

令和 6 年 6 月 25 日現在

機関番号：34453

研究種目：基盤研究(B)（一般）

研究期間：2020～2022

課題番号：20H01716

研究課題名（和文）教育現場における健聴者・聴覚障害者間の円滑な意思疎通を支援する音声言語基盤技術

研究課題名（英文）Infrastructural spoken language technology to support smooth communication with hearing-impaired people in education

研究代表者

小林 彰夫（Kobayashi, Akio）

大和大学・情報学部・教授

研究者番号：10741168

交付決定額（研究期間全体）：（直接経費） 13,900,000円

研究成果の概要（和文）：発音の不明瞭なろう・難聴者（聴覚障害者）にとって、音声認識を備えた音声インターフェースは認識性能の低さから使いにくいものであった。そこでこの問題を解決するために、本研究では50名のろう・難聴者の音声を収集したコーパスを整備し、読み上げ文や自由発話（会話）を多数収集した。また、整備したコーパスを用いてろう・難聴者の音声認識性能の改善を行った。その結果、ろう・難聴の読み上げ文の発話に対して、聴者には及ばないものの有意な認識誤りの削減となった。

研究成果の学術的意義や社会的意義

聴覚障害学的な観点からのコーパス作成はこれまでも世界的に行われてきたが、音声認識を含む音声言語処理への応用を視野に入れた50人ほどの大規模コーパスは例がなく、聴覚障害者を対象とした音声言語処理技術の進展に大きく寄与するものである。重度聴覚障害者の発話は多様性があり、既存の音声認識ソフトウェアで十分な認識性能に達する者もいれば、調音運動の不安定性から個人適応のような手法によっても性能の改善が限定的であるなどの知見が得られており、聴覚障害者を支援する音声インターフェースを備えたデバイスの実現に向けて前進した。

研究成果の概要（英文）：For deaf and hard-of-hearing (DHH) individuals with ambiguous articulation, speech interfaces with automatic speech recognition (ASR) have been challenging due to insufficient ASR developed for normal hearing. In order to solve this issue, a corpus of 50 DHH individuals' speech was developed, and a large number of read and spontaneous speech, i.e., conversation, were compiled in this study. The developed corpus was also employed to improve the ASR performance for DHH. The results showed a significant reduction in phoneme errors for the DHH speech in read speech, although not as significant as for normal hearing.

研究分野：音声言語処理

キーワード：ろう・難聴 音声認識 音声コミュニケーション

## 1. 研究開始当初の背景

健聴者と聴覚障害者とのコミュニケーションでは通常、筆談、手話通訳、音声認識といった支援手段が用いられる。しかし、筆談は障害者の作文能力が必要であり、日本語能力の不足がちな聴覚障害者にはハードルが高い。手話通訳は、意思伝達手段として優れているが、通訳者は全国で3700人ほどしかおらず、日常的にサポートを受けられるわけではない。音声認識は聴者が意思を伝えることは可能だが、正しく発話する能力が不十分な聴覚障害者は、音声認識の恩恵を受けることが難しい。

一方で、聴覚障害者のコミュニケーション手段はさまざまである。厚生労働省の調査によれば、手話を日常的に用いているのは聴覚障害者の概ね25%にすぎず、音声コミュニケーションを行う者が50%以上（補聴器・人工内耳装用と筆談の総計）と報告されている[1]。また、平成30年の雇用障害者数は53万人と過去最高であり、障害者雇用率制度により一定の障害者の雇用が民間企業に義務づけられていることから、聴者との意思疎通の機会が今後増加することが予想され、聴者と聴覚障害者が円滑なコミュニケーションをはかるための支援が必要となる。

音声認識技術はこれまで、聴こえる側の音声をテキストに変換して聴覚障害者に提示することによって障害者支援を行ってきた。既存の音声認識は情報保障を受けるための聴者が利用する技術であって、調音上の課題を抱えた聴覚障害者にとってみずからがこれを利用しコミュニケーション手段や社会参加の方法として利用されたことはなかった。

## 2. 研究の目的

本研究は、聴者と聴覚障害者用の円滑なコミュニケーションを行うための、聴覚障害者の音声認識技術の開発が目的である。一般に聴覚障害者のコミュニケーションに関しては、手話認識の研究、あるいは聴者の音声をテキストに変換して字幕等で提示する情報保障の研究がある。本研究は、こうした従来の研究とは一線を画し、聴覚障害者の発話内容の推定に音声そのものを利用する点に新規性を置く。すなわち、構音上の問題が認められたとしても、障害者特有の発話の性質を用いて発話内容を推定する。本研究により、聴覚障害者の音声の特徴や、聴覚障害者向け音声インターフェースの開発に関する新たな知見を得る。

## 3. 研究の方法

### (1) 聴覚障害者音声コーパス

聴覚障害者の音声の分析や音声言語処理への応用を行うためには、音声コーパスの構築は不可欠である。聴覚障害者を対象としたコーパスは、音声学的な観点からの解析のためのコーパスが存在するが[2]、音声言語処理を目的としたコーパスではない。音声言語処理の分野では、本邦に限っていても聴者の発話を収集した音声コーパスが数多く存在する。音声言語処理の研究で標準的に用いられるコーパスの中には、読み上げのデファクトスタンダードであるATR音素バランス文が含まれているものが多く、聴覚障害者の音声認識性能を聴者のそれと比較するには、聴覚障害者による標準的な文章の読み上げ音声を収集すべきであるといえ、本研究ではまず研究に供する音声コーパスの整備を行った。

### (2) 音声認識手法

音声認識手法に関しては現在、深層学習の枠組みを用いて音響特徴や音声波形そのものと文字・語を対応づけるend-to-end（エンド・ツー・エンド）音声認識手法が数多く提案されている。End-to-end手法の採用には多くの利点がある。例えば、従来の音声認識では、音素の系列から単語への変換で用いられる発音辞書を人手により作成する必要があった。単語が10万以上ともなれば、発音辞書の作成は極めて高コストとなる。End-to-end音声認識であれば、文字や語を直接推定するため発音辞書は不要である。End-to-end音声認識の利点を活かすことにより、低コストのソフトウェアの研究開発が可能となる。

End-to-end音声認識では、その多くがエンコーダー・デコーダーモデルとよばれる方式を採用している。エンコーダー、デコーダーはそれぞれニューラルネットワークを用いて実装される。エンコーダーは音響特徴のフレームの系列を受け取って、音声の特徴を端的に表すような多次元のベクトルに変換する。一方、デコーダーにはさまざまな実装があるが、トランスデューサー[3]を採用した手法では、エンコーダーから出力される多次元ベクトルと出力済みの過去の文字・単語の系列から当該系列に引き続く文字・単語を予測する。トランスデューサーの特徴は、CTC（Connectionist Temporal Classification）に比べ単語系列を予測に用いるため頑健な認識が可能であることから、本研究で採用することとした。

### (3) データ拡張による音声認識

聴覚障害者や超高齢者のような非典型的な音声を対象とした音声認識では、非典型音声から聴者への声質変換に基づいて音声認識を行う手法[4]や、Cycle-GANにより若年層から非典型的な超高齢層への声質変換を行い[5]、学習データを拡張することによって音声認識性能を改善するといった手法が提案されてきた。本研究では、上述の聴覚障害者および聴者による音素バランス文の発話からパラレルコーパスを構成し、パラレルコーパスを利用した声質変換手法によりデ

一タ拡張を試みる。従来法である Cycle-GAN による声質変換はパラレルコーパスを必要としない利点がある一方、Voice-Transformer [6] のような Transformer ベースの声質変換手法では、パラレルコーパスを構成する制約はあるものの、外部より話者情報が容易に導入できることや、ターゲット音声ソース音長の長さには拘束されないといった特徴がある。本研究では Voice-Transformer を用いてソース音声を聴者/聴覚障害者、ターゲット音声を聴覚障害者として声質変換を行い、拡張した聴覚障害者の音声データを用いて学習することによって、音声認識性能の改善を試みる。Voice-Transformer による声質変換では、Transformer エンコーダーおよびデコーダーに話者性を表す話者ベクトルを入力することにより、many-to-many (あるいは any-to-many) の声質変換が行われる。したがって、適切な話者ベクトルをデコーダーに入力することにより、生成される音響特徴量を変化させることが可能である。

#### (4) 音韻特徴を利用した音声認識

聴覚障害者発話の音声認識の誤り傾向を調べるうえで、音素は単語を他の単語と区別するための最小単位であることから有用である。音素は音韻論という観点から調音によって特徴付けられる。音韻の特徴とは、音素よりも小さな単位で、音素間にまたがる特徴を捉えるために使用される。例えば、/s/と/z/という音素を比較すると、どちらも歯茎摩擦音であるが、有声/無声という対立関係がある。音韻特徴を用いることで、音素を別の観点から評価する尺度を導入することとなり、その結果、曖昧な聴覚障害者音声の音素を識別できる可能性がある。

本研究における音韻特徴は、音素を特徴付ける複数のクラスとして定義される。したがって、音素と音韻特徴を表すクラスは、調音特徴を用いて音素認識を行うニューラルネットワークの出力シンボルとなる。そこで、音韻特徴を5つのグループ(有声/無声、母音調音時の口形、子音の調音位置など)に分け、グループ内で対立関係にあるいずれかの素性を活性化して出力するように設計し、one-hot ベクトルとして出力するようにした。これを従来の音素出力としてマルチタスク学習を行う。

## 4. 研究成果

### (1) 聴覚障害者音声コーパスの整備

研究代表者らは、ATR 音素バランス文を始め、JNAS (新聞記事読み上げ音声コーパス) の新聞読み上げ、会話としての自由発話を中心に聴覚障害者の発話の収集を行った。音声の収録にはヘッドウォーンマイクロホン、ダイナミック/コンデンサーマイクロホンなどを用いた。収録場所は、外部騒音の少ない部屋あるいは無響室 (COVID-19 以前)、ドアを解放した室内でアルコール消毒が可能な簡易ブース (COVID-19 以後) を利用した。また、リモート会議ソフトウェア zoom を用いてオンラインで収録を行なった話者も存在する。

読み上げ文は、51 人の聴覚障害学生から 23.5k 文 (整備途上のものを除く) の収集を行った。また自由発話は、学生のべ 18 名に対して 3.4 時間 (対話としての合計時間 6.1 時間、整備途上を除く) となった。

### (2) 音声認識における聴者と聴覚障害者の性能比較

まず、聴覚障害者音声コーパス作成の序盤で収集したデータ (12 名) を用いて音素認識実験による性能調査を行った。実験では、ESPnet [7] のハイブリッド 注意機構/CTC によるネットワークを用いた。ネットワークは、VGG を前段に持つユニット数 320、4 層の双方向 LSTM のエンコーダーおよびユニット数 320 の location-based の注意機構を持つデコーダーとした。入力音響特徴量は、80 次元のメルフィルタバンク特徴量とした。評価データは健聴者 6 名、聴覚障害者 12 名とした。どちらも ATR 音素バランス文 B・C セットを読み上げたものである。また、注意機構/CTC のマルチタスク学習では、CTC 損失に対する重みを 0.0、0.5、1.0 と変えて実験した。

実験結果 (音素誤り率) を図 1 に示す。図中、聴覚障害者のベースラインとは、聴者の音声データのみを用いて学習したモデルによる評価 (聴者の評価と同様のモデル) を指す。また、障害者適応は評価話者を除く障害者音声データを用いてベースラインのモデルパラメータをチューニングしたモデルによる評価結果、個人適応は B セットに対して C セットを、C セットに対して B セットを的応用データとして障害者適応モデルをチューニングしたモデルによる評価結果である。図中の誤り率は B, C セットの誤り率の平均である。

聴者と聴覚障害者の認識結果を比較すると、聴者の音声のみから学習したモデルを用いた結果 (ベースライン) では、聴覚障害者の音声認識性能が極めて低いことがわかる。一方、聴覚障害者につ

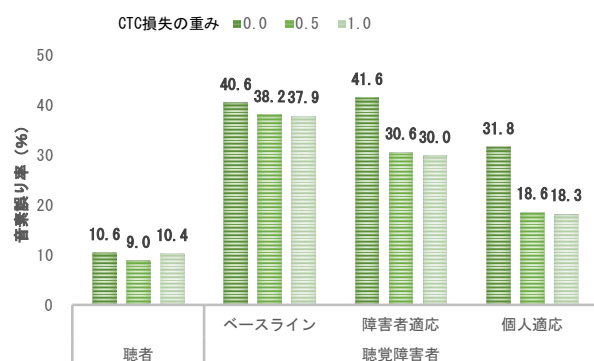


図 1 聴覚障害者と聴者の音素誤り率の比較

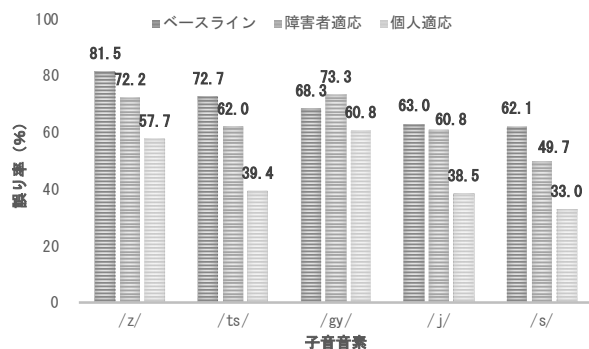


図 2 聴覚障害者の子音音素の誤りの比較

### (3) データ拡張による聴覚障害者の音声認識

声質変換に基づくデータ拡張を行って学習データを増やし、これを用いて学習したモデルによる実験結果を図 3 に示す。図中の数字は音素誤り率を表す。「ベースライン」は聴者音声のみから学習した Conformer-Transducer モデルでの評価、「障害者」は先述の聴覚障害者音声コーパスを用いて、ベースラインのモデルパラメータを更新したモデルによる評価、データ拡張はそれぞれ、

- ① 手法 1: 聴覚障害者 (4 名) のみで声質変換モデルを学習し、障害者音声のデータ拡張を行って得たデータを用いて、ベースラインのモデルパラメータを更新して評価
- ② 手法 2: 聴覚障害者および聴者 5 名の音声で声質変換モデルを学習し、障害者音声のデータ拡張を行って得たデータを用いて、ベースラインのモデルパラメータを更新して評価
- ③ 手法 3: 聴覚障害者および聴者 5 名の音声で声質変換モデルを学習し、障害者音声のデータ拡張を行って得たデータを用いて、ベースラインのモデルパラメータを更新して評価

また、図中のクロズドは評価データに声質変換を行ったデータの話者を含み、オープンは含まない場合の評価である。

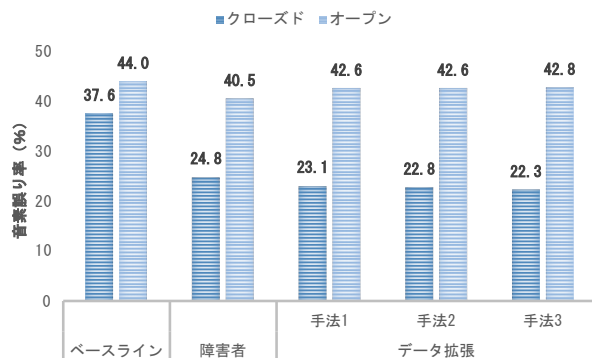


図 3 データ拡張による音素誤り率の比較

で音素誤り率が削減していることがわかる。障害者音声のみで学習した「障害者」が音素誤り率 24.8%であったのに対し、10 人の聴者のデータを用いて声質変換のモデルの学習を行い、データ拡張を行った手法 3 が音素誤り率 22.3%となった。これは 10.1%の誤り削減率である。また、前節で述べたコンフュージョンが起きやすい音素 /z/ や /ts/ など大幅に誤りが削減された。一方でオープンデータでは、データ拡張による音素誤り率の改善は得られておらず、コンフュージョンも解消されなかった。このことから、データ拡張に基づくモデルの学習によって、個人の持つ音韻性に対して頑健になるものの、聴覚障害者に共通する音韻の曖昧性に対しては頑健にはなっていないと考えられる。

### (4) 音韻特徴を用いた聴覚障害者の音声認識

音韻特徴を用いた聴覚障害者の音声認識では、評価対象として 14 名の聴覚障害者による音素バランス文 (計 699 文) を用いた。また比較のため、聴者 120 名による音素バランス文 5,996 文を用いて評価した。実験に用いたニューラルネットワークは Conformer-Transformer とし、前節で述べた音韻特徴および従来の音素に対して表 1 に示す損失関数の組み合わせを変更して実験を行った。表中、CE はクロスエントロピー損失を、AdaCos は深層距離学習で用いられる損失関数である [8]。提案法-1 はデコーダー側に音韻特徴に基づく損失関数を追加して学習、提案法-2 はエンコーダー側にも音韻特徴に基づく損失関数を用いた。

いて、図中の 3 つのモデルのいずれを用いる場合であっても注意機構を使うより CTC 損失単独で学習した方が、音素誤り率が低くなった。聴覚障害者の不明瞭な発話では、入力に制約がかかる前向き後ろ向きアルゴリズムに基づく CTC を損失関数としたほうが頑健に学習できると考えられる。

図 2 はベースラインで音素誤り率の高かった子音 (B, C セットで頻度 10 以上) について、音素誤り率の変化を表したものである。聴覚障害者音声を加えたモデル (障害者適応) では音素誤りの削減は小さく、大幅な誤りの削減には話者適応を行う必要がある。

声質変換で用いたデータと評価データの話者が一致しているクロズドの結果を見ると、ベースラインでは 37.6%の音素誤り率であったものが、聴覚障害者の音声を学習に追加することにより 24.8%まで低下しており、聴覚障害者の音声に対してもモデル個人適応が有効であることがわかる。一方で、オープンなデータを評価した場合、誤り率の削減は小さくなった。聴覚障害者にみられる発話の特徴は誤り率の削減に反映されていると考えられるが、障害者個人の発話手法の多様性が原因で、誤りの削減に限界があったものと考えられる。データ拡張手法は、手法 1 から手法 3 にわたって、声質変換の学習に用いた聴者の異なり数を増やしているが、聴者の数を増やすにしたがって

表 1 損失関数の組み合わせ

	エンコーダー		デコーダー	
	CTC (音素)	CTC (音韻)	CE (音素)	AdaCos (音韻)
ベースライン	○		○	
提案法-1	○		○	○
提案法-2	○	○	○	○

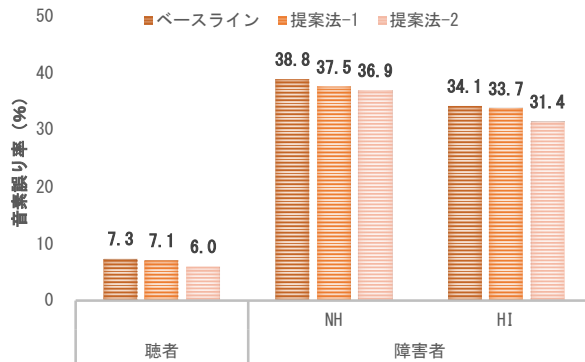


図 4 音韻特徴による音素認識誤り率の比較

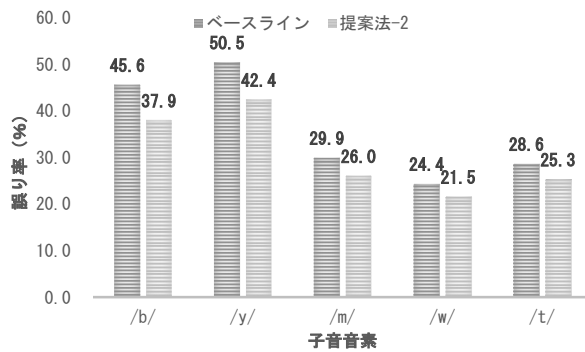


図 5 代表的な子音音素の誤り削減の比較

図 4 に音素誤り率を示す。図中、NH は聴者の音声のみから学習したニューラルネットワークを、HI は聴覚障害者の音声を用いて追加学習したネットワークである。ベースラインに対してデコーダー出力に音韻特徴による AdaCos 損失を加えた (提案法-1) 場合、NH モデルによる障害者音声の音素誤り率はベースラインの 33.8% から 36.9% となり、ベースラインに比べて 3.4% の誤り削減率となった。比較で行った聴者の評価音声では、ベースラインが 7.3%、提案法-1 が 7.1% であり、誤り削減率は 2.7% となって聴覚障害者同様、認識誤りを削減できている。

一方、障害者の音声を用いて追加学習した HI モデルを用いて障害者音声の評価した場合、ベースラインでは音素誤り率が 34.1% であったのに対し、提案法-1 では 33.7% となり、1.2% の誤り削減となった。エンコーダー出力に対しても音韻特徴による損失 (提案法-2) を加えると、NH モデルによる障害者音声の音素誤り率は 36.9% となり、ベースラインに対して 4.9% の誤り削減となり、NH モデルのベースラインと比較した場合 7.9% の誤り削減率となった。エンコーダーおよびデコーダーの両者に音韻特徴による損失を加えた場合に誤り削減効果が高く、聴者・聴覚障害者両者に対して音素誤りを削減したことから、音韻特徴は音素に基づく損失に対して相補的に機能しているのではないかと考えられる。また、子音に比べて母音の方が単純な調音であり、音韻特徴による効果は母音に対して高かった。子音の誤りについて調べた。図 5 は、ベースラインと提案法-2 を比較して、子音の誤りに対する誤り削減率の上位 5 つを示したものである。有声両唇

破裂音/b/が最も誤りが削減された。これは両唇鼻音/m/との混同が解消されたためである。また、両唇接近音/w/、硬口蓋接近音/y/も誤りから回復したことがわかる。/w/と/y/のような接近音は、調音において母音に近いため、おそらく母音と同様、音韻特徴により誤り削減の恩恵を受けているとみられる。また、話者別の音素誤り率を調べたところ、提案法-1 は一部の話者で誤りを削減できなかったが、提案法-2 はすべての話者で誤りを削減した。このことから、聴覚障害者に対して頑健な音声認識を行うためには、エンコーダー/デコーダーモデルの学習では、両者に対して音韻特徴のような相補的な情報を用いて学習すべきであると考えられる。

引用文献

[1] J. Bigham et al., Proc. ACM SIGACCESS, pp. 383-384, 2017.  
 [2] L. Mendel et al., An update on hearing loss, pp. 103-128, 2015.  
 [3] A. Graves, arXiv:1211.3711, 2012.  
 [4] 下笠ほか, 音講論集, pp. 1055-1056, 2021.  
 [5] Chen et al., Proc. Interspeech, pp. 4828-4832, 2021.  
 [6] Kameoka et al., IEEE Trans. ASLP, pp. 656-670, 2021.  
 [7] Watanabe et al., Interspeech, pp. 2207-2211, 2018.  
 [8] Zhang X. et al., Proc. CVPR, pp. 10823-10832, 2019.

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 0件 / うちオープンアクセス 0件）

1. 著者名 Kitaoka Norihide, Chen Bohan, Obashi Yuya	4. 巻 -
2. 論文標題 Dynamic out-of-vocabulary word registration to language model for speech recognition	5. 発行年 2021年
3. 雑誌名 EURASIP Journal on Audio, Speech, and Music Processing	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) 10.1186/s13636-020-00193-1	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 Kobayashi Akio, Onishi Junji, Nishizaki Hiromitsu, Kitaoka Norihide	4. 巻 -
2. 論文標題 End-to-End Speech to Braille Translation in Japanese	5. 発行年 2022年
3. 雑誌名 2022 IEEE International Conference on Consumer Electronics (ICCE)	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/icce53296.2022.9730468	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 Kobayashi Akio, Yasu Keiichi, Nishizaki Hiromitsu, Kitaoka Norihide	4. 巻 -
2. 論文標題 Corpus Design and Automatic Speech Recognition for Deaf and Hard-of-Hearing People	5. 発行年 2021年
3. 雑誌名 2021 IEEE 10th Global Conference on Consumer Electronics (GCCE)	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/gcce53005.2021.9621959	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 Daiki Mori, Kengo Ohta, Ryota Nishimura, Atsunori Ogawa, Norihide Kitaoka	4. 巻 -
2. 論文標題 Advanced language model fusion method for encoder-decoder model in Japanese speech	5. 発行年 2021年
3. 雑誌名 Proc. APSIPA ASC 2021	6. 最初と最後の頁 503-510
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 Koharu Horii, Meiko Fukuda, Kengo Ohta, Ryota Nishimura, Atsunori Ogawa, Norihide Kitaoka	4. 巻 -
2. 論文標題 End-to-end spontaneous speech recognition using hesitation labeling	5. 発行年 2021年
3. 雑誌名 Proc. APSIPA ASC 2021	6. 最初と最後の頁 1077-1081
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 Wang Yu, Leow Chee Siang, Kobayashi Akio, Utsuro Takehito, Nishizaki Hiromitsu	4. 巻 -
2. 論文標題 ExKaldi-RT: A Real-Time Automatic Speech Recognition Extension Toolkit of Kaldi	5. 発行年 2021年
3. 雑誌名 Proceedings of the 2021 IEEE 10th Global Conference on Consumer Electronics (GCCE 2021)	6. 最初と最後の頁 346-350
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/GCCE53005.2021.9621992	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 Leow Chee Siang, Hayakawa Tomoaki, Nishizaki Hiromitsu, Kitaoka Norihide	4. 巻 -
2. 論文標題 Development of a Low-Latency and Real-Time Automatic Speech Recognition System	5. 発行年 2020年
3. 雑誌名 Proceedings of the 2020 IEEE 9th Global Conference on Consumer Electronics (GCCE 2020)	6. 最初と最後の頁 464--467
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/GCCE50665.2020.9291818	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 Wang Yu, Leow Chee Siang, Nishizaki Hiromitsu, Kobayashi Akio, Utsuro Takehito	4. 巻 -
2. 論文標題 ExKaldi: A Python-based Extension Tool of Kaldi	5. 発行年 2020年
3. 雑誌名 Proceedings of the 2020 IEEE 9th Global Conference on Consumer Electronics (GCCE 2020)	6. 最初と最後の頁 470--473
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/GCCE50665.2020.9291717	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 Chen Jiahao, Nishimura Ryota, Kitaoka Norihide	4. 巻 9
2. 論文標題 End-to-end recognition of streaming Japanese speech using CTC and local attention	5. 発行年 2020年
3. 雑誌名 APSIPA Transactions on Signal and Information Processing	6. 最初と最後の頁 1--7
掲載論文のDOI (デジタルオブジェクト識別子) 10.1017/ATSIP.2020.23	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 Arongna, Naomi Sakai, Keiichi Yasu and Koichi Mori	4. 巻 63
2. 論文標題 Disfluencies and Strategies Used by People Who Stutter During a Working Memory Task	5. 発行年 2020年
3. 雑誌名 Journal of Speech, Language, and Hearing Research	6. 最初と最後の頁 688--701
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17件 (うち招待講演 0件 / うち国際学会 0件)

1. 発表者名 小林彰夫、安啓一
2. 発表標題 ラジオ音声を対象とした音声強調における擬似学習データ利用の検討
3. 学会等名 第86回情報処理学会全国大会
4. 発表年 2024年

1. 発表者名 藤江匠汰、安啓一、小林彰夫
2. 発表標題 混合音声から抽出した難聴者の発話の聞き取りやすさ に関する客観的および主観的な評価による検討
3. 学会等名 情報処理学会アクセシビリティ研究会
4. 発表年 2024年



1. 発表者名 高橋快斗、木内貴浩、若林佑幸、太田健吾、小林彰夫、北岡教英
2. 発表標題 話者適応による聴覚障害者音声認識の評価
3. 学会等名 情報処理学会第151回音声言語情報処理研究発表会
4. 発表年 2024年

1. 発表者名 高橋快斗、木内貴浩、若林佑幸、太田健吾、小林彰夫、北岡教英
2. 発表標題 聴覚障害者音声における自己教師あり学習に基づく音声認識の評価
3. 学会等名 日本音響学会春季研究発表会
4. 発表年 2024年

1. 発表者名 小林 彰夫、大西 淳児、西崎 博光、北岡 教英
2. 発表標題 読み上げ文を対象としたEnd-to-End音声点訳
3. 学会等名 日本音響学会講演論文集
4. 発表年 2021年

1. 発表者名 森 大輝、太田 健吾、西村 良太、小川厚徳、北岡 教英
2. 発表標題 タスク外音響情報を付加したEnd-to-End音声認識モデルの設計
3. 学会等名 日本音響学会講演論文集
4. 発表年 2022年

1. 発表者名 堀井 こはる, 福田 芽衣子, 太田 健吾, 西村 良太, 小川厚徳, 北岡 教英
2. 発表標題 非流暢ラベルを用いた言い淀み整形End-to-End音声認識
3. 学会等名 日本音響学会講演論文集
4. 発表年 2022年

1. 発表者名 森 大輝, 太田 健吾, 西村 良太, 小川 厚徳, 北岡 教英
2. 発表標題 Encoder-Decoder音声認識モデルにおける暗黙的言語情報の置換法
3. 学会等名 日本音響学会講演論文集
4. 発表年 2021年

1. 発表者名 堀井 こはる, 福田 芽衣子, 太田 健吾, 西村 良太, 北岡 教英
2. 発表標題 言い淀みを考慮した自由発話のEnd-to-End音声認識
3. 学会等名 日本音響学会講演論文集
4. 発表年 2021年

1. 発表者名 森大輝, 太田健吾, 西村良太, 小川厚徳, 北岡教英
2. 発表標題 End-to-end音声認識モデルにおける暗黙的言語情報の置換法
3. 学会等名 音学シンポジウム,
4. 発表年 2021年

1. 発表者名 レオ チーシャン, 王 宇, 小林彰夫, 宇津呂武仁, 西崎博光
2. 発表標題 Kaldiベースのストリーミング音声認識システムの開発
3. 学会等名 日本音響学会2021年秋季研究発表会講演論文集
4. 発表年 2021年

1. 発表者名 陳 家浩, 西村 良太, 北岡 教英
2. 発表標題 Uni-directional LSTM と Local Attentionを用いたストリーミング音声認識
3. 学会等名 日本音響学会講演論文集
4. 発表年 2020年

1. 発表者名 小林彰夫, 安啓一, 西崎博光, 北岡教英
2. 発表標題 聴覚障害者の音声データの収集と音素認識による評価
3. 学会等名 日本音響学会研究発表会講演論文集
4. 発表年 2021年

1. 発表者名 安啓一, 種子田尚人, 石井悠貴
2. 発表標題 若年の補聴器・人工内耳装用者における語音検査と自身の発話明瞭性の関係
3. 学会等名 日本音響学会秋季研究発表会講演論文集
4. 発表年 2020年

1. 発表者名 寺澤 洋子, 水野 真由美, 山本 雄也, 大中 悠生, 石川 嘉秀, 松井 淑恵, 安啓一
2. 発表標題 加齢性難聴に伴うポピュラーソングの印象変化の検討～模擬難聴を用いて
3. 学会等名 日本音響学会秋季研究発表会講演論文集
4. 発表年 2020年

1. 発表者名 大中悠生, 安啓一, 松井淑恵, 寺澤洋子
2. 発表標題 加齢性難聴者のサイン音認知の研究にむけて
3. 学会等名 情報処理学会研究報告
4. 発表年 2020年

1. 発表者名 寺澤洋子, 相馬翔太, 安啓一, 平賀瑠美
2. 発表標題 聴覚障害者の調性認知に関する検討
3. 学会等名 情報処理学会研究報告
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	北岡 教英  (Kitaoka Norihide)  (10333501)	豊橋技術科学大学・工学(系)研究科(研究院)・教授   (13904)	

6. 研究組織（つづき）

	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	西崎 博光  (Nishizaki Hiromitsu)  (40362082)	山梨大学・大学院総合研究部・教授    (13501)	
研究分担者	安 啓一  (Yasu Keiichi)  (70407352)	筑波技術大学・産業技術学部・准教授    (12103)	

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