

音声認識アプリを活用したシステムの設定方法とその留意点の検討 — 少人数におけるゼミでの実践を通して —

三森伸一郎

宮城教育大学大学院教育学研究科高度教職実践専攻 修了

I はじめに

平成 28 年 4 月 1 日に、「障害を理由とする差別の解消の推進に関する法律」(障害者差別解消法)が施行され、障害者の状況に応じた社会的障壁の除去のための合理的配慮を行う必要があるとされている。しかし、聴覚障害のある学生が在籍している大学では、学校生活上における支援が充実しているところもあれば、不十分のところもあるのが現状である。特に少人数での音声によるコミュニケーションは難しい。しかし一方、学習者同士の議論に活用可能と想定されているアプリも開発されており、その活用が有望視される。

宮城教育大学では、「UD トーク」(Shamrock Records, Inc.)を 2014 年 11 月にトライアル導入をし、2015 年 1 月から教育機関プランで法人契約している(松崎 2017)。宮城教育大学大学院修士課程の少人数討論型の授業を対象に、UD トークを活用した支援システムの運用についての検討を行った(松崎 2017)。しかし、少人数での議論場面で UD トークを活用した支援システムの実践事例の報告は少ないのが現状である。そこで、本研究は、UD トークを活用したシステムの設定方法とその留意点を明らかにすることを目的とした。

II 研究の方法

1 音声認識システムの概要

音声認識アプリは、教育機関プランで法人契約をしている UD トークを用いた。UD トークをインストールしたスマートフォンを 4 台、ハンドマイク 3 本、認識結果表示用タブレット 1 台、タブレット画面を表示するテレビ 1 台を使用した。テレビは、授業担当教員、受講者全員がタブレットで表示されている認識結果が分かるように表示した。机と椅子の配置等の詳細は、図 1(左端 1 回目)の通りである。

2 調査対象

音声認識システムの活用を実施している大学院教育学研究科高度教職実践専攻の授業 1 科目を対象とする。本授業では、教員 2 名、健聴学生 4 名、筆者である聴覚障害学生 1 名(以下、M)で、研究に関する議論が行われている。M は、先天性感音性難聴を有しており、聴力レベルは右 90dB、左 105dB である。なお、音声認識アプリの活用期間は、2017 年 4 月 - 2018 年 1 月である。

3 調査の手続きと分析方法

健聴学生 4 名を対象に、音声認識システムに関する質問紙調査を 10 回行った(2017 年 5 月 - 11 月)。1・2 回目は、音声認識システムのメリットとデメリットについて自由に記述してもらった。3・4・5・6・8 回目は、音声認識システムの改善点や改善案についての自由記述を要請した。7・9 回目は、新しいシステムを導入した時の効果の記述を要請した。1~9 回目の対象分の記述内容を KJ 法(1967 川喜田)により整理・分類し、音声認識システムの改良を行った。最後の 10 回目の調査では、音声認識アプリの活用(図 1 - 10 回目)を踏まえて、授業担当教員も音声認識システムに関する「ハンドマイクの使いやすさ」「認識結果等を表示したテレビ画面の見やすさ」「ルールを守ることができたか」「机と椅子の配置の良さ」「キーボードの使いやすさ」の 5 段階評価とそれぞれの設定方法についての記述、「修正方式と要約方式のどちらがよいか」についての設問への回答と、その理由の自由記述も要請した。

Ⅲ 研究の結果

1～9回目の対象者の記述内容をもとに、表1のようにラベルを抽出し、大カテゴリーと小カテゴリーを作成した。以下、大カテゴリーと小カテゴリーを用いて、音声認識システムの改良の経緯について述べる。

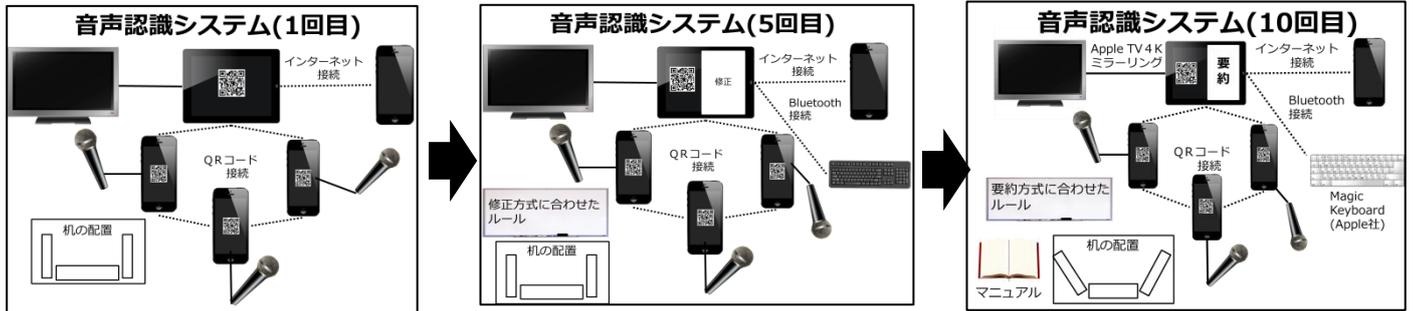


図1 音声認識システムの改良の経緯

表1 対象者の記述から整理・分類したカテゴリー

回数	大カテゴリー	小カテゴリー	ラベル数
1・2回目	音声認識システムにおける話者の行動	認識結果を見るのが難しい	3
		マイクの操作を忘れてしまう	3
	音声認識システムの誤認識・誤変換の原因と防止策	誤認識・誤変換が起こる理由	2
		誤認識・誤変換を防ぐ方法	4
	音声認識システムを活用することで生まれた意識	緊張感	2
話し言葉の視認	参加しているという意識	2	
		相手に伝えようとする意識	1
3・4回目	画面を見ることに関する課題	文字の大きさに関する課題	4
		話者が感じている課題	1
		訂正者が感じている課題	1
	音声認識システムのルール作りの必要性	UDトークのルール	2
		誤認識・誤変換の訂正のルール	3
訂正方法の追加・改善	訂正方法の追加	2	
		使用機器に関する改善	2
5回目	環境の調整		4
	修正方式の変更に関する提案		4
6回目	修正方式と要約方式のそれぞれのメリットとデメリット		7
	今後の方向性に関する提案		2
7回目	従来のキーボードの使いにくさ		3
	要約を担当している人への配慮点		3
	要約方式が向いていると実感		1
8回目	新しいキーボードに変更したことでの効果性		3
	要約を担当している人が自由に発言できるように配慮することの効果性		3
	要約できない時の対処方法		1
9回目	要約を担当することで生じる難しさ		2
	音声認識システムに対する慣れ		1

1・2回目の回収した対象者全員分の記述内容から18個のラベルを抽出し、四つの大カテゴリーを作成した。ここでは、「音声認識システムの誤認識・誤変換の原因と防止策」のみを取り上げる。「音声認識システムの誤認識・誤変換の原因と防止策」とは、健聴学生は話し手の音声情報と起こされた文字情報を比較することができるため、音声情報が正しく文字に起こされていない誤認識・誤変換を引き起こす原因を推論し、それを改善するための行動を起こしたことである。例として、口の形を意識したり、ゆっくりとはっきりと話したりするなどが挙げられる。しかし、健聴学生が誤認識・誤変換の回数を減らす行動を起こしたのにもかかわらず、Mは誤認識・誤変換が多いことから話し手の話している細かい内容を理解することが難しい現状があった。そこで、健聴学生が誤認識・誤変換の修正を行うことができるように、キーボードをBluetoothで接続した「iPad pro」(Apple)をテレビ画面に接続し、左画面にUDトークを表示し、右画面にメモ機能のある「Evernote」(Evernote Corporation)を表示し、Evernoteに誤った言葉を正しい文字に入力できるようにした。修正者を担当するのは一名であり、本授業を受講している健聴学生4名にお願いした。なお、話し手ごとに修正者担当する人を変えた。

3・4回目の回収した対象者全員分の記述内容から15個のラベルを抽出し、三つの大カテゴリーを作成した。ここでは、「音声認識システムのルール作りの必要性」のみを取り上げる。「音声認識システムのルール作りの必要性」とは、音声認識システムを効果的に活用するための方法を検討し、それを教員や健聴学生に提示する必要があることに対応するものである。そこで、UDトークのホーム画面にあるヒントと健聴学生から訂正方法のバリエーションを増やしてほしいという記述を踏まえた上で、図2-5回目のようにルールを検討した。ここまで改良した音声認識システムは図1(中央5回目)のとおりである。

5回目の対象者3名の記述内容から、8個のラベルを抽出し、二つのカテゴリーを作成した。ここでは、「環境の調整」のみを取り上げる。「環境の調整」とは、話者と修正者が音声認識システムを効果的に活用するために、文字の大きさ等の環境を調整することである。また、誤認識・誤変換を正しい言葉に修正する

のに、時間を要するという記述も得られた。そこで、上記をもとに、誤認識・誤変換を修正するためのメモアプリを「Evernote」から「文字数(DRIP PRODUCTS LIMITED LIABILITY CO.)」に変更し、文字を大きくすることができるようにした。さらに、アンケートから誤認識や誤変換を正しい文字に変換し、表示するまでに時間がかかってしまうという記述が得られたことから、6回目から、誤認識・誤変換を正しい言葉に修正する修正方式から話者の話を要約する要約方式に変更することも検討した。なお、要約方式はメモアプリである「文字数」でキーボードによる文字起こしを行い、文字情報はテレビの右画面に表示した。

6回目の対象者4名分の記述内容から、9個のラベルを抽出し、二つのカテゴリーを作成した。ここでは、「修正方式と要約方式のメリットとデメリット」のみを取り上げる。修正方式は、誤認識・誤変換を正しい言葉に修正するのみであるため、すぐに修正した結果を画面に表示することができるメリットと話者の話の内容を一字一句覚えなければならないため、修正者への負担が大きくなるというデメリットがある。要約方式は、要約を担当する人が話者の話を聞くことに集中できるメリットと話者の話のある程度終えてから、取り組むために文章が画面に表示されるまでの時間がかかるというデメリットがある。そこで、上記を踏まえ、修正方式と要約方式に関する資料(資料1 参照)を作成し、要約方式のみの実践を行うことにした。

7回目の対象者2名分の記述内容から7個のラベルを抽出し、三つのカテゴリーを作成した。ここでは、「従来のキーボード(下記参照)の使いにくさ」と「要約を担当している人への配慮点」のみを取り上げる。「従来のキーボードの使いにくさ」とは、全角と半角の切りかえのボタンが分からなかったり、以前からキーが小さかったりするなど従来のキーボードが馴染まないことである。「要約を担当している人への配慮点」とは、要約を担当している人が話の内容を理解することができなかったり、質問や意見を発言することができなかったりするなどが起こらないようにする方法を検討することである。そこで、音声認識システムを四つの改良を行った。一つ目は、従来のキーボードである「Wireless Keyboard minisizeTK-FBP052」(ELECOM)をよりキーの大きい「Magic Keyboard-日本語」(Apple)に変更したことである。二つ目は、修正方式よりも要約方式の方が取り組みやすいという記述があったため、これまで取り組んできた修正方式から要約方式へ変更したことである。三つ目は、要約を担当している人も議論に参加できるように、その配慮点をルールに盛り込んだことである。図2の「ルール(7回目)」のとおりである。四つ目は、変換時の文字を大きく、太くしたことである。なお、キーボードの変更と要約を担当している人への配慮点を設けたことにより、8回目の記述内容から対象者からその効果性に関する記述が得られた。

8回目の対象者3名分の記述内容から7個のラベルを抽出し、三つのカテゴリーを作成した。ここでは、「要約できない時の対処方法」のみ取り上げる。「要約できない時の対処方法」とは、要約を担当している人が話の内容を理解できなかったときに「・・・」と打ち込むことである。ここでは、改良は行わなかった。

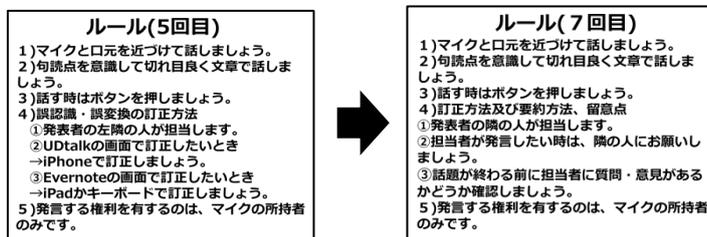


図2 ルールの改良

9回目の対象者2名分の記述内容から3個のラベルを抽出し、二つのカテゴリーを作成した。カテゴリー内にある「要約を担当することで生じる難しさ」とは、8回目の大カテゴリーにある「要約できない時の対処方法」とほぼ同じ記述である。そこで、以下の三つの改良を行った。一つ目は、要約を担当している人が分からなかったときは、「??？」と打ち込むように資料に加えたことである。二つ目は、以前に対象者からテレビ画面と前の座席に座っている人が重なってしまい、画面が見づらいという記述があったため、机と椅子の配置を四角形から馬蹄形へと変更したことである。三つ目は、テレビに「iPad pro」(Apple)の画面をよりきれいに表示するために、「Apple TV 4K」(Apple)の導入を行ったことである。

10回目の結果は、音声認識システムに関する「ハンドマイクの使いやすさ」「認識結果等を表示したテレビ画面の見やすさ」「机と椅子の配置の良さ」「キーボードの使いやすさ」「修正方式と要約方式のどちらがよいか」の観点に基づいて、音声認識システムの設定方法に関して参加者の感じていることについて述べる。「ハンドマイクの使いやすさ」は、スイッチの操作が容易であること、話すことに集中してしまうと、マイクを持つことを忘れてしまうことがあることが明らかになった。「認識結果等を表示したテレビ画面の見やすさ」は、以前よりも文字が大きくやや太いため、見やすいこと、要約を担当している人は役割を果たすことで精一杯であるため、音声認識システムと要約方式によって起こされた文字情報が表示されている画面を見ることができないことが分かった。「ルールを守ることができたか」は、音声認識システムの活用を繰り返すうちに自然にルールを守ることができたこと、それを忘れてしまうことがあるとのことであった。さらに、「話者はTV画面を見ながら話す」というルールを加えればよいのではないかという提案もあった。「机と椅子の配置の良さ」は、議論も参加しやすく、画面も見やすいという効果があることが分かった。「キーボードの使いやすさ」は、Magic Keyboard—日本語(Apple)に変更したことで使い心地が良くなったという記述を得た。「修正方式と要約方式のどちらがよいか」は、修正方式の方がやりやすいと答えた人は1名、要約方式の方がやりやすいと答えた人は3名、無回答が1名であった。修正方式を選択した人は、修正されると配慮して話そうと思う効果があることを理由として挙げていた。要約方式を選択した人は、自分の中で話しを再構成してまとめる方が自分に合っている理由と自分も議論に加わることができるという理由を挙げていた。無回答を選択した人は、修正方式は直すことに必死になってしまうというデメリットがあり、要約方式は要約ができなかった時に混乱してしまうというデメリットがあると挙げていた。

新しい音声認識システムの具体的な設定方法は、(右端10回目)のとおりである。

IV 考察

1回目と比較して、聴覚障害学生も健聴学生も音声認識システムを意識した議論を行うことができるようになった。少人数の議論場面における音声認識システムを検討する上で、聴覚障害学生に細かい内容を理解するための一助となる支援方法を組み込む必要があると思われる。例えば、要約方式を導入するなど挙げられる。支援方法を検討する際には、聴覚障害学生が音声認識アプリで起こされた文字情報から議論の全体像と細かい内容を理解できるかどうかを確認した上で、聴覚障害学生と健聴学生との対話を通して、検討することが重要である。

さらに、議論場面で活用するためには、「音声認識システムの誤認識・誤変換が生じた時のための支援方法を検討すること」、「全員が参加しやすい環境づくり」、「音声情報を正しく文字に変換することのできるルールづくり」の視点が必要であることが示唆された。

最後に、少人数におけるゼミの当事者であるMの感じたことについて述べる。UDトークは、大学院入学時から活用している。当初は、音声を文字に変換するスピードが速いことと誤認識と誤変換への対処方法が分からなかったことから、議論の全体像や細かい内容を把握することができないことが多かった。しかし、健聴学生からのアンケートをもとに音声認識システムの設定方法を検討していくことで、音声認識システムに対する健聴学生の姿勢が変わっていった。例えば、ゆっくりとはっきりと話すようになったり、音声認識システムの改良に関する提案をいただいたり、Mの理解度を尋ねゼミにフィードバックしたりするなどである。そうすることで、誤認識と誤変換が出てくる回数が減り、Mが理解しやすい

ものになっていった。以上のことから、音声で文字入力を行う健聴学生が使いやすいように環境を設定していくことも必要な視点だと思われる。

なお、本研究で得られた基礎的知見をもとに、「複数人における議論場面の音声認識システムの活用事例に基づいた参考資料-宮城教育大学教職大学院-」を作成した(資料1)。聴覚障害教育の関係者もしくは聴覚障害者への情報提供のために作成したものである。本資料をきっかけに、音声認識アプリをどのように学校生活へ取り入れるかを検討する際の判断の一助になることを期待している。

V 引用・参考文献

- 大塚善樹(2017) 少数者ゼミにおける聴覚障害学生のための学習支援システムの検討 東京都市大学横浜キャンパス情報メディアジャーナル 2017.4 第18号
- 川喜田二郎(1967) 発想法 創造性開発のために 改版 中公新書
- 内閣府(2014) 障害を理由とする差別の解消の推進に関する法律
- 根本匡文(2006) PEPNet-Japan TipSheet 聴覚障害幼児・児童・生徒を囲む教育環境④ 日本聴覚障害学生高等教育支援ネットワーク
- 松崎丈(2017) 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—少数者討論授業を事例に— 宮城教育大学情報処理センター研究紀要
- 三好茂樹・黒木速人・川野純大・白澤麻弓・石原保志・小林正幸(2007) 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 mar.2007
- 三好茂樹(2008) PEPNet-Japan TipSheet 音声認識技術を用いた情報保障⑩ 日本聴覚障害学生高等教育支援ネットワーク
- UD トーク HP(2018) <http://udtalk.jp/> (最終確認期日 2018.1.23)

複数人における議論場面の音声認識システムの活用事例に基づいた参考資料

—宮城教育大学教職大学院—

宮城教育大学大学院高度教職実践専攻 三森伸一朗

0. はじめに

本レジュメでは、一斉授業方式による活用は検討せず、複数人における議論場面を想定しているものである。

1. 音声認識とUDトーク,誤認識・誤変換

音声認識とは、人が話した言葉を、機械等が文字に変える方法のことである。UDトークとは、「『共有』を目的とした話し手から聞き手へのコミュニケーションをサポートするためのアプリ」(UDトーク HP 2018)である。誤認識・誤変換の例としては、以下のとおりである。それが生じると、聞き手に話し手の意図を伝えることができない可能性がある。

誤認識:話し手「す」→UDトーク「つ」 誤変換:話し手「用事をお願いします」→UDトーク「幼児をお願いします」

2. 音声認識システム

音声認識システムは、修正・要約方式を加えない利用方法と修正・要約方式を加えた利用方法がある。

(1) 修正・要約方式を加えない利用方法

その場での情報保障を音声認識に全て任せるという形である。これを図式化したものは、図1のとおりである。



図1 直接利用による方法の

(2) 修正・要約方式を加えた利用方法

修正方式による利用方法は、誤認識・誤変換を修正者が訂正をした上で、聞き手に正しい文字情報を伝達するという形である。要約方式による利用方法は、音声認識による誤認識・誤変換を踏まえた上で、話者の発言を要約し、聞き手に伝達するという形である。修正方式と要約方式で用いるアプリケーションは、「文字数」(DRIP PRODUCTS

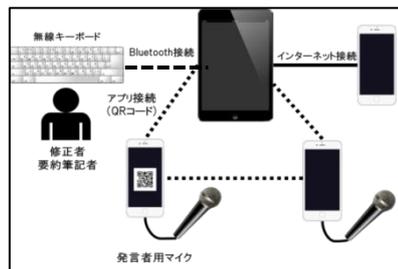


図2 修正方式と要約方式の利用



図3 アプリとアプリの接続方

LIMITED LIABILITY CO.)である。これらを図式化したものは、図2のとおりである。

(3) 接続方法

1) 直接利用による方法

- ① iPhone 2台, ハンドマイク 1本を用意する。
- ② 全ての機器で UD トークを開き, 発言者用マイクに接続した iPhone 1 台に表示された「トークを公開する」をタップし, QR コードが表示されるのを待つ (図3)。
- ③ その他の機器は, 全て「トークに参加する」を押し, QR コードを読み取る。

2) 間接利用による方法

- ① iPad pro 1 台, iPhone 3 台, ハンドマイク 2 本, Magic Keyboard 1 機を用意する。

②iPad と iPad pro はインターネットに接続していないため,iPhone1台と iPad1台をインターネット共有によって,



図4 iPad pro をインターネットにつなげる

接続する(図4)。



図5 キーボードのスイッチ(ON状態) 図6 キーボードを Bluetooth で接続する方

③キーボードのスイッチ(図5)を入れ,iPhoneもしくは iPad に Bluetooth 接続する(図6)。

④全ての機器で UD トークを開き,発言者用マイクに接続した iPhone1 台に表示された「トークを公開する」をタップし,QR コードが表示されるのを待つ。

⑤その他の機器は,全て「トークに参加する」を押し,QR コードを読み取る。

⑥二画面設定にする場合は,UD トークの画面を表示した iPad pro を下からスワイプし,「文字数」を左画面までスワイプする(図7)。

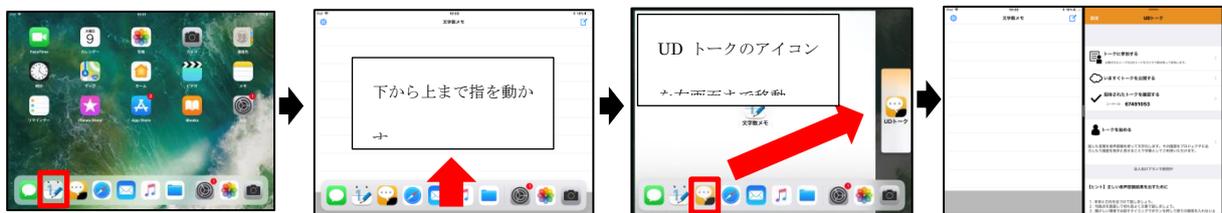


図7 二画面設定の方法—iPad pro のみ

3)備考

①PC に認識結果を出して,キーボードで変換する方法もある。

②宮城教育大学はネットワーク上の関係で,PC に認識結果が表示されるのに時間を要する。そのため,タイムラグの短い iPhone もしくは iPad による修正の方法がおすすめである。

③音声認識システムの設定方法は,人それぞれである。私は図8,図9のように設定した。ルールをポスター等で掲示すると,話者全員がルールを守ろうとする意識が生まれることがある。マイクの本数は,2人に1本設定した。

④認識結果を大画面に表示する方法がある。話者全員が自分の認識率をチェックすることができる。そのためには,Apple TV もしくは USB-C Digital AV Multiport アダプタが必要である。

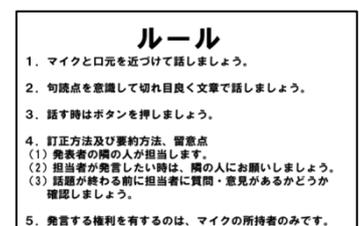
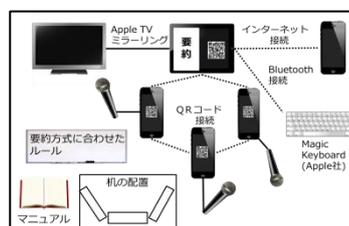


図8 音声認識システムの 図9 ルールの例

4. 音声認識システムを正しく活用するための方法

(1)話者

正しい音声認識結果を出すために的確な話し方が求められる。具体的には、以下の3つ(UDトークアプリ ホーム画面 引用)を実践すれば、高い認識率で文字が表示される。

- 1) マイクと口元を近づけて話しましょう。
- 2) 句読点を意識して切れ目良く文章で話しましょう。
- 3) 騒がしい環境では話すタイミングでボタンを押して周りの雑音を入れないようにしましょう。

修正者が不在の場合は、以下の方法で修正することができる。

1) 文章レベルによる言い直しによる方法。2) UD トークの修正したい文章をタップすると、文字の訂正ができる機能を活用する方法。

(2)修正者

修正者とは、音声認識の認識結果を表示した画面上にある誤認識・誤変換を正しい文字に修正する人のことである。修正する方法は、主に三つある。

- 1) iPhone に表示されている UD トークの画面で訂正を行う方法。
- 2) テレビ画面もしくは iPad pro に表示されているメモアプリ「文字数」に Bluetooth で接続したキーボードで正しい文字に修正する方法(図10)。
- 3) 誤認識・誤変換の部分を言い直す方法。

修正時の留意点は、優先順位を意識することである。具体的には、以下のとおりである。

- 1) 意味が推測しにくい文章がある時は、文章ごとの修正をする。

例,(正)「明日の授業内容は、模擬授業と討論をします」
(誤変換)「明日の需要無いようは、裳着従業と党論を。」

- 2) 誤認識の時は、修正をする。

例,(正)「幼児(ようじ)」 →(誤認識)「料理」

- 3) 誤変換のときは、修正は不要である。

例,(正)「幼児(ようじ)→(誤変換)「用事」

(3)要約筆記者

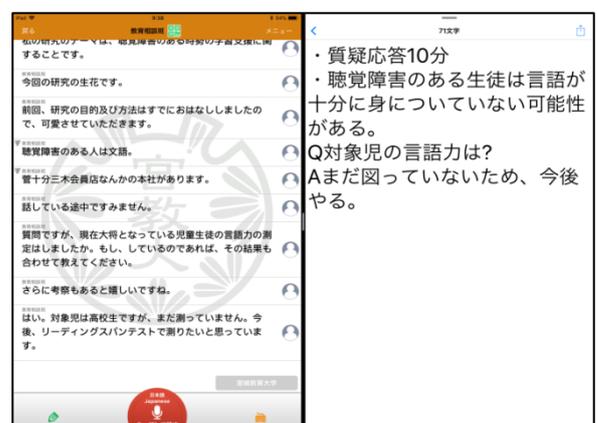
要約筆記者とは、話者の音声情報と音声認識の認識結果をもとに、議論場面の内容を要約する人のことである。要約する方法は、テレビ画面もしくは iPad pro に表示されているメモアプリ「文字数」(DRIP PRODUCTS LIMITED LIABILITY CO.)に Bluetooth で接続したキーボードで正しい文字に修正する方法である。(図11)



図10 修正の例



(要約前)



(要約後)

図11 要約の流れの例

要約時の留意点は、以下のとおりである。

1)議論の時には名前を入力してから、内容を打ち込む。

例、田中／障害理解教育を進めるためには、①障害の疑似体験、②教員の共通一致が必要・・・。

2)質疑応答の時には、以下のように打ち込む。

・質問→Q:本実験の対象者は何人？

・応答→A:20人

3)一人の話者につき、一文でまとめることができるように工夫する。

4)分からなかったときは、「??？」と打ち込む。

(4)修正者・要約筆記者自身が発言をする場合

1)挙手をする。2)キーボードを隣の人に渡す。3)議論が終わったとき、キーボードを隣の人から受け取る。

5. その他

UD トークのホーム画面の設定内に「音声認識辞書に単語を追加する」がある。そこに使用頻度の高い単語を登録しておく、認識率が上がる。

6. さいごに

UD トークは聴覚障害者に音声情報を正しく伝達できるとは限らない。そのため、マニュアルどおりに従うのではなく、実際に活用した上で自分に合った方法を探ることが大切である。