

音声認識アプリを活用した支援システムの構築に関する検討 — 少数討論型授業を事例に —

松崎 丈

宮城教育大学 特別支援教育講座

宮城教育大学では、2014年にサーバー型音声認識を実装したアプリ「UDトーク」を活用した支援システムの導入を始めた。本年度は、聴覚障害のある大学院生が受講する少数討論型の授業に導入した。本研究では、UDトークの支援システムの運用に関する観点を探索的に検討するために、大学院修士課程における授業3科目での導入事例を対象に質問紙調査を行い、KJ法で分類した。結果として、主に少数討論型授業におけるUDトークを活用した支援システムを構築するために、教員等の音声入力技術や専門修正者の修正技術の育成やネットワーク環境の整備等の観点が必要であることが明らかになった。その上で、今後UDトークを活用した支援システムの構築に向けて取り組むべき事柄について考察した。

キーワード: 聴覚障害学生支援、音声認識アプリ、UDトーク、少数討論、支援システム

1. はじめに

近年、全国各地の大学で音声認識技術を導入した聴覚障害学生支援の試みが行われ始めている。大学における音声認識技術を活用した支援に関する先行研究を概観すると、スタンドアロン型の音声認識を活用した支援システムと、サーバー型の音声認識を実装したアプリを使った支援システムに二分できると思われる。

前者は、2000年代に構築された支援システムである。他の機器に依存せず単独で動作するスタンドアロン型音声認識を利用しており、音声入力の相当なトレーニングを受けた者を活用することで十分な認識精度を確保する。つまり、教員等の音声を直接入力するのではなく、特定のトレーニングを受けた者が復唱すること[1]、かつ修正も行うこと[2]で音声認識率を高く維持することが標準的な手法として採用されている[3-6]。また、聴覚障害のある学生の立場を考慮した情報保障の質的向上のために、誤認識の種類やその対処の検討[7・8]、支援者が円滑に作業を実施するための支援技術の検討[2]、支援者の復唱及び修正の方法に関する検討[6]、字幕の提示方法[9・10]などの検討がなされてきた。そうしたシステムの運用の質を維持するために復唱者や修正者を少なくとも各2名体制で確保したり、PC

やネットワーク機器等を現場でセッティングすることが必要になる[11]。そのため、人数や時間等の制約が伴いがちな聴覚障害学生支援でどのように利用していくかが課題になっていたともいえる。

後者のシステムは、2010年代から登場したサーバー型音声認識を実装したアプリを活用したものである。この型の音声認識は端末スペックに依存しないため、スマートフォンやタブレットでも利用可能であり、クラウド環境を用いるために高速かつ高精度の音声認識が実現可能となる。不特定の話者の音声にも対応でき、特定の復唱トレーニングも前者ほど必要ではなくなってきた。聴覚障害分野では、この音声認識を実装したアプリ「UDトーク[12]」が開発された。このアプリには、QRコードによる端末間の容易な接続やWebによるリアルタイム単語登録などの機能も追加されている。大学においては認識精度の高さや様々な機能の追加により、前述した課題を解消できることが期待され、実際にUDトークを活用した支援の報告が出てきている[13・14]。後者の支援システムは、前者とは異なり、教員や他受講生も直接音声入力を担うことになる。しかしながら、現時点で教員や他受講生も含めてUDトークを活用した支援システムをどのような観点で構築するのかについては不明な点が多く、大学における活用事例の収集とともに検討を進める必要がある。

宮城教育大学では、UD トークを 2014 年 11 月にトライアル導入し、2015 年 1 月から教育機関プランで法人契約して利用を開始した。聴覚障害学生支援では、2017 年度から宮城教育大学大学院修士課程及び教職大学院で聴覚障害のある学生が受講する授業で先行的に実施している。その授業では少人数での討論によるものが多く、教員や受講生の音声を直接入力し、修正は修正者や、必要に応じて他の受講生も担っている。

そこで本研究では、この少人数討論型の授業事例を対象とし、その支援システムの運用をどのように行っているのかについて探索的に調査することで、主に少人数討論型授業における UD トークを活用した支援システムの構築にむけた基礎資料を得ることを目的とする。

2. 調査方法

2.1 調査対象

2017 年度に UD トークによる支援を実施した大学院修士課程の授業 3 科目(各 15 コマ)の事例を対象とする。具体的には、各授業の担当教員、聴覚障害学生、健聴学生、修正者の 4 者及び全授業への修正者の派遣や備品の手配等に関わったコーディネーターを対象に質問紙調査を行った。

なお、UD トークによる支援は次のように行った。機材は、宮城教育大学しょうがい学生支援室の備品で、UD トークをインストールしたスマートフォン (iPhone 6) を 1 台、音声入力用のハンドマイク (iRig Voice) を 1 本、編集入力ソフトウェア「UD トーク for Windows」をインストールしたノート PC (Microsoft Surface Pro4 128GB) を 1 台、認識結果表示用のタブレット (iPad 2 Wi-Fi 64GB) を 1 台利用した。無線のネットワークを経由して各端末で認識結果を共有し、ノート PC で編集の上、タブレットに表示した。修正者は授業 1 コマに対して 1 名派遣しており、手書き及びパソコンノートテ

クの経験がある者とした。

本調査の対象者は、授業 3 科目の担当教員 3 名 (T1、T2、T3)、聴覚障害のある学生 1 名 (DS)、健聴学生 2 名 (HS1、HS2)、修正者 2 名 (S1、S2)、コーディネーター 3 名であった。DS の情報保障利用経験は、手書きノートテイクとパソコンノートテイクの両方とも 4 年 9 ヶ月であった。修正者の S1 は手書きノートテイクの実践経験年数が 3 年 8 ヶ月、パソコンノートテイクが約 1 年 9 ヶ月であり、S2 は手書きノートテイクが 1 年、パソコンノートテイクが 2 年 10 ヶ月であった。表 1 に、各授業科目と対象者との関係を示す。

表1 各授業科目と対象者との関係

対象者	授業 A	授業 B	授業 C
担当教員 3 名	T1	T2	T3
聴覚障害学生 1 名	DS		
健聴学生 2 名	HS1、HS2		
修正者 2 名	S1・S2*	S2	S1・S2

*実際はいずれか 1 名のみ派遣。

2.2 調査の手続きと分析方法

質問紙調査は、自由記述による調査用紙を作成の上、メールに添付して各対象者に送付し、返信してもらおうという流れで行った。

質問内容は、スタンドアロン型音声認識に関する先行研究で検討されてきた事柄を踏まえて、①音声認識技術の利用について、②音声入力及び復唱について、③修正について、④認識結果について、⑤その他 (UD トークの性能や活用方法など) の 5 項目を作成した。なお、1 回目の回答に対し、さらに詳しく質問もした。

こうして得られた対象者全員分の記述内容は、KJ 法[15]に基づいてカードに記述し、本研究のリサーチクエスションと照らし合わせながらカテゴリ化を行った。次に、カテゴリ間関係を分析し、ストーリーラインを作成した。

3. 結果と考察

対象者全員分の記述内容から、226 個のラベルが抽出された。そのラベルの内容の類似性から、60 の小カテゴリーが設定された。さらにその小カテゴリー一見出しの類似性から、21 の中カテゴリー、さらに6 の大カテゴリーが設定された(表 2)。6 の大カテゴリーは、【UD トークに対する認識】【音声入力及び修正の方略の形成】【UD トークの性能を活かす

環境整備】【専門支援者の活用】【少人数討論での活用の効果と課題】【より効果的な運用のために必要な機能】である。

以下、6 の大カテゴリーと 21 の中カテゴリーを用いてストーリーラインを述べる。

大カテゴリー及び中カテゴリー間の関係を、図 1 に示す。

【UD トークに対する認識】として、対象者は、UD

表2 少人数討論型授業におけるUDトークを活用した支援システムに関するカテゴリー

大カテゴリー	中カテゴリー	小カテゴリー	ラベル数	
UDトークに対する認識	現行の音声認識技術の高さや利用価値を実感	UDトークの音声認識率が非常に高く、リアルタイムで速く知ることができる	2	
		話者が注意深く使用すれば従来の文字通訳(要約型)よりも伝わる情報が増加	2	
	教育機関における情報保障ニーズとの整合	タブレットやスマートフォンなど持ち運びやすくフリック入力可能な端末の活用	13	
		QRコードで端末同士の接続が容易にできる	3	
		修正用ソフト(UDトーク for Windows)が使いやすい	3	
		認識結果の保存・送信機能が容易にできる	3	
	音声入力に伴う様々な制約への気づき	複数の話者が同時に発言することの困難	6	
		普段意識しない発音への疲労感と困難感	4	
		視線や身振りなど見える情報の減少	4	
			議論の盛り上がりへの対応の限界	3
音声入力及び修正の方略の形成	音声入力の質の維持方略	適度な速さで句読点で句切って明確に発音する	17	
		マイクの向きや距離への注意の維持	5	
		話すタイミングでボタンを押して回りの雑音を入れないようにする	4	
	修正方略の形成	専門用語や外来語、授業のキーワードは優先して修正	9	
		繰り返し誤認識が発生する用語、大体意味がわかる用語やフィルターの修正は後にする	6	
		発言内容と認識結果が大幅に異なる場合はキーボード直接入力への切り換え	2	
		利用者との間で微妙な修正対象への対処をその都度決定	2	
		10人以下の少人数かつターンテークが明確なディスカッション	10	
UDトークの性能を活かす環境整備	利用する授業の種類・条件	認識率が高い人の一方的な講義	6	
	認識結果の表示に影響するネット環境の整備	支援が難しい場面(映像視聴等の講義や支援者の人数不足・活動時間外等)	2	
		変換から確定までの時間の短縮化による修正作業の円滑化	6	
		テザリング機能の状態による認識結果の表示速度の変動	5	
専門修正者の活用	修正スキルの基本条件	内容理解の有無に関係なく少し遅れて正確に再現する技術の習得	6	
		タッチタイピング技術は最低必要	5	
		誤認識が多発した発話に対するパソコンノートテイクへの切り換え	3	
	修正者2名ペアでの配置	単語登録スキルの習得	2	
		多発した誤認識群から修正対象を検出する精度を維持	4	
	負担軽減の配慮の必要性	誤認識の多発に対して話した内容の記憶量に基づいて修正できる範囲の拡大	2	
誤認識が増大した時のみ修正者1名への負担が倍増。		8		
		スクリーン表示や多くの利用者がある時は修正者への精神的な負担が増大	2	
少人数討論での活用の効果と課題	受講生の発言行動への教育的効果	発言内容を筋道立てたものに編集	8	
		感情的(抑揚のある)かつ不要な発言の減少	2	
	聴覚障害学生の発信行動への教育的効果	音声認識アプリや端末の扱いに精通	3	
		自身の発言内容を効果的に発信するICTスキルの習得	3	
		遠慮せずにニーズを発信する技術の習得	2	
	参加者全員に及ぼす影響の発現	文脈から発言の意図を推理する力の習得	2	
		誤認識や接続不安定で発言者の思考や議論の流れが妨げられる	3	
全員で利用するための取組の必要性	アイコンタクトや顔を見て会話する機会が減少する	2		
	UDトーク利用意図への共通理解及び使用方法の熟知	7		
	使用者全員でシステムの構築、整備と片付けを実施	4		
	発言者が自身の認識結果を復唱・修正を行う	4		
		スクリーン表示で認識結果を共有し、理解できなかったと言いつい易い雰囲気づくり	4	
より効果的な運用のために必要な機能	音声認識技術の向上	音声認識率の向上	5	
		複数の話者が同時に発言する機能	4	
		雑音環境下でも認識できる機能	2	
	修正や認識結果の表示等のインターフェース機能の新たな追加	認識結果の正誤を確認してから表示できる機能	6	
		複数名で共同で修正できる機能	4	
	音声や点字への変換機能の追加	修正した結果の文字色や認識率が低い結果の発生を色で知らせる機能	3	
		より正確な読み上げ機能(入力文ごとに分けて自分のタイミングで読み上げる)	5	
		点字表示もできる機能	2	
		ワイヤレス周辺機器の活用	Bluetoothのキーボードやマイクの利用	2

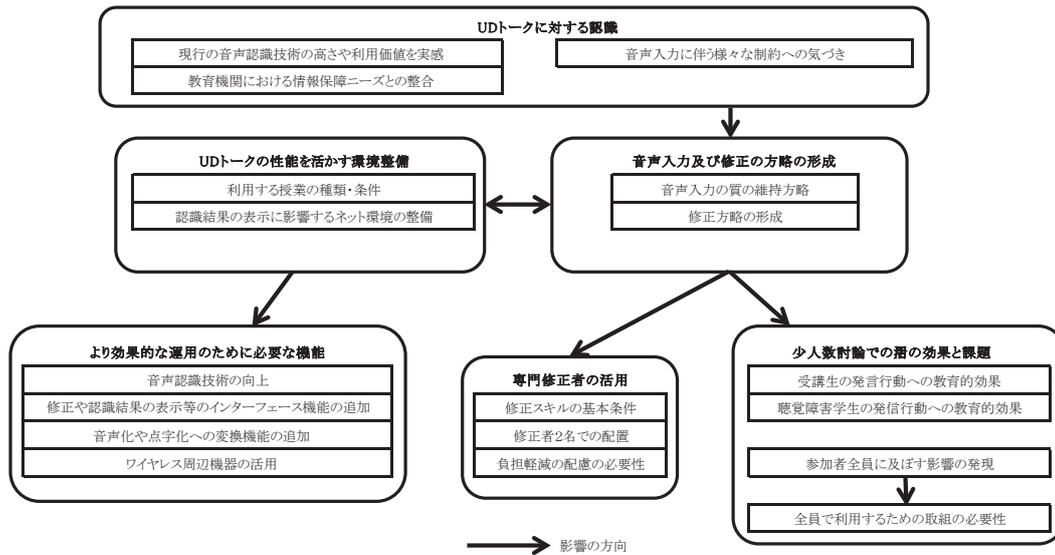


図1 カテゴリー間の関係を示した概念図

トークの音声認識率が非常に高いことや話者の発話内容を全文表示、しかもタイムラグを感じさせないほどリアルタイムで速く知ることができるために〈現行の音声認識技術の高さや利用価値を実感〉するとともに、授業の合間の休憩時間(宮城教育大学では10分)にタブレットやスマートフォンなど持ち運びができる端末同士を接続してシステムを容易にセッティングできるといった〈教育機関における情報保障ニーズとの整合〉もできていると認識していた。その一方で、複数の話者が同時発話できないことや普段意識しない発音を求められるために不慣れから疲労感や困難感を感じるなど音声認識技術は十分に完成されていないことに気づくといった〈音声入力に伴う様々な制約への気づき〉もあった。

その上で、教員や受講生が発言した内容を適切な認識結果として表示するために、【音声入力及び修正の方略の形成】を図っている。〈音声入力の質の維持方略〉では、小カテゴリーにあるように認識しやすい話し方(例えば、早口ではなく適度な速さで明確に発音する、句読点や文の切れ目にあたる位置での小さな間を空けるなど)、マイクの向きや距離への注意、音声入力をOFFにしておくことなどが挙げられた。なお、復唱に関する取組の記述は皆無

であった。〈修正方略の形成〉では、修正対象の優先順位の設定、誤認識の多発には一つひとつ修正せずパソコンノートテイクで発言内容を要約するなどが挙げられた。これと併せて、【UDトークの性能を活かす環境整備】も同時に重要であることが確認された。平日の日中は学内の無線LANネットワークが不安定になりやすいため、しょうがい学生支援室が複数台購入したiPhone 6のテザリング機能を使って無線のネットワーク環境を立ち上げるといった〈認識結果の表示に影響するネット環境の整備〉をしている。それでもネットワークが不安定になり、認識作業が止まったり認識結果の表示が遅くなったりしたという。また、本年度より本格的に聴覚障害学生が受講する授業に音声認識技術を活用した支援をした経験から、音声認識率が高い教員による一方的な講義や10人以下の少人数による討論といった〈利用する授業の種類・条件〉を考慮して導入を進めていくことが望ましいだろうとのことである。

本研究で対象となった教員3名の授業は少人数討論型の授業であり、他の受講生HS1, HS2もUDトークを利用して発言した。その結果、【少人数討論での活用の効果と課題】で教員が、〈受講生の発言行動への教育的効果〉と〈聴覚障害学生の発言行動への教育的効果〉を指摘している。例えば、

音声入力する受講生は音声認識率を維持するために内容を筋道立てて発言することや、聴覚障害学生は認識結果から発言の意図を捉えたり自身の文字による発言を工夫して発信する力を高める機会になる。しかし一方で、全員がUDトークを使って討論している時に誤認識や接続不安定で討論の流れが止まるほどの〈参加者全員に及ぼす影響の発現〉が見られているため、全員でUDトークのシステムを上手に運用できるような〈全員で利用するための取組の必要性〉があるといえる。

それから、学術用語や専門的な説明が飛び交う授業では、授業に参加している当事者のみの音声入力や修正では限界があり、【専門修正者の活用】に取り組んでいる。専門修正者には、①タッチタイピング技術、②単語登録スキル、③誤認識の多発に対するパソコンノートテイクへの切り換え、④話された発話内容を少し遅れても再現できる技術の習得が必要であり、これらが〈修正スキルの基本条件〉となるであろう。特に④の技術は、誤認識、誤変換や抜けている箇所を検出して修正・補填する点で、発話内容を自分で整文するパソコンノートテイクとは異なる技術であるといえる。また、90分間と長時間で修正作業の質を維持するために現行の修正者1名体制よりも〈修正者2名ペアでの配置〉とすることと、〈負担軽減の配慮の必要性〉を考えた取組が必要であると考えられる。

さらに、【より効果的な運用のために必要な機能】に関しても、複数の話者の発言を同時入力したり雑音環境下でも認識できる〈音声認識技術の向上〉、複数名で修正の共同作業を行ったり不適切な誤認識等を修正してから初めて表示する機能など〈修正や認識結果の表示等のインターフェース機能の新たな追加〉、様々な信号で出力できる〈音声や点字への変換機能の追加〉、Bluetoothのマイクやキーボードで円滑に利用できる〈ワイヤレス周辺機器の活用〉が必要であるということが明らかになった。

4. 総合考察

本研究では、少人数討論型授業で実施したUDトークの支援システムの運用について探索的に調査した。その結果、スタンドアロン型の音声認識を利用した支援システムでは、教員等の音声を直接入力せず、特定のトレーニングを受けた者が復唱・修正を行うことが標準的な手法となっているが、UDトークのようなサーバー型音声認識のアプリに関しては、むしろ教員等が音声を直接入力し、専門修正者が修正することが標準的な手法となることが示唆された。その場合、教員や受講生は一定の質で音声入力するための方略を身につけるとともに、専門修正者は、タッチタイピング技術等の習得だけでなく、デカラージシャドーイングのように音声を聞いた後、遅れて復唱する技術トレーニングを受けておくことが必要となる。また、サーバー型音声認識はネットワーク環境に依存しているため、長時間でのネットワーク環境(無線)の安定化をどうするか検討する必要がある。また、授業や大学行事等の使用場所によっては、スタンドアロン型の音声認識を活用した支援システムと使い分けることも考慮する必要があると思われる。以上から、UDトークを活用した支援の質的向上や問題の改善につなげる手引きの開発や研修会の開催に取り組むことが課題になると考えられる。

また、聴覚障害学生や他の受講生への教育的効果も得られることから、UDトークを授業に導入する際は、そうした教育的意義を全員で共通確認するとともに、UDトークを十分に活かせるように従来の音声会話における調整行動を共有するための教育を実施する必要もあるといえる。例えば、他の受講生に対しては、前述の音声入力の方法だけでなく内容を筋道立てて発言したり同時発話が起きないようにしたりする意識を高めさせるとともに、聴覚障害学生には、教員や他受講生にUDトークの支援に関する自身のニーズを発信したり自身の発言を

効果的に行える方略を検討してもらうなどが挙げられる。

さらに、UDトークをより効果的に運用するために必要な機能が複数挙げられたが、2017年1月時点で、①複数名での修正作業、②修正箇所の色を別の色で表示、③編集した結果のみ表示、④音声や点字での出力はすでに実現可能である。一方、同時発話や雑音下での音声入力に関する音声認識技術の向上については、現時点では活用する側の音声入力方略やワイヤレス周辺機器の活用も含めた環境整備で補う必要があると思われる。

以上から、主に少人数討論型授業においてUDトークを活用した支援システムをどのように構築するかについて多角的な視点で基礎資料を得ることができた。今後は、討論型以外の授業へのUDトークの活用事例も収集し、聴覚障害学生だけでなく教員や他受講生も加わって全員で授業のユニバーサルデザインを実践できるように引き続き研究を進めていきたい。

5. 引用文献

- [1] 勝丸徳浩他: 講義音声認識に基づくノートテイクシステム, 電子情報通信学会技術研究報告, vol. 109(259), pp.25-30(2009).
- [2] 三好茂樹他: 音声認識技術を利用した字幕作成担当者のための支援技術とそのシステム開発, 筑波技術大学テクノレポート, vol. 14, pp.145-152 (2007).
- [3] 金澤貴之他: ICT を活用した聴覚障害学生支援—キャンパス間連携入力と音声同時字幕システムの活用事例から—. 群馬大学教育実践研究, 26, pp.107-117(2009).
- [4] 松崎丈・藤島省太: 聴覚障害学生支援における音声認識を活用した通訳システムの構築—利用者の観点に基づいた字幕呈示の検討—. 宮城教育大学紀要, pp.191-203(2008).
- [5] 松崎丈他: 音声認識技術を活用した聴覚障害

学生支援:教室内外及び遠隔地における通訳システムの構築. 宮城教育大学情報処理センター年報, 16, A1-A6(2009).

[6] 三好茂樹他: 音声認識字幕における円滑な連携作業を実現するためのソフトウェア開発と情報保障者の技能, ヒューマンインタフェース学会研究報告集, 11(2), pp.171-178(2009).

[7] 牧原功他: 音声認識技術による字幕運用の課題—音声言語を文字化することの問題. 群馬大学留学生センター論集, 7, pp.33-50(2008).

[8] 加藤伸子他: 聴覚障害学生のための音声認識字幕における誤認識の与える影響についての検討, ヒューマンインタフェース学会研究報告集, vol. 17(12), pp.75-78 (2015).

[9] 入倉啓輔他: 音声認識技術を利用した聴覚障害学生のための講義保障システム, 社団法人映像情報メディア学会技術報告, vol. 31 (14), pp.29-32 (2007).

[10] 磯野春雄・山口淳平: 音声認識技術を利用した聴覚障害学生のための字幕講義システム, 日本人間工学会大会講演集, vol. 45, pp.204-205 (2009).

[11] 三好茂樹他: 音声認識によるリアルタイム字幕作成システム構築マニュアル, 日本聴覚障害学生高等教育支援ネットワーク(2009).

[12] UDトーク: <http://udtalk.jp/>

[13] 皆川雅章: ICT を用いた情報保障の取組み—学生・教職員協働による実践事例—, コンピュータ&エデュケーション, vol. 40, pp.26-31 (2016).

[14] 松崎丈: 音声認識技術を活用した通訳システムの検討, しょうがい学生支援室 平成 26(2014)年度年次報告書, pp. 8-10 (2016).

[15] 川喜田二郎: 発想法—創造性開発のために, 中公新書(1967).