東北大学・工学研究科・准教授 (11301)	
研究分担者	千葉 祐弥 (CHIBA Yuya) (30780936)	東北大学・工学研究科・助教 (11301)	2017～2019

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 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
---------	---------