

## 第3章

# AIを用いた音声認識アプリの機能と未来像 —インクルーシブな社会の実現のために—

Shamrock Records株式会社代表取締役 青木 秀仁

## 1 AIを用いた音声認識アプリ 「UDトーク」について

このアプリはAIを使用したクラウド型音声認識サーバーを経由して、人が発した音声を自動で文字にするものである。1対1の会話から、多人数のミーティング・会議・講義など様々なコミュニケーションに活用出来るほか、議事録作成にも用いることも出来る。

iPhone、iPad等を使い音声認識をするため、高価な機器や複雑な設定などは一切不要であり、スマホでアプリをダウンロードすることで直ぐに利用が可能になる。また、講演会場などでは、ヒアリング・ループや音響設備とつなぎ、音声を入力することで図3-1のようにリアルタイムで字幕を表示することも出来る。

多言語の翻訳および音声認識や音声合成を実装することで、障がい者バリアフリーだけではなく言語バリアフリーに、漢字かな変換機能を使うことで世代間バリアフリーに、多目的に使うことのできるコミュニケーションアプリである。

講演会などではスクリーンに2言語同時に表示することも出来、また手持ちのスマートフォンやタブレットで自分の言語を設定し見ることも出来る（図

図3-1 研究会での講演



出典：筆者撮影

図3-2 UDトークの利用画面

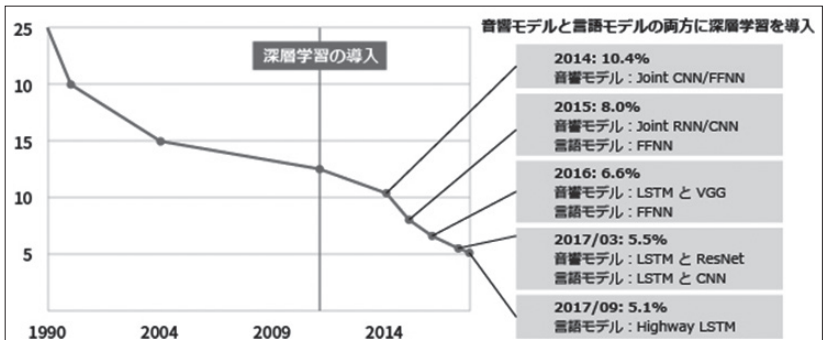


出典：Shamrock Records株式会社

3-2)。聴覚障がい者対応と外国人対応を区別せず一緒に「ユニバーサル対応」として行うことが可能となるものである。

なお、音声認識自体は特段真新しい技術ではなく、古くから研究が進められてきた分野であるが、どうしても認識率が向上しなかったことから、10年程前に急激に開発が廃れた歴史がある。しかしながら、ディープラーニングのおかげで、近年認識率が急激に向上し実用化されるに至っている（図3-3）。

図3-3 音声認識率の推移



出典：IBM proVISION 92号 p 48

## 2 UDトーク利用状況

第三次AIブームが到来した昨今、官民間問わずAI利用についての報道は盛んになされているが、ふたを開けてみるとAIを用いたシステムは多くが現状実験段階に留まっており、実用段階に達していないものが多い。そんな中、UDトークはすでに製品として流通しており、ビジネスとして成立している段階にある。事実、既に多くの自治体や大学等で活用がされており、また年々利用者数も増加している。現在、日本の名だたる企業がUDトークを使っており、大学だけでも約100の大学が聴覚障がい者の学生のために導入している。2018年4月時点で日本の大学は全国で768<sup>1</sup>あり、その中で聴覚障がい者の学生を受け入れている大学は400～500程度であると推測されている。以上のことから、聴覚障がい者を受け入れている大学におけるUDトークの導入率は少なくとも20%程度であり、かなり高いシェアを占めていることになる。

導入件数について、弊社で詳細な調査は実施していないが、概ね400件を超えていると推計している<sup>2</sup>。

続いてダウンロード数についてであるが、こちらについても詳細な調査は行っていないが、だいたい20万ダウンロード程度ではないかと推測している<sup>3</sup>。また、このうち8割以上の利用者がリピートして利用していることがわかっている。このようなアプリの多くは、一度利用しただけで消されてしまうものがほとんどであるが、その中でもUDトークは高いリピート率を誇っていると言える。このリピート率はLINEやFacebook等の有名アプリとほぼ同水準である。

---

1 教育情報センター2018年7月号（旺文社）

2 2018年7月31日時点。

3 2018年7月31日時点。

### 3 UDトークの様々な機能

#### (1) 音声認識と音声合成でバリアフリーコミュニケーション

UDトークは、人間が発した音声で文字を表示するものであることから、基本的には聴覚障がい者の利用を念頭においているが、機能はそれだけにとどまらない。

他の1つ目の機能として、音声の読み上げ機能がある。ボタンをタップするだけで字幕を読み上げてくれる機能が付いており、目の不自由な方でもコミュニ

ケーションが可能になるように配慮されている。筆談機能も備え付けられている(図3-4)ほか、キーボードで入力した内容を読み上げることも可能である。発音障がいがある方や、「ろう」の方で喋るのが苦手な方等がUDトークを自分の声の代わりに活用することが出来るのである。そうすると聞こえない方、喋れない方と目が見えない方が、自分たちだけでコミュニケーションを取ることが出来る。要するに、これらの機能を駆使すれば様々なハンデを持った人が、他人の助けを借りなくとも、自らコミュニケーションをとることが可能となるのである。

これはまさにイノベーションの賜物であり、今までできなかったことがテクノロジーで可能になったということである。

図3-4 筆談機能



出典：Shamrock Records株式会社

#### (2) 多言語翻訳と音声認識でグローバルコミュニケーション

UDトークは障がい者のみならず、外国人の方とコミュニケーションも可能とするものである。というのも、アプリには数多くの外国語が実装されており、

瞬時に翻訳を行ってくれるのである（図3-5）。それぞれの翻訳の精度もかなり良く、一昔前までの技術ではかなり不自然な翻訳しか行えなかったが、近年のディープラーニングの技術により、既に説明した音声認識の技術のみならず、翻訳技術も格段に向上し実用に耐えうるレベルに達したのである。

この翻訳機能を使えば、外国へ旅行に行った際の翻訳はもちろんの事、講演等でも利用可能な精度での翻訳を行ってくれる。

ところで、実はUDトークには大阪弁への翻訳機能が実装されている（図3-6）。今まで説明した各種機能については、順次「〇〇の用途で使います」と説明してきたが、この大阪弁への翻訳機能に限って言えば、開発者自身も特段の用途を目的として開発したわけではなく、利用者から「何に使うか？」との問い合わせもしばしばある状況である。ただ、とある法人ユーザーの方から聞いた話によれば、「つまらない会議は大阪弁にして乗り切っている」とのことであった。

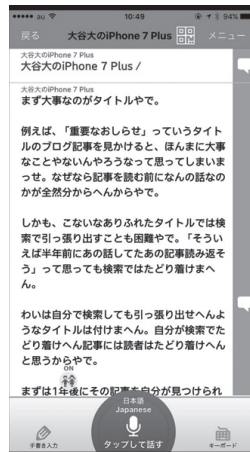
遊び感覚で作った機能ではあったが、そのユーザーに非常に喜んで頂いたことから、このようなアプリ作成においては、「福祉のために」とか「困った人を

図3-5 多言語翻訳



出典：Shamrock Records株式会社

図3-6 大阪弁への対応



出典：Shamrock Records株式会社

助けるために」などというような聖人君子である必要は全くなく、何か面白いと思った機能を適宜作って実装していけば、ユーザーの方で用途を見つけて使ってくれる場合があるということを感じた。また、別に福祉系のサービスだからといって、変に身構える必要はなく、普通の人でも使いやすいとか、面白いとか、まさにユーザー本人が他の人に紹介をしたくなるようなものをつくれれば必然的に広がっていくのではないかと感じている。

### (3) ひらがな変換で世代間コミュニケーション

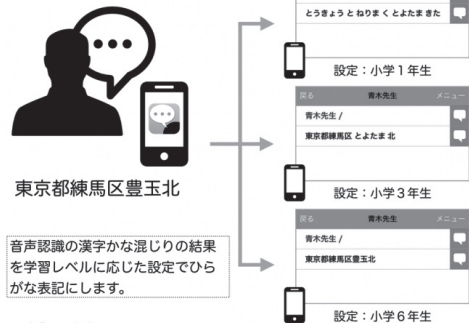
続いての機能は主に子どもとのコミュニケーションを想定した機能である。

漢字変換を、設定した学年（小学1年生～小学6年生まで）にあわせて行ってくれるものであり、つまり、初等教育の学習レベルに応じた漢字変換を行うものである。例えば1年生に設定すると、ほとんどがひらがなで表記されるが、小学3年生に設定すると一部、そこまで習熟が予定されている漢字については変換される。さらに小学6年生に設定すると同様の処理が実施される。図3-7のように、小学3年生に設定すれば一部が漢字変換され、

図3-7 ひらがな変換の画面



漢字かな変換機能イメージ図



©Shamrock Records, Inc.

出典：Shamrock Records株式会社

小学6年生に設定すると全て漢字に変換されていることがわかるだろう。例えば、耳の聞こえない小学校1年生、3年生、6年生に対して同時にコミュニケーションをとる場合において、1回の動作でそれぞれに学習状況に適した漢字変換をした内容を送ることも可能である。

また、漢字の上にフリガナを付けることが可能であり、漢字学習のツールとしての機能も果たすことが出来るのである。

このひらがなで表記する機能は、実は耳が聞こえない子供だけではなく、外国人向けに使われていることもある。日本語が出来る外国人であっても、漢字が読めない場合についてはひらがな変換を利用しているケースもあるようである。また、先ほどの事例と類似するが、漢字を勉強している外国人についても漢字を学習状況に合わせてひらがなに変換する機能が用いられているとのことである。

#### (4) ウェアラブルデバイスとの連携

専用のゴーグルを着用しシアターモードにすることでARのように視界の中に字幕を映し出すこともできる(図3-8)。字幕であると、喋っている人と字幕の画面を行ったり来たりする必要があり、この課題を解決するために活用さ

図3-8 カメラとシアターモード



出典：Shamrock Records株式会社



れているのがこのモードである。

この機能は、大学の講義を聴きながらノートをとる際に重宝されている。というのも、先生が喋っている時には字幕を見る必要があることから、プリントを見る、またはプリントに書き込むといった動作が同時に行えないことが問題となっていた。

ところがこの機能を使えばカメラ越しにプリントを見ながら書き込みをできるようになり利用者からの評判は大変良い。

実は先ほどの大阪弁への変換機能と同じでこれについても作成者自身このような利用方法を想定していたわけではなく、機能を実装した後で利用者の知恵から生み出された用途である。

## 4 自治体におけるUDトーク活用事例

以上、UDトークについて様々な機能を照会してきたが、最後にUDトークが導入された神奈川県川崎市及び三重県伊勢市の事例を紹介する。

### (1) 神奈川県川崎市議会の事例

川崎市議会では、聴覚障がい者や耳の遠い方々でも、気軽に議会の傍聴が出来るようすること等を目的としてUDトークの導入を決定した。2019年2月の議会から運用を開始している。同様の取組みについて、国内では福井県議会が発言内容を文字化している事例があるが、これについては人間が発言を聴きながら手入力して文字を出力しており、AIによる市議会における議員の発言内容の字幕表示は全国の初の取組みである。議会への導入については、他にも導入を検討している自治体が複数あり、今後は従来の手書きや人が行っている作業の転換が進んでいくことが予想される。アプリ作成者としても良い運用を期待しているところである。

## (2) 三重県伊勢市の導入事例<sup>4</sup>

伊勢市では、聴覚障がいのある人などのコミュニケーション支援の充実を目指し、2018年9月からUDトークを導入している。

来庁者の窓口対応について、今までは、手話の出来ない聴覚障がいのある人や耳の聞こえにくい人には、筆談等で対応していたが、窓口での事務手続きの説明などでは、筆談では時間がかかり、説明が伝わりにくいことがあった。また、外国人の窓口対応においては、英語や中国語などの多言語に堪能な職員が少なく、来庁者に対して丁寧な意思疎通が図れないといった課題があった。

しかし、UDトークの導入以降は、会話内容をタブレット端末にリアルタイムに表示できることから、聴覚障がいのある人や外国人の多言語翻訳等において、お互いにストレスを感じることなく、スムーズな窓口対応ができるようになった。UDトークを通して行われた会話内容のログは保存できることから、今後は、要約筆記に代わる講演会や議事録作成など、多方面での利用を検討している。

図3-9 伊勢市UDトーク運用事例



出典：Shamrock Records株式会社

4 UDトーク運用レポート：三重県伊勢市を引用。但し書式については一部訂正を加えている。<http://udtalk.jp/category/news-release/>

## 5 インクルーシブな社会の実現のために

インクルーシブ (inclusive) とは「包括的な」「包み込む」と言った意味である。社会的弱者やマイノリティとして捉えられている人を含めて、あらゆる人々をどのように社会に包括していくかについては、過去幾度となく議論されてきたテーマである。

インクルーシブな社会の実現に向けた課題は、ざっくり言ってコミュニケーションの課題であると考えている。そして、このコミュニケーションの課題に関しては、近年の科学技術の進歩に伴い、多くはテクノロジーで解決できる段階に達していると感じている。例えば先程から説明してきたように、UDトークを使って字幕を出しておけばよい。

要するに、スマートフォンを使うことにより、個別への対応が可能なのである。1人1人に対して補助が出来ればそれはそれで問題ないことかもしれないが、そのような対応は人員的にも困難であるから、テクノロジーを使って解決していくことが望ましい。

ところで、聴覚障がい者の方が講演会などのイベントに参加する場合、主催者が聴覚障がい者専用の席を特別に用意するケースがある。しかし、誰も「障がい者専用」など大きく掲げてある席に座りたいとは思わないであろう。他にも、手話通訳を用意している場合もあるが、実は聴覚障がい者のなかで手話を理解できる人の割合は非常に低いと言われている<sup>5</sup>。従って、手話通訳があるからと言って、聴覚障がい者の方がそのようなイベントに参加しやすくなるかと言えば必ずしもそのようなことはない。また、参加するほうも特別な配慮をしてもらうことで気を使わせてしまうことが心理的な負担になることは容易に想像できるだろう。

では、それらの課題を解決するにはどうすれば良いか。一番の解決策は、主催者が字幕を備え付けることではないだろうか。字幕を必要とする人が参加す

5 1996年に実施された厚生労働省身体障害者・児実態調査によれば、聴覚障がい者で「手話ができる」と答えたものは14.1%：厚生労働省HP ([https://www.mhlw.go.jp/www1/toukei/h8sinsyou\\_9/1-1.html](https://www.mhlw.go.jp/www1/toukei/h8sinsyou_9/1-1.html))

るか否かに関わらず、設備として様々な配慮や対応があるというのがあるべき姿である。そのためには、UDトークを利用し字幕をスクリーンで用意する、たったこれだけでよいのである。

そうすれば、参加者もわざわざ事前申請する必要もなく、また主催者側も特別な席を用意する必要もなくなるのである。先に述べた通り、スマホなどちょっとした用意があれば、そしてちょっとしたやる気さえあれば実現できることである。

是非そのような対応が当たり前になって欲しいと願うばかりである。



## おわりに



最後に「UDトーク」の今後の展望についてであるが、単刀直入に言って「UDトーク」が役目を終える日が来ることを願っている。

なぜなら「UDトーク」が役目を終えるとき、それは、少なくとも障がい者と健常者がコミュニケーションをするにあたって大きな不具合がなくなり、会話のバリアフリーが実現している世の中になっていることを意味するからである。そうになったら少し寂しい気もするが、その瞬間を早く見たいと考えている。

※本稿は2018年7月31日に実施された、第2回都市自治体における人工知能の利活用のあり方に関する研究会で講演頂いた内容に、事務局で加筆・修正を加えて作成したものである。