

音声認識システムと 日本語研究・日本語教育への応用

——テレビ・ラジオ番組など数種を例として——

神 作 晋 一

1. はじめに

現在、PCやスマホ、タブレットなどの端末で使用できる音声認識システム⁽¹⁾が開発されている。一般にはWEB検索や動画サイトの自動字幕機能、スマートスピーカー⁽²⁾や携帯電話の音声アシスト機能⁽³⁾といったAIアシスタント等に利用されている。

こうした音声認識システムには、①音声に音声で答えるもの、あるいは音声と文字の相互変換(②音声→文字、③文字→音声⁽⁴⁾)となるものがある。これらは特定の大学や研究機関などのような特殊な状況でなくても、一般的な個人利用の電子機器・端末(PC、スマホ、タブレット)などで恩恵を受けられる状況に進んできている。

その中②のような音声の自動文字テキスト化は、さまざまな学術研究上や活字メディア上で、インタビューや対談などにおいて話者の発話の文字起こしを時間も労力もかけて行っている現状を考えれば、ニーズは高いものと考えられる。

しかし、音声から自動的に文字テキストを生成させる点において、たとえば著名な動画サイト(YouTube等)の字幕機能を見る限り、その精度は必ずしも完全なものではなく、正確にするにはさらなる修正などが必要なレベルであり、課題を残している。

とはいえ、こうした状況の下で、日本語研究では生の音声データをテキスト化処理できれば分析等に有効であり、また、日本語教育では学習者などの発話が自動的に文字テキスト化されるようになれば、フィードバックなどにも有効であると考えられる。

本研究ではgoogleの音声認識システムを利用し、日本語のテレビやラジオ番組の音声の文字テキスト化が、日本語研究や日本語教育で利用するという観点から、どのくらい可能か、いくつかの事例ケーススタディとしてではあるが、分析・考察することを目的とする。

2. 先行研究・概要など

音声認識とは、たとえば河原達也編(2016)によれば「入力音声から言語情報を抽出し、

文字列（日本語の場合はかな漢字列）に変換する処理」(P.1)である。この一連の処理を支える条件や要因としては、以下の4つ⁽⁵⁾が述べられている。

1. 利用話者：特定の話者から不特定多数の話者への対応。
2. 語彙サイズ：少数の語彙から大量、単語連鎖をモデル化した大規模な発話データへ。
3. 発話スタイル：単語発音から連続、発話のバリエーションへの対応。
4. 使用環境：マイク利用から自然発話、生活環境への対応。

それぞれの項目の要点を見出しの後に要約したが、これまで、それぞれの研究が積み重なってきて、より広い条件に合致するところまできたということが言える。

こうして開発された音声認識の技術を実際に使用するための装置として、アプリケーション（アプリ、APP）があるが、河原氏によれば、使用環境・語彙サイズ・発話スタイルなどの観点から「プロット」すると、代表的なものとして、以下の8つが挙げられるという。

1. テキスト入力（ディクテーション）：タイプやワープロにおける音声入力（キーボード入りに習熟、人前で音声は…⁽⁶⁾）
2. 音声によるコマンド入力（カーナビ・ゲーム機など）：特定のコマンド実行からクラウドサーバ環境になり、用途が広がる。
3. 音声による情報アクセス（電話応答装置・携帯端末）：「siri」や「しゃべってコンシェル」、アシスタントソフト、AIスピーカーなど
4. 音声による会話（人間型ロボット・エージェント）：現在は自由な話し言葉への対応が困難で単純な会話のみ、非定常な雑音源で利用者の音声を拾うのに苦勞する。
5. 音声の書起し（会議録・講演録・字幕付与）：人間同士の自然なコミュニケーションを対象とするのは困難。パブリックスピーキング（ニュース番組や議会、学会講演など）のモデル化で実用的なレベルに。テレビ番組の字幕付与や議会の会議録で実用化。
6. 音声の検索・マイニング：長時間・大規模の音声データからのマイニング・分析（コールセンターなど）
7. 音声翻訳：外国人とのコミュニケーションを支援するもの、同時通訳、音声の書き起こしと関連。外国語の音声アーカイブの検索など
8. 語学学習支援⁽⁷⁾：外国語の発音や会話スキルの支援（例：日本人にとっての英語）。間違いやすいパターンのモデル化。コンテンツの作成においては語学教室や教師との連携が必要。

それぞれ項目ごとの要点をかなり簡縮化してまとめたが、まず日本語研究についてはコーパスのようなものを作成する⁽⁸⁾という意図で5の音声の書き起こしや6の音声の検索、マイニングというのがそれにあたるだろう。6でもかなり有効であるが、処理や分析、論文などで文字列としてアウトプットすることを考えると、5の書き起こしというのはまだまだ求められることになると思う。そして日本語教育においては5や6も教材作成などに有効であり、さらに8のような外国語習得の支援というところにたどり着く。

今回は音声認識システムによる文字テキスト化に着目し、いわゆる発話の文字起こしが通常の環境でどのくらい可能なのか考察し、何がどのくらいできるのかを検証してみたい。

3. 調査対象

3.1.使用環境

今回は PC などのコンピュータの利用ができる人なら誰でも取り組みやすいものとして「Google ドキュメント <https://www.google.com/intl/ja/docs/about/>」を利用した。

「Google ドキュメント」は、まずブラウザに「クローム chrome」を使用し、Google のトップページ <https://www.google.com/>を開き、画面右上のアイコンから「ドキュメント」を開く。ドキュメントでは作成画面の「ツール」のタブ→「音声入力」でマイクのアイコンをクリックした後、音声での入力が可能となる。

今回はノート PC (windows10環境) で chrome から Google ドキュメントを開き、スマホ (iphone) で音声ファイルや動画ファイルを再生し、PC本体にある内蔵マイクあるいは USB 接続のマイクに音声を近づけて行った。後述するいくつかの音声を認識させたが、音が聞き取れず止まる⁽⁹⁾ということはほとんどなかった。

3.2.内容 (コンテンツ)

使用する内容 (コンテンツ) はラジオ番組数種とテレビ番組から選んだ。いわゆる加工した教材ではなく、生教材や生の言語資料として加工・処理できるかというのを図ったものである。

ラジオ番組 1 : 「飯田浩司の OK!cozy up」(ニッポン放送)⁽¹⁰⁾ ※ニュースと解説・コメント

(2018年11月27日放送分から抜粋、約10分)

<https://www.youtube.com/watch?v=unanBa1yPn8> 34:13~43:33⁽¹¹⁾

出演者: 飯田浩司 (司会、ニッポン放送アナウンサー)、有本香 (コメンテーター、ジャーナリスト)

ラジオ番組 2 : 「飯田浩司の OK!cozy up」(ニッポン放送) ※アナウンサー 2 人の対話

(2019年03月07日放送分から抜粋、約 4 分)

<https://youtu.be/0VwvsPvAqcE> 11:13~15:32

出演者: 飯田浩司 (司会、ニッポン放送アナウンサー)、新行市佳 (ニッポン放送アナウンサー)

ラジオ番組 3 : 「サンドウィッチマンの東北魂」(ニッポン放送)⁽¹²⁾ ※メール紹介とコメント (対話)

(2019年7月7日、podcast 配信分から抜粋、約12分)

出演者: サンドウィッチマン (伊達みきお、富沢たけし)

<https://omny.fm/shows/sand/playlists/sand/embed?style=cover>

テレビ番組 1 : 「真相深入り虎ノ門ニュース」(DHC テレビ)⁽¹³⁾ ※ニュースとコメ

ント

(2018年11月29日放送分から抜粋、約8分)

出演者：居島一平（司会）、有本香（コメンテーター、ジャーナリスト）

【DHC】11/29(木)有本香×居島一平【虎ノ門ニュース】(虎8チャンネル、niconico⁽¹⁴⁾)

<https://www.nicovideo.jp/watch/1543475222> 80:04~81:24,88:14~95:39⁽¹⁵⁾

4. 認識結果と分析・考察

4.1. ラジオ番組

4.1.1. ラジオ番組1（ニュースと解説、コメント）

この番組はいわゆるニュース解説番組で、ニュースの読み上げの部分とコメンテーターの解説・コメントの部分で構成されているものである。まずはニュースの場面について示してみる。

認識結果（ニュースの部分）

台湾の選挙結果を受け中国政府が中国との関係の発展を望む人々の願いだとコメント

台湾で24日2020年に行われる総統選挙の前哨戦とされる統一地方選挙が全土で行われ与党民進党が大敗し中国に融和的な姿勢を示す最大野党の国民党が躍進しましたこの結果を受け中国政府は中国との関係の発展を望む人々の願いを反映したものだ」とコメントを発表しております

筆者による聞き取り（修正）

「台湾の選挙結果を受け、中国政府が『中国との関係の発展を望む人々の願いだ』とコメント。」

「台湾で24日、2020年に行われる総統選挙の前哨戦とされる統一地方選挙が全土で行われ、与党民進党が大敗し、中国に融和的な姿勢を示す最大野党の国民党が躍進しました。この結果を受け中国政府は、『中国との関係の発展を望む人々の願いを反映したものだ。』とコメントを発表しております。」

これを見ると、認識率が驚くほど高い。表記などの不具合もなく、そのまま句読点や区切り記号を補充すれば、全く問題ないほどである。プロのアナウンサーが話したということもあるだろうが、マイクや音量などの環境も良かったのであろう。

認識結果（解説トーク、対話の部分） [抄録] 35:30~36:48（約1分20秒）

今年の8月の終わりにですね台湾に行きましてその時はですね台湾のあの国立政治大学っていう大学があるんですけどもそこでその前にほんのね政治家なんかを勉強しておられる学生さんを相手に講演をさせてもらいました大学院学生さんですの方に通常はですねだいたい研究者の方が行くことが多いんですけども私はその現実の音日本の政治正解を取材してるものっていうことであの向こうの人達にとってみれば一番そのほらリアルの所触ってるという意味でですねまあ興味を持って色々聞いて下さった

んですね同時にやっぱりその台湾の今誠司ってどうなってんのって意外にね台湾の情報ってすごくたくさんの方に入ってくるんですけどあの人の行き来も当然多いからすごく台湾のこと我々知ってるってイメージあるんだけど文化面はすごく伝わってくるわけですよ何か言わの食事はおいしいとこです台湾のドラマとか芸能とかですなだけ台湾の選手ってどうなってるかて実は案外わかりづらくてね相当せんでええ蔡英文刺さったけれども実はその正式版画板弱化したと必ずしもそうとは言えないわけですよ言えなかったんですよ

筆者による聞き取り（修正）

（あの一、実はね、これ、私は）今年の八月の終わりにですね、台湾に行きまして、【はい】その時はですね、台湾のね、あの一国立政治大學っていう大学があるんですけども、そこでその、まあ日本のね、えー政治やなんかを勉強しておられる学生さんを相手に【はい】あのまあ、講演をさせてもらいました。あ、でまあ、大学院の学生さんですね、主にね。まあ、通常はですね、だいたい研究者の方が行くことが多いんですけども、まあ、私はその現実のね、日本の政治・政界を取材している者っていうことで、あの向こうの人たちにとってみれば、一番、そのほら、リアルなところを触ってるっていう意味でですね、まあ興味を持っていろいろ聞いてくださったんですね。同時に、やっぱり台湾の今政治ってどうなってんのって…意外にね、台湾の情報ってすごくたくさん日本に入ってくるんですけど、で、あの、人の行き来も当然多いから、すごく台湾のこと我々知ってるってイメージがあるんだけど…【ええ】文化面はすごく伝わってくるわけですよ【うーん】台湾の食事はおいしいとかですね、【はい】まあ、台湾のまあドラマとか芸能とかですね…だけど、台湾の政治ってどうなってってかて、実は案外分かりづらくて【はい】で、総統選で【確かにそうですね】、えー、蔡英文さん勝ったけれども、実はその政治基盤が盤石かっていうと、必ずしもそうとは言えないわけですよ、言えなかったんですよ。

ここでは紙幅の都合もあり、コメンテーターによる解説の部分を一部だけ抜粋してみた。なお、認識結果がおかしい箇所とそれに対応する筆者の読み取り部分には下線を施した。また、司会者や聞き手ののあいづち等は【 】で示してある。（これ以降の挙例も同様とする。）

さすがにフィルターのような部分や、司会者の相槌のようなものなどはあまり拾えてないが、文脈を追って把握できる程度には認識されていることが分かる。明らかな変換ミスや話者の発音が少し不明瞭だったり、音声の読み取り時の不具合だったりと思うようなところもあるのだが、それでも聴きなおしながら、余裕をもって直せるレベルであることが分かる。

4.1.2.ラジオ番組2（情報提供、コメント、対話）

この部分は「モーニングライフアップ」という番組のコーナーの一つで、アナウンサーあるいはスタッフが取材した事項を紹介するものである。紹介する箇所はアシスタントの

新行市佳が中心に話し、飯田浩司アナが聞き手に回っているものである。

認識結果 (抜粋) 12:48～15:32 (2分44秒)

タピオカブームには三つの理由があるんだそうです。一つ目は何と言ってもやっぱりインスタ映えだったけど会えるのがいるんですよ。タピオカドリンクってちょっと可愛かったりもするじゃないですかとつぶつぶ学校検索してみると、理由はここ10年位で台湾旅行や台湾の食に対する人気が急激に高まっていることも関係しているようです。なろうどの2018年の人気海外旅行先ランキングでは台湾が菌を抜いて堂々の一位になったんです。そうなんだね、はい、抜いたんだよ確かにね行きも帰りも飛行機も満席でした来ないで来ましたもんね、健康意識の高まりですが先ほどもお話ししたんですが今回のタピオカブームはミルクティーやお茶に入れて飲むというドリンクとしてのタピオカお茶=体にいいという風なイメージがあるので健康意識の高い女性たちが中心になってお茶専門店に足を運ぶようになったからではないかと推測されています。ちなみに有楽町銀座界隈にも出て専門店が4点ほどありましてそうです。何もタピオカというよりはお茶を、ちょっとしたらっぱい食べ方ですね。コンビニも専門店なりに種類が豊富になってきています。体験したことないと言うか初めて見てもいいのかもしれない。今再びブームになっているかを取り上げました。

筆者による聞き取り (修正)

新行：タピオカブームには三つの理由があるんだそうです。一つ目は何と言ってもやっぱりインスタ映えだと。

飯田：出た！映(ば)えるの？

新行：ばえるんですよ。タピオカドリンクってちょっと可愛かったりもするじゃないですか、そのつぶつぶがこう…

飯田：つぶつぶが大きいもんね、今。

新行：そうなんです。実際ですね。インスタグラムで#ハッシュタグ「タピオカ」と検索してみると、なんとですね、73万件もの投稿が…。

飯田：そんなにあるの？

新行：あるんだそうですよ。【へー】2つめの理由はここ10年位で台湾旅行や台湾の食に対する人気が急激に高まっていることも関係しているようです。

飯田：は一なるほどね。

新行：あのう2018年の人気海外旅行先ランキングでは台湾がハワイを抜いて、【ほう】堂々の一位になったんです。

飯田：あ、そうなんだね。ハワイ抜いたんだ【とうとう】いや確かにね、行きも帰りも飛行機ほぼ満席で、【そうでした？】そう。

新行：こないだ(台湾に)行きましたもんね。【そうそうそう】そして、三つ目の理由なんですが、【うい】健康意識の高まりです。先ほどもお話ししたんですが、今回のタピオカブームはミルクティーやお茶に入れて飲むというドリンクとし

てのタピオカ【おお】お茶=体にいいという風なイメージがあるので健康意識の高い女性たちが中心になって、お茶専門店に足を運ぶようになったからではないかと推測されています。

飯田：なるほど。

新行：ちなみに有楽町銀座界限にもですね、【はいはい】専門店が4店ほどありまして…

飯田：へー、そうなんだ。

新行：そうなんです。いずれもタピオカというよりはお茶をこう前面に押し出していて…

飯田：ああ、まあ、そうだよな。台湾実はお茶の名産（地）でもあるんだよね。【たしかにそうなんですよね】あの凍頂烏龍茶とかもあるけれども【はい】結構高い山があるから…

新行：そのお茶とタピオカで楽しむというのが今のブームということなんですね。

飯田：うーん、なるほどね。

新行：はい。そのタピオカなんですけど…今コンビニでも買えるようになっていてすよね。【そうなんだ！】ちょっと買って来たんですけど…。

飯田：コーヒーとかさ、マウントレーニヤとかと同じところにあるね。【そうですね】同じような感じ。ちょっといただきますね。（飲んでいる）ちょっと吸ったらタピオカ一杯出てきた。あ、でもお茶自体がおいしい。【おいしいですよ】ミルクティーとしてとっても美味しい。

新行：他にもですね、タピオカ杏仁ミルクティーとか、（タピオカ）いちごみるくティーなどなどですね、コンビニも専門店なみに種類が豊富になってきています。【ほうー】まだね、体験したことないという方も、まずはコンビニのタピオカから始めてみてもいいのかもしれない。【なるほど】今朝のモーニングライフアップは今再びブームになっているタピオカを取り上げました。

全体的にはマイクが反応しなかったのか、認識されていないところも多いが、6～7割くらい反応しているというところであろうか。なお、認識結果で下線を施したのは、聞き役の飯田アナウンサーの部分である。中心となって話す方の話の方がずっと認識されている。やはり話し手が複数いると難しいということになるようである。

4.1.3.ラジオ番組3（メール紹介、読み上げ、コメント）

ラジオ番組の特徴の一つにリスナーからのメール（はがきやFAX）を読むことがあるが、メールを読んでいるときは、一種のパブリックスピーキングと同じなので、比較的認識しやすいことが予想される。

メール1 認識結果（抜粋） 1:37～1:56（19秒）

さんとのコラボ動画見ました YouTube charisma のひかるさんですらさんを前にし

て喉がないほど緊張していましたこれからもサンタさんお疲れさんそしてペヤングさん応援していきたいと思います

メール1 筆者による聞き取り (修正)

伊達：山形県のイイ？さんからいただきました【ありがとうございます】「YouTube ひかるさんとのコラボ動画見ました。YouTube 界でカリスマのひかるさんですらサンドさんを前にして喉が通らないほど緊張していました。えー、これからもサンドさんはもちろん、ひかるさんそしてペヤングさんを応援していきたいと思います」ということです。

「 」で囲んだところがメールの内容であるが、メール1については、ところどころ抜けている。一部活舌のあまり良くないところがあると考えられるのだろうか。

メール2 認識結果 (抜粋) 2:36~4:53 (2分17秒)

現在オーストラリアで競馬の騎手をしてるの子ですオーストラリア毎週キャストでレース前の車の中で楽しく拝聴させていただいております震災が起きたあの日自分はオーストラリアにいたのですが岩手競馬場の方等からその日の様子と被害を聞いておりました自分にも何かできないかと考えたとき JFA 全日本サッカー協会が主催する夢先生というプロジェクトに参加する機会に恵まれましたこれは現役中または引退したアスリートたちが小学校五年生もしくは中2の生徒を対象に学校に訪れ夢を叶えるために自分は何をしてきたかを語るというものです1時間目に体育館で生徒と一緒に運動しながら仲良くなった後次の1時間は教室で生徒たちと語り合います僕はそこに先生として訪れることになったんですが自分が訪れた福島の小学校の子供たちからはたくさんの元気をもらいましたそれではいけない話題もあるので自分は過剰に神経質になっていたところもあったんですが子供たちはその後時にせず未来のことだけを考えポジティブな子供達の言葉に逆に励まされましたまた海外に住む人からのイメージでは福島県原発事故後の放射性物質の影響であそこに行くのは危険だとか東北地方東北地方の食べ物は食べ東北地方の食べ物食べない方がいいなどと誤った情報を受け取ってしまいそのまま信じてしまってる人がまだまだ多いようです今僕にできることはこうして現地に訪れた経験をもとにそういった形の誤解を解くことと元気な子供達にもと世界に飛び出して行って動画を発信していくことの大切さを伝えることだなと感じております僕は中学校卒業後中卒でオーストラリアに渡ったんですがこんな中卒な自分でも夢をかなえることができるんだということをこれからも子どもたちに伝えたいです次は是非とも岩手県の学校にも訪れてみたいですねちなみに自分の苗字は富沢です Pro を始めてサンドイッチマンの二人を見た時から他人ではないような気がしてました

メール2 筆者による聞き取り (修正)

富沢：「現在オーストラリアで競馬の騎手をしてるノジリ (野尻?) です。【オーストラリア、競馬の騎手】毎週ポッドキャストでレース前の車の中で楽しく拝聴させてい

ただいております。【ありがとうございます】震災が起きたあの日、自分はオーストラリアにいたのですが、岩手競馬場の方達からその時の様子と被害を聞いておりました。自分にも何かできないかと考えたとき JFA 全日本サッカー協会が主催する夢先生というプロジェクトに参加する機会に恵まれました。【はい】これは現役中または引退したアスリートたちが小学校五年生もしくは中2の生徒を対象に学校に訪れ、夢を叶えるために自分は何をしてきたかを語るというものです。【ふーん】1時間目に体育館で生徒と一緒に運動しながら仲良くなった後、次の1時間は教室で生徒たちと語り合います。【うん】僕はそこに先生として訪れることになったんですが、自分が訪れた福島の小学校の子供たちからはたくさんの元気をもらいました。【うん】触れてはいけない話題もあるのではと自分は過剰に神経質になっていたところもあったんですが、子供たちはそんなことを気にせず、未来のことだけを考えポジティブな子供達の言葉に逆に励まされました。【うん】また海外に住む人からのイメージでは福島県原発事故後の放射性物質の影響であそこに行くのは危険だとか東北地方東北地方の食べ物は食べ、東北地方の食べ物は食べない方がいいなどと誤った情報を受け取ってしまいそのまま信じてしまってる人がまだまだ多いようです【うん】今僕にできることはこうして現地に訪れた経験をもとにそういった方たちの誤解を解くことと元気な子供達にもっと世界に飛び出して行って動画を発信していくことの大切さを伝えることだなと感じております。【ふーん】僕は中学校卒業後中卒でオーストラリアに渡ったんですが、【すげえな】こんな中卒な自分でも夢をかなえることができるんだということもこれからも子どもたちに伝えたいです。次は是非とも岩手県の学校にも訪れてみたいですね。ちなみに自分の苗字は富沢です。【あら】初めてサンドイッチマンの二人を見た時から他人ではないような気がしてました。」 えー珍しい【ありがとうございます】

これは、富沢が読み手であり、それがほとんど認識されている一方、【 】内の伊達のあいづちなどの反応は認識結果に出ていない。しかしながら、メールの本文自体の認識結果はかなり良くなっており、修正する場合も容易であった。

メール3 認識結果 (抜粋) 6:34~7:27 (53秒)

こちら神奈川県相模原市の方ですねラジオネームガッツ石松先日です北海道に帰ってきました今回は初めて旭山動物園と白い恋人パークに行きました後ジンギスカンを初めて店で食べましたほら家でみんなで食べるものでしょうでもでも私一人者家族ないしそんな語り口がそうそうだろうねーか帰らないし食べる機会が減っちゃってでもやっぱり松尾ジンギスカンが好きと言うかそれ以外のジンギスカン食べたことない遠くにも家で家でしか食べない郷土料理家で鹿ってのがもう家になってますけどな感じで見るしか食べない郷土料理銘柄が家によって分かれる何かありますかっていうことですね

メール3 筆者による聞き取り (修正)

伊達：こちらはですね、神奈川県相模原市の方ですね。ラジオネームガッチュ石松（イシマチュ）さんからです。【ありがとうございます】「先日です。北海道に帰ってきました。今回は初めて旭山動物園と白い恋人パークに行きました。後ジンギスカンを初めて店で食べました。【ふーん】ほら、家でみんなで食べるものでしょう。【ンフフ】でも、でもほら、私一人じゃん。家族いないし【そんな急な語り口になるの?】（そうなってるんですよ。）実家帰らないし食べる機会が減っちゃって。でもやっぱり松尾ジンギスカンが好き。と言うかそれ以外のジンギスカン食べたことない。東北にも、え一家出しか、家でしか食べない郷土料理、「家でしか」っていうのがもう「家出」になってますけどね、漢字でね。）家でしか食べない郷土料理・銘柄が家によって分かれる何かありますか」っていうことですね。

これは、伊達を読み手であり、それがほとんど認識されている一方、【 】内の富沢のあいづちなどの反応は認識されていない。また、（ ）内は読み手の伊達が途中でメールの本文にない説明をしているが、その部分は中途半端な認識結果になっている。

メール 4 認識結果（抜粋） 10:00～11:18（1分18秒）

ホーチミンで日系の幼稚園を経営しておりますこちらに来て5年目になりました日本を感じたくて4日前に iPod を弄り回していてサンドイッチマンさんのこの番組に遭遇日本にいた時は運転中に J WAVE しか聞いてなかったのでニッポン放送は高校以来ですあんまりに面白いこととあの震災に関してのその後の情報や震災を通じて今思う事などたくさんの方の意見が胸を打ちました当然あの日の自分も思い出しました東京でしたがこちらにこちら日本と時差2時間のため私は毎日7時半に寝てしまいますベトナム時間でない夜7時半で夜中目が覚めて山道を聞くという子四日間の生活リズムを聞いてくれてるんだありがたいですね私も特派員に加わることはできますかももちろんですもちろんです今すごい勢いで発展中です便利と不便が混在中で真面目にやられてません後空いていただけるのであればまた投稿させていただきます以上ホーチミンからです

メール 4 筆者による聞き取り（修正）

富沢：ペンネーム、ピンクレモンさんですかね。「私は現在ベトナムのホーチミンで日系の幼稚園を経営しております。【ベトナムのホーチミン、ほほう】経営…【すごいな】こちらに来て5年目になりました。日本を感じたくて4日前に iPod を弄り回していてサンドイッチマンさんのこの番組に遭遇。日本にいた時は運転中に J-WAVE しか聞いてなかったのでニッポン放送は高校以来です。あんまりに面白いこととあの震災に関してのその後の情報や震災を通じて今思う事などたくさんの方の意見が胸を打ちました。【うーん】当然あの日の自分も思い出しました。東京でしたが、えーこちらにこちら日本と時差2時間のため、私は毎日7時半に寝てしまいます。ベトナム時間でね、夜7時半。で夜中目が覚めてサンドを聞くというここ四日間の生活リズムです。【へー聞いてくれてるんだ】（ありがたいですね。）【うーん】こんな私も特派員

に加わることはできますか。」もちろんですよ。【もちろんです】「ベトナムは今すごい勢いで発展、進化中です。便利と不便が混在中で真面目にやられてられません。【へー】加えていただけるのであればまた投稿させてください。以上ホーチミンからでした。」

やはり、メール本文の認識結果は良いが、聞き手のあいづちや反応はあまり出ていない。

4.2.テレビ番組

テレビ番組1は、ラジオ番組1と内容やコメンテーターが同じものを選んでみたものである。ニュースの部分とコメンテーターの解説部分（一部対話になっている）を示す。

認識結果（ニュースの部分）

台湾総統選の前哨戦となる統一地方選は24日投開票され与党の民主進歩党は同等のが常である高尾台中の二つの直轄市などで止め22のケント紙の蝶ホストの13から6に減らしてます大敗を期しました最大野党国民党は大勝し次期総統選での政権奪還へ向けて大きな足がかりを築いた形ですよ年金制度改革などに対する有権者の不満から蔡英文総統の支持率が低迷したことが響いたということです大総統は初心者の皆さんに心からお詫びするとので大敗の責任を取って気もする民進党主席を辞任すると表明しました。

筆者による聞き取り（修正）

台湾次期総統選の前哨戦となる統一地方選は24日投開票され、与党の民主進歩党は同党の牙城である高雄や、台中の2つの直轄市などで敗れ、22の県と市の首長ポストを13から6ポストに減らす大敗を喫しました。最大野党、国民党は大勝し、次期総統選での政権奪還へ向けて大きな足掛かりを築いた形です。年金制度改革などに対する有権者の不満から蔡英文総統の支持率が低迷したことが響いたということです。蔡氏は「支持者の皆さんに心からおわびする」と述べ、大敗の責任を取って兼務する民進党主席を辞任すると表明しました。

これを読んだのは、厳密にはプロのアナウンサーではないが、認識率は高いと思われる。読み取りの音声の状況によってはもう少し正確になる可能性もあるだろう。

認識結果（解説トーク、対話の部分）【抄録】 89:38～91:04（約1分半）

あのこの前私夏に台湾の国立政治大学とか淡江大学にお邪魔してですねまあいろんな方と書かれていますかで繋がりができていろんな方と今回のこの件に関してもメールとかチャットとかでやりとりしたんですけれどもこれはねその人その人の皆さん政治に詳しい人なんだけどその人その人のおっばいタッチ政治的な立ち位置によって見方が違うんですね例えばねその国立政治大学大学院んで学んでた人たちとか特に社会人なんだけれどもそのページを大学院で学んでるって言うな人なんかは今は台北で祖

国その立官庁に勤めていたりあるよその都市銀行に勤めていたりというエリートが多いんですよそういう人たちというのは政治的に来ちゃうリベラルで連載を強くしてたって人もこの人たちの言い方を借りれば蔡英文のこの2年間の改革というのは私たちが期待したものよりもはるかに生ぬるいとかこのねやっぱりその国民党時代の取得けいししゃこういう人たちの二等なれば既得権益を切ってくれるもんだと思ってたとかかなりやっちはいたんですけどもねそこがまず不満

筆者による聞き取り

あのう、この前、私夏に台湾の国立政治大学とかあの淡江大学にお邪魔してですね、まあいろんな方と関わり合いと言いますかですね、繋がりができて、まあいろんな方と今回のこの負けに関しても、ええとメールとかチャットとかでやりとりしたんですけども、これはね、その人その人の、皆さん政治に詳しい人なんだけどその人その人のやっば立ち位置、政治的な立ち位置によって見方が違うんですよ。例えばね、ええと、その国立政治大学、大学院で学んでた人たちとか特に社会人なんだけれどもその政治を大学院で学んでるっていうような人なんかは、まあ言ってみれば台北で、すごくその、えー、例えば官庁に勤めていたり、あるいはその投資銀行に勤めていたりというエリートが多いんですよ。で、そういう人たちというのは政治的に非常にリベラルで、で蔡さんを非常に強く支持をしてたって人が多いんです。【ほー】この人たちの言い方を借りれば、蔡英文のこの2年間の改革というのは私たちが期待したものよりもはるかに生ぬるい。【ああ、生ぬるい】もっと過去のね、やっぱりその、ええまあ、国民党時代の、えー、その既得権益者、【ええ】こういう人たちの不当なね、既得権益を切ってくれるもんだと思ってた。【はあ】まあ、かなりやっちはいたんですけどもね、で、そこがまず不満。

解説の部分はさすがに丁寧な修正が必要にはなってくるものだが、論旨が把握できるくらいには認識・文字化されていることが分かる。

5. まとめと今後の課題

以上、今回はケーススタディとして、googleの音声認識システムを利用したラジオやテレビのニュース解説番組の文字テキスト化(ディクテーション)について述べた。

音声認識のテキスト化については、ハード⁽¹⁶⁾とソフトなどいくつかの面で課題があるとはいえ、いわゆるニュース原稿なら90%あるいはほぼ完全に近い認識率であり、対話形式で進む解説の発話でも、かなりの認識ができていた。単語レベルなどの細かい分析までは手が回らなかったが、全体としては、例えば後で手作業を用いて直すという場合、リアルタイムでというのは難しくても、ある程度時間をかけてもいいという環境なら、テキスト化の下処理としては十分に作業がなされる⁽¹⁷⁾ことが分かった。

今後、他の話者や他のジャンル(例:ドラマ、アニメなど)においても実験を行い、どの程度の認識率を得られるか。教師としての教材づくりや学生(学習者)に作業をさせるような場合などに、どのくらい有効かということも検証してみる必要があるかと思う。

それらは次の機会を俟ちたい。

【主要参考文献】

- 河原達也 (2013) 「音声認識技術の現状と将来展望」『電気学会誌』 133-6
- 河原達也、峯松信明 (2013) 「音声情報処理技術を用いた外国語学習支援」『電子情報通信学会論文誌』 Vol.J96-D, No.7, pp.1549-1565
- 河原達也 (2014a) 「第1回 実用化進む音声認識 システムの構成要素を概観」(「音声認識・対話技術の基礎と応用」)『日経エレクトロニクス』 2014.5.26
- 河原達也 (2014b) 「第2回 音声認識に新潮流 ビッグデータや DNN を活用」(「音声認識・対話技術の基礎と応用」)『日経エレクトロニクス』 2014.6.9
- 河原達也 (2014c) 「第3回 音声認識・対話のアプリケーション 成功の鍵は必然性や自然性」(「音声認識・対話技術の基礎と応用」)『日経エレクトロニクス』 2014.6.23
- 河原達也 (2014d) 「第4回 話し言葉をテキスト化するシステム 会議録の作成や字幕付与への展開」(「音声認識・対話技術の基礎と応用」)『日経エレクトロニクス』 2014.7.7
- 河原達也 (2014e) 「第5回 音声対話システムの実際 siri はどのように成功したか」(「音声認識・対話技術の基礎と応用」)『日経エレクトロニクス』 2014.7.21
- 河原達也 (2016) 「第一章 音声認識の概要」『ITtext 音声認識システム (改訂2版)』オーム社
- 野口悠紀雄 (2016) 『話すだけで書ける究極の文章法 人工知能が助けてくれる!』講談社

【注】

- (1) 音声認識システムについての現状は、例えば河原達也 (2014a, 2014b, 2014c, 2014d, 2014e) などに解説されている。
- (2) Amazon のアレクサやgoogleアシスタントなど。
- (3) iOS の siri, docomo のしゃべってコンツェルンなど。
- (4) 厳密にはいわゆる音声出力、音声合成の機能である。例えば峯松信明氏らによる OJAD (韻律読み上げチュータズキクン) などがある。<http://www.gavo.t.u-tokyo.ac.jp/ojad/phrasing/index>
- (5) この4つの条件については河原達也 (2013) でも解説されている。
- (6) 一方、野口悠紀雄 (2016) では音声入力で当該の著書を執筆した過程など記して、音声入力の効能を説いている。
- (7) 河原達也、峯松信明 (2013) など。
- (8) 大規模コーパスなどは作成や運用に時間がかかる。リアルタイムの資料を利用したいような時は音声認識で文字化の処理が行われれば、かなりメリットがある。
- (9) 準備・選定の段階では音源によって認識が全くされなかったものもあった。
- (10) 番組公式サイト。無料ポッドキャスト Podcast あり <https://www.1242.com/cozy/>
- (11) 放送局の公式 YouTube チャンネル <https://www.youtube.com/channel/UCTLUf5>

otEetTThajFIHG_iQ/featured

- (12) 番組公式サイト。無料ポッドキャスト Podcast あり <http://www.1242.com/sand/>
- (13) <https://dhctv.jp/>
- (14) 有料での登録が必要。YouTube では最新 2 週間分のみ。
- (15) ラジオ辯組 1 と同様、2018年11月24日に行われた台湾の統一地方選挙を取り上げたものを採用した。先にラジオ番組を決めた結果、テレビ番組も同種の内容を選ぶことにしたという次第である。どちらも、番組としての展開はニュース（原稿）を読み上げたのちにコメンテーターの解説と司会との対話という性質のものである。
- (16) マイクの指向性（無指向性、単一指向性、双指向性など）によっても結果が変わってくるかどうか。
- (17) また、フィルターのようなものをどこまで反映させるかによっても作業の労力が変わってくる。

《付記》

本論文は『『A I と日本語教育』国際シンポジウム』（台湾・淡江大學、2019年 3 月 9 日）と「第26回国學院大學日本語教育研究会」（國學院大學、2019年 7 月27日）で発表したものに加筆修正したものである。論文掲載にあたり、発表の場でご意見等をくださった方々、『語文論叢』の編集委員の方々に感謝申し上げます。

（かんさく・しんいち 南臺科技大學應用日語系助理教授）