

スピーキング指導可視化ツール開発と指導意思決定支援 AI 自動赤ペン添削で弱点を図示する学習履歴マイニング

田淵龍二¹ Ryan Spring²

¹ ミント音声教育研究所 (tabuchiryuji@nifty.ne.jp)

² 東北大学

Abstract: 英語教育 4 技能でもっとも遅れているスピーキングテストを改善するために開発した赤ペン添削を使った英語音声指導授業を行った。課題文を生徒が発声するとわずか 3 秒で添削して返す。生徒は判定結果の赤ペン箇所（音素や単語）に注意しながら直ちに再挑戦する。教員が学習履歴マイニング結果を 10 秒で俯瞰してエラーの多い語彙や表現を把握することで、最適化したクラスレッスンを同じ授業内で返す意思決定を支援できた。

1. はじめに

本論考は、自動音声認識（ASR）と、スピーキングテキストの双方向添削と、テストデータ（学習履歴）から有用な知識を取り出すデータマイニングの 3 つの技術を使ったスピーキングテストの実施結果を、指導支援の観点から研究したものである。

双方向添削とは、読み上げ用英文（以下たんに課題文）と読み上げ音声を ASR（Automatic Speaking Recognition）が書き起こした文（以下たんに結果文）双方の相違点をテキスト形式で併記する方法である。課題文は話し手が読み上げたテキスト、結果文は ASR が書き起こしたものである。課題文は話し手の意図をテキスト化したもので、結果文は聞き手が受け止めた意味内容をテキスト化したものである。

スピーキング双方向添削で収集した学習履歴から生徒の活動状況を臨場感を持って再現する可能性を追求した。臨場感を再現するには、活動内容と時間の 2 つの要素が不可欠だと考えた。スピーキングの核は課題文の読み上げと評価であり、時間はそれらの開始時間や所要時間である。

授業で使ったスピーキングアプリ NatTos では 3 万件余りのデータセットを収集した。

感染症蔓延による 2020 年オンライン授業への急変の中でありながら、どんなクラス授業が行われ、個々の生徒がどのように試験に取り組み、どんな発声だったかをまるでそばで見ているかのように再現することを目指した。結果は予想を超えて成果を上げることができたのでその様子を紹介する。

2. 方法

スピーキングテストの方法や環境を列挙する。

- ツール： 双方向赤ペン添削スピーキングテスト
- 名称： NatTos ナットーズ
- 公開： 2020 年 6 月
- アクセス： <http://www.mintap.com/nattos/>
- 特徴： 全自動双方向添削、3 秒レスポンス
- 機能： ASR による結果文と課題文を併記して相違点を赤ペンで視覚化する
- データ： 1 件あたり約 20 次元
- 媒体： iPhone, iPad, Android、パソコン（Windows）
- 登録： 無登録、料金無料
- 対象： 東北大学 1 年生 3 クラス, 100 人余り
- 科目： 英語
- 期間： 週一 5 週間（2020 年 11 月～2021 年 1 月）
- 形態： オンライン
- 運営： 90 分授業の冒頭と末尾に 15 分の小テスト
- 課題文： テーマに合わせて自作、5 回分 69 本
- 週ごとのテーマ： 子音、母音、ストレス、脱落と連結、リズム
- 文長： 単語、句、短文、長文
- 文レベル（文のみ 53 本）： 高校 1 年±2
- 課題配布： 小テスト用 URL をウェブで配布
- リード音声： 課題を開始すると課題文を TTS (Text To Speech) で読み上げる
- 課題遂行： リード音声を聴取後、読み上げ開始
- 補助音声： 必要に応じてリード音声を再生可能
- 制限： 同じ課題文は最大 5 回で打ち切り
- 指示： 小テストについて、特段の指示はない
- 指導： 前テストの結果を直後の発音指導に反映
- 課外： 授業中に終わらなかったときは宿題

3. 結果

小テストで収集した履歴の基本情報を列挙する。

- データベクトル： 10 数次元、32,790 件
- 受験者： 3 クラス約 100 人分

3.1 クラス授業開始直後の様子

クラス授業としてスピーキングテストをおこなった。テストを開始して 5 分ほど過ぎた頃に教師が学習履歴を開くと Fig.1 のような進行状況が確認できる。24 名がテストを始めている。題名 pronunciation 2 の下にユニット名が word 1 から sentence 2 まで縦に 4 つ並びそれぞれ 4



Fig.1 テスト開始 5 分後の進行状況

本ずつの課題文にアクセスする円形アイコンが並んでいる。円形アイコンにはひまわりや楕円形コインが描かれ、大きさや色とその下の数値が達成度と人数を示している。一番早い生徒は第 3 ユニットの第 3 問目 (B) に到達している反面、ユニット 1 (A) がまだ終わらない生徒が 7 名ほどいることも分かる。

3.2 クラス授業終了後の確認

授業後にテスト運営の様子を時系列で振り返ることができる。Fig.2 は折れ線グラフを縦にしたもので、時間表示が目盛りになっている。普通は左外に記載される目盛りがグラフの内側に回り込んだ格好である。横軸は添削数で左下を原点とし右に向かって増大する。目盛りは記載されていないが、その代用として時刻表示文字列の末尾にある括弧入り数値が度数である。たとえば一番下の文字列は 2020 年 12 月 14 日 14 時 36 分台の 1 分間に 7 個の課題文が始まったことを示している。折れ線のピークが 14 時 43 分過ぎと 15 時 57 分過ぎの 2 ヶ所に見られ、その 2 つの山は上下 10 分ほどで谷になっていることから、14 時 40 分ごろから 16 時 10 分ほどの間にそれぞれ 20 分弱の間スピーキングテストを行ったと読み取れる。

90 分授業であったことから、授業の出だしと締めくくりで 15 分ほどずつテストが行われたことと符合している。谷間の 60 分間ほどはレクチャー時間であったが、よく見ると数名がテストを行っていた形跡が見える。遅刻してきた生徒かもしれない。

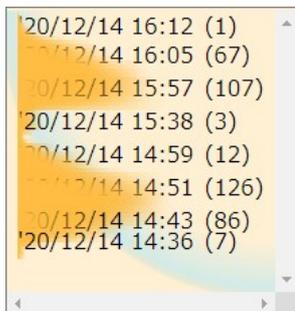


Fig.2 テスト運営状況

3.3 課題文の達成度を概観する

課題文への取り組みが始まると、受験者は同じ課題を 5 度まで挑戦できる。5 度以内にクリアすると次の課題へと進む。クリアできなくてもその課題は打ち切って次に進む。クラス全体を一人の生徒に見立てて達成状況を概観できるのが Fig.3 の棒グラフである。全体が 100%である。右側の桃色が失敗で 52%と読める。したがってその左側 5 区画の合計 48%が成功したことになる。成功したのが何回目かは、左から何番目かでわかり、1 度目の成功が 22%、2 度目が 11%で、8%、5%と続き、最後の 5 度目にぎりぎり成功したのは 2%であった。右上方の数字から 825 回の添削が行われたことが分かる。

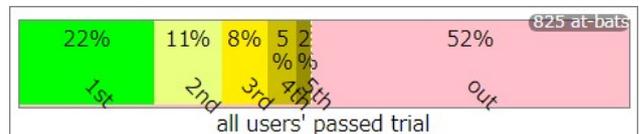


Fig.3. 課題文の達成状況を概観する到達度チャート

この図を始めて見たとき、とても驚いた。それまでぼんやり思っていたことが数値化されて出てきたからである。通常のスピーキングテストは 1 度きりの音声で受験者の学力を判定する。この図で言えば左端の 22%が正解と評価される。しかし、例えば次の左の 11%は最初は失敗したが 2 度目に成功している。その力量は今まで捨てられていたことになる。この力量とは修正力である。自分が発した声在意図したものとは違って相手に伝わったことがわかったとき、正しく言い換える力、あるいは言い直す選択肢を所持している力である。



Fig.4 双方向赤ペン添削

3.4 双方向赤ペン添削の様子

1 度目の間違いの様子だったのかを知ることは、次に成功するために不可欠だ。そこで双方向赤ペン添削で生徒がどんな画面を見ているかを Fig.4 に示す。1 度目の添削が終わったところである。A が読み上げ用英文 (課題文)、B が発声音声を ASR が聞き取った結果文、C が正しさと速さの評価値である。生徒はこれを確認したあと下の赤丸をタップして再

挑戦する仕組みだ。

添削箇所を見てみよう。A では ran に ran と赤が入っている。ran を意図した発声が実際には別の音声で聞こえたことを示している。ではどんな音だったのかが知りたいはずだ。それにはすぐ下の B にある run を見れば良い。/æ/ の音が弱かったのだと悟った生徒はきっと次には成功させようと思うに違いない。

3.5 間違い探しから、どう聞こえたかへ

ここで従来のスピーキングアプリと比べてみよう。一昔前までは大規模な CALL 教室でしか使えなかったレベルのサービスがスマホで気軽に使えるようになった。たとえば 2018 年から提供されている進研ゼミの英語スピーキングアプリは会員限定サービスで、音声を評価し添削してくれる (Fig.5)。表示された英文を読み上げると、レッスン終了後に言い間違った単語が彩色され、改善ポイントが示される。隣に英語母語話者がいなくても、「ここが違ったよ」と学習者に教えてくれる便利なアプリだ。



Fig.5 片方向スピーキングアプリ例[1]

図の A を見ると should に赤が入っていることから、受験者は should が失敗したと分かる仕組みだ。どんな音声だったのかは、録音しておいた自分の声を再生できるようになっている。間違いを指摘するだけでないところは工夫が見られる。しかし学習途上の生徒には荷が重いかもしれない。やはり「こう聞こえたよ」と文字で直接的に教えてもらった方が話が早いだろう。話し手と聞き手のインタラクティブでわかりやすい情報交換を心がけたい。

3.6 修正力を発揮させた双方向赤ペン添削

さて話を達成度グラフ Fig.3 にもどす。自分の声が相手にどう聞こえたのかを知ることがその後の発声による影響を与えているのかについて、Fig.3 は肯定的な証拠を与えてくれる。1 度目の発声で失敗した受験者 78% (100%-22%) のうち 14% (11%÷78%) が 2 度目に成功している。さらに 2 度目も失敗した者 67% (100%-22%-11%) のうち 12% (8%÷67%) が 3 度目に成功している。このようにして 1 度目に失敗した 78%のうち 32% ((11+8+5+2)%÷78%) が 5 度目までに成功させていることが分かる。この 32% は 1 度目に成功した 22%より高い比率である。この

間誰からも指導を受けていない (はず) ことから、自力で工夫して解決したと考えられる。「君の声はこう聞こえたよ」と言う話し手と聞き手のインタラクティブな情報交換が、学習者の自力修正力を引き出した可能性が推察される。

3.7 自力修正を示唆する事例

赤ペン添削による話し手と聞き手のインタラクティブな情報交換によって受験者が積み重ねたスピーキングの様子を具体例で見よう。

3.7.1 自力修正が成功した例

Fig.6 は自力修正力がうまく働いた例である。この図からは、ある受験者が同じ課題を 2 回実施し、1 回目は 2 度失敗したあとに 3 度目で成功し、2 回目には 1 度目にいきなり成功させている様子がわかる。それは図の上半分の背景色が灰色の欄の下 2 行 (A,B) から見て取れる、白抜き番号 1 から始まる行 (A) が 1 回目の受験、番号 2 の行 (B) が 2 回目である。各行にいくつかずつ並んでいる正方形が添削結果を視覚化したもので、左から順に 1 度目、2 度目…である。正方形は左右真二つに彩色されていて、左の緑が正しさ、右の青がなめらかさである。緑の縦長矩形に書かれた数字を見れば、何度目に成功したかが分かる仕掛けになっている。滑らかさを示す青の長矩形は上辺いっぱい伸びていることから、1 秒間に 3 単語以上の円滑な発声であったことがわかる。

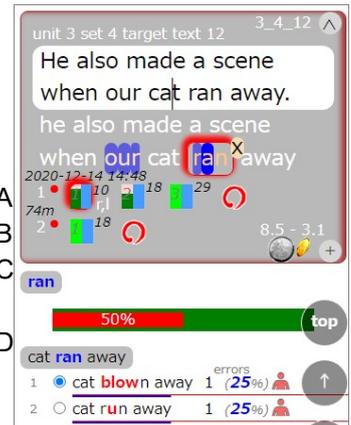


Fig.6 自力修正例 (1.1)

さて、1 回目の成功に続けて 2 回目を行えばうまくいくのは当たり前と思われるだろう。しかしこの図には 1 回目と 2 回目がどのくらいの時間間隔において発声されたかを示す情報が埋め込まれている。それが行頭白抜き番号のすぐ上にある黒文字の数字である。1 回目は 2020-12-14 14:48 と読める。授業開始時間は 14:40 であったから、授業冒頭の小テスト中の発声だ。2 回目の文字は 74m とある。これは 1 回目開始時から 74 分 (minute) 後、つまり 16:02 に 2 回目を開始したことを物語っている。授業終了時刻は 16:10 でであったから授業終了前の小テスト中とわかる。つまり前テスト終了後のレクチャーを 1

時間ほど受けたあとの後テストである。1 時間余りの間をおいた発声であることから「できて当たり前」とは言えないだろう。

次にそれぞれの発声がどの程度の間合いで行われたのかを見てみよう。それは正方形アイコンのすぐ右肩の数字でわかる。この数字はリード音声を聞いたあとの時間経過を示している。2 度目以降は直前との差が間合い時間となる。1 回目は、10 秒後に添削に回している。失敗を知ったあとの 2 度目までは 8 秒、3 度目は 11 秒である。直前の間合い 8 秒より長い 3 秒の間に、録音前に空読みして練習をしていたと考えるのは深読みだろうか。残念ながらそこまではデータに入っていない。

さて、図の下段に移ると、ここには課題文にある語彙 ran の失敗の様子が記述されている。1 行目 (C) の ran がそれを示す。2 行目は対象語彙 ran の失敗が失敗全体の半分であることを示している。3 行目 (D) は特徴的な失敗例の表示枠に入ったことを示し、すぐ次の 2 行がその事例である。

特徴的な失敗例とは、対象語彙の前後は正しく聞き取られた場合の失敗事例を指した分類である。こうした分類概念によって、この失敗に注目して音声指導や自力修正をするように提案している。弱点克服が目的である。

番号 1 にチェックが入っている。これにより、この類型失敗をした発声に対応した上の欄の正方形に赤い枠が付く仕組みになっている。ここでは 1 回目の 1 度目がそれにあたる。この失敗例を見ると、R 音 ran が L 音 blow になってしまったとわかる。

ここでは教師用学習履歴の話だが、生徒は Fig.7 にあるような赤ペン添削を見ている。生徒はこの赤ペンから R/L エラーだと自覚し、そこへ修正すれば合格だと考えたことだろう。画面下半分の枠には R/L エラーへの対処法がコメントされている。



Fig.7 赤ペン添削画面 (1)

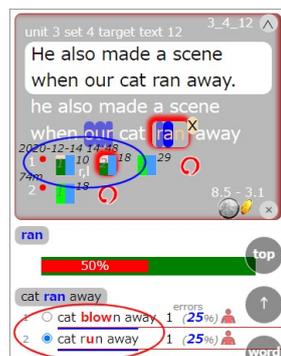


Fig.8 自力修正例 (1.2)

さて話を Fig.6 にもどす。図の最下段にある番号 2 をチェックすると cat run away のエラー類型についての表示に切り替わる。それを Fig.8 に示す。赤丸がチェックを切り替えた様子を示し、青丸内の赤い縁取りが 1 度目から 2 度目の正方形に移動していることがわかる。1 度目で失敗した R/L エラーは修正されたのだが、今度は ran が run になっていて、別の æ/ʌ エラーにハマってしまった。

このときに受験者が目にする双方向添削を Fig.9 に示す。1 度目と比べると Blown だったのが run になっていて R/L エラーは修正されている。しかし別の問題が発声した。our が I have a となり、ran が run になっている。受験者の気持ちを想像してみよう。R/L エラーだったので R 音に集中しようと頑張った舌の動きを変化させたのだが、その影響が他の 2 箇所にも現れてしまった。一方では our の R 音が気になる余り不自然な口の動きになって R 音が A 音を誘発し、他方では ran の R 音に集中して ran の A 音がおろそかになったのだろう。そこで口慣らしの空読みを 1、2 度おこなってから 3 度目に成功させたと推察される。それが 2 度目から 3 度目の間合いが 8 秒から 11 秒に伸びた現象として観測されたのであろう。

こうした自力修正過程が、74 分後に成果として現れ 1 発合格となったと考えるのは合理的だろう。

3.7.2 自力修正に成功し損ねた例

Fig.10 は自力修正で結果が出なかった例 (以下 B とする) である。この事例を先の成功例 Fig.8 (以下 A とする) と比較するとき、顕著な違いは失敗した語彙の多さと同じ失敗の多さである。例 A は our, ran



Fig.9 赤ペン添削画面 (2)



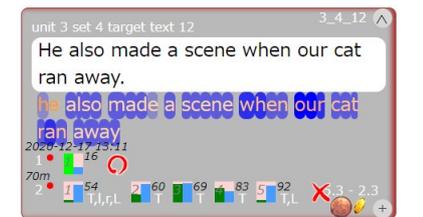
Fig.10 自力修正例 (2)

の2個だが例Bは he, made, a, scene, when, our, cat, ran, away の 9 個と 4.5 倍である。こうした失敗情報は Fig.10 の青楕円で囲った部分に背景色と文字色の濃淡で示されている。背景色青の濃淡はアルファベットごとについて音素レベルのエラーを示し、文字色赤の濃淡は語彙レベルのエラーを示している(この2種類のエラーについては後に触れる)。ほぼ全部の語彙や文字が彩色されていることから聞き手が理解に努めても意味不明となってしまうだろう。

次に個々の添削結果を示す正方形に注目する。特徴的エラー類型として he also があり、チェックされた she also が 11 回もあったとされる。正方形のアイコンに赤い枠がついている部分である。あわせて 12 度の添削で he also がうまく発声できたのは最後から 2 個目だけであった(この he/she エラーについては後に触れる)。時系列を見ると、1 回目は 13:13 に 1 度添削を受けたあとといったんこの回を終了し 50 秒後に再開して一気に 5 度おこなって全滅している。68 分後に 3 回目に入って失敗したあと再び終了し、35 秒後に 4 回目に入り一気に 5 回発声して全滅して終えている。2 回目までが前テストで 3 回目からが後テストである。経過と結果だけを見ると自力修正に失敗し続けたことになるが、正方形に埋め込まれた緑と青の棒グラフに注目すると、この受験者のポリシーと修正力が見えてくる。まず緑棒グラフに注目すると 2 回目は過半に満たないものが続くが、4 回目は過半を超えている。明らかに失敗ヶ所が減ったのである。そして右側の青棒グラフは 12 回すべてが天井まで伸びている。1 秒間に 3 単語以上の速さである。この受験者は、

とにかく一息にどんだんしゃべることを是としているようだ。

そこで今度は話速を遅くして対応した事例 Fig.11 を示



回目	1	2	3	4	5	
度目	1	1	2	3	4	5
正しさ	1	4	9	7	1	1
wps	1.1	2.6	2.8	2.9	1.7	2.5

Fig.11 自力修正例 (3)

す。各正方形の右側青棒グラフに長短の差があることが見て取れる。前テストの 1 回目に 1.1wps というとてもゆっくり発声して成功している。後テストでは 2.6wps と並の速さでうまく行かないことが続いたせいか 4 回目に 1.7、5 回目と 2.5 と調整したが結局失敗で終わっている。

3.7.3 自力修正に音声を利用した例

自力修正に音声を使った生徒も見受けられた。その一例を Fig.12 に示す。課題文を開いたときに流れるリード音声は希望すればいつでも何度でも聞き直すことができる。この補助音声を利用した記録は図中の赤い点である。各回の最初の赤点は必ず流れるので除外すると、この例

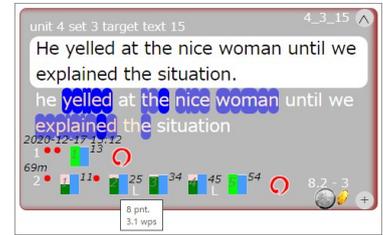


Fig.12 自力修正例 (4)

では 1 回目の最初に聞き直している。その後 2 回目の 2 度目を発声する前にもう一度聞き直している。よく見ると 1 回目より得点が 5 点から 8 点にアップしている。この得点は表面には記載されていないがマウスで正方形アイコンをポイントするとデータがポップアップするので読み取ることができる。この受験者の場合、音声を聞き直すことで調子を掴んだようで、3 度目 4 度目と高得点を得ながら 5 度目に合格していた。

3.7.4 到達度チャートと自力修正力の個人差

到達度チャートは、課題の達成状況を視覚化したもので、すでに Fig.3 に示している。このチャートは、スピーキング力向上を目的にスピーキングテストを活用する自力修正力を観察し、クラス指導や個人指導に役立てることを目的としている。ここでは 4 週目のテストでの到達度チャートを Fig.13 に紹介する。紙面の都合で 1 度目成功率が最大だった生徒、平均、最小だった生徒に絞った。最大と最小は 26 回以上の添削実施者から選んだ。1 度目成功率の最大の受験者と最小の受験者の差異は 5 倍程度もある一方で、2 度目以降の成功率で見ると、1 度目の成功率が低い方が伸び率が大きいと見受けられた。個人差はあるものの、今回のスピーキングテスト方式では自力修正力を発揮した生徒が多いことを示唆している。



Fig.13 比率で見た到達度チャート/平均と最大最小

4. 事後研究/反省と改善

4.1 『赤ペンミス』の原因は ASR か受験者か

5 週間に及ぶスピーキングテストを終えておこなったアンケートに「NatTos の精度が悪く、何度 he と発音しても she になって苦戦した」との感想があった。本人の音声も英語母語話者が聞いたところ、he ではなく she と聞こえ、問題は ASR ではなく、話し手の音声だと判明した。この出来事はスピーキングテスト開発者と教師に 2 つのことを教えてくれた。

- (1) ASR は発展途上でもあることから、ASR の判定だけでなく、実音声の聞き取りが欠かせない
- (2) データマイニングは、現場情報と照合しながらの補正作業が欠かせない

またこの出来事は教育研究者に新しい知見を与えてくれた。それが he/she エラーである。he/she エラーとは HE 音が SHE と聞き取られる事象である。そこで課題文 69 本の内 8 本に he あり、そのうち文頭付近に he がある 6 本について解析した。該当する課題文を Fig.14、解析結果を Fig.15 に示す。

#	解析した課題文一覧
1	But he thought that we had paid for a pizza hat.
2	He also made a scene when our cat ran away.
3	He yelled at the nice woman until we explained the situation.
4	He was going to talk about his research with his colleagues.
5	So, he didn't know any of the men and women there.
6	Tomorrow, he is going to have to apologize to the people that he let down.

Fig.14

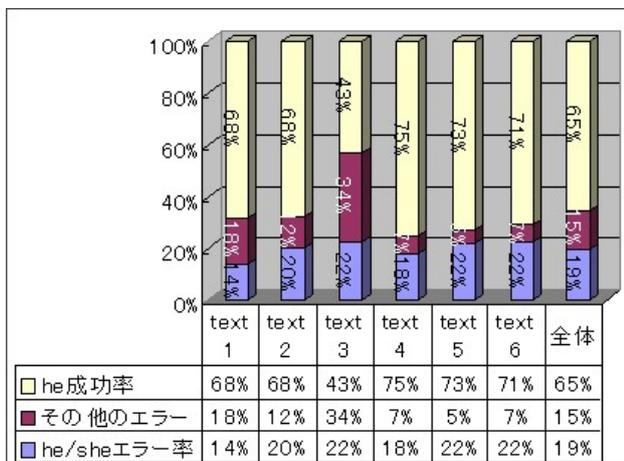


Fig.15

解析した添削 3,551 本のうち 689 本に he/she エラーが観察された。Fig.15 の右端にある全体の傾向に見られる通り、he/she エラー率 19%はその他のエラー率 15%を上回っている。これは he/she エラーがあるばかりに合格し損ねた受験者が多数存在することを示唆している。そこで受験者ごとに he/she エラーの出現の様子を調べた。

4.2 母語干渉を示唆した he/she エラー

調査をできるだけ正確にするために 6 課題文すべてを実施した (添削を受けた) 受験者のみに絞り込んだところ、対象は 73 名 3,002 件となった。結果を Fig.16 に示す。横軸はある受験者の全エラーのうち占める he/she エラーの割合を 10%刻みで区分して、縦軸はその指定範囲に該当する受験者の人数である。例えば右端にある小さな赤い矩形は he/she 転移率 (HE の発声 SHE と認識されてしまう割合) が 80%より大きく 90%以下である人数が 2 人であることを示している。この 2 人は極度に he/she 転移してしまうことが分かる。左端の長めの矩形緑は he/she エラーが 0%である人数が 10 人であることを示している。この 10 人は今回のスピーキングテストで he/she 転移を示さなかった生徒である。

グラフ全体を俯瞰するとエラー率が 10%以下である受験者が 18 人と最大で、ここを境に左右に減少している。この 18 人は he/she 転移をまれにしか起こさないとわかる。図の矩形赤の範囲 (he/she エラー率が 50%より大きい受験者) の 6 人は、he/she 転移率が 5 割を超えているので、2 回に 1 回以上の割で頻繁に he/she 転移を起こしている。

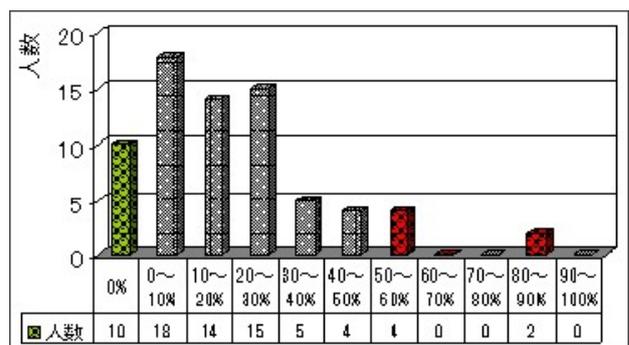


Fig.16

以上のことから、HE 音の発声時に S 音が交じり SHE と認識される弱点 (he/she 転移) を持つ日本人が一定数存在することが見て取れた。これは、「日比谷」と「渋谷」を混同する日本語の特性が英語音声に投影された母語干渉であることを示唆している。

4.3 見過ごされた he/she エラー

スピーキングツール NatTos には語彙クラウドと呼ぶ弱点把握システムがある。添削で赤が入った語彙を頻度順に文字の大小で視覚化したデザインだ。生徒が発声に失敗した語彙を担当教師が瞬時に判断してクラス指導に活かすことを目的として語彙クラウドを設置している。統計的には頻度順に一覧表にして番号を振っておけばよいのだが、文字の大小で頻度を表した方が次の点で優れている。

- (1) 大きな字ほど自然に素早く目に入る
- (2) 画面をスクロールしなくても全体を見通せる

しかし、1人の生徒が he/she エラーを指摘するまでに気づけなかった。つまり語彙クラウドに欠陥があった。旧版と改善された新版を Fig.17 に示す。

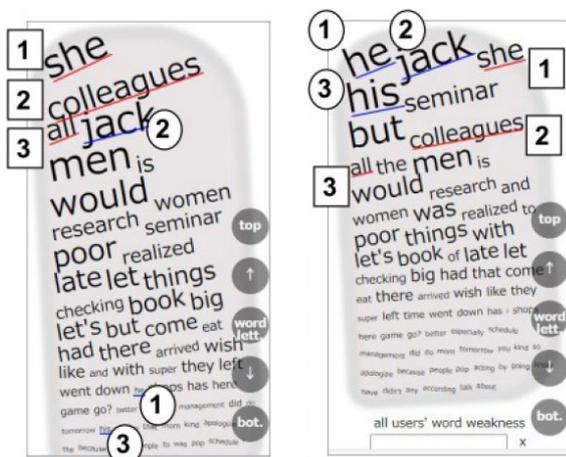


Fig.17 旧版の語彙クラウドと新版 (右)

新旧で語彙 he の位置がどう変わったのかに注目する。まず四角で囲った 1,2,3 で示した赤線語彙が新版(右)では少し順位が下がって大きさも少し小さくなっている。それに対し旧版では小さくて見えないくらいだった丸 1 青線の he と丸 3 の his が一躍トップに躍り出ているのがわかる。これは急激な変化である。何が変わったのだろうか。

4.4 新版で追加された評価基準の仕組み

欠陥の由来は双方向赤ペン添削の仕組みにあった。

双方向添削は話し手と聞き手の両方でどこがずれているかを文字(音素)単位で視覚化したもの (Fig.18) である。このとき

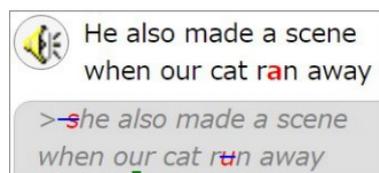


Fig.18 he/she エラーのある添削例

ran/run エラーは課題文と結果文に記録されるが、he/she エラーは結果文にしか記録されない。つまり旧版では話し手エラーを課題文の記録だけを見て計算していたのだ。新版ではこの点を修正し聞き手に記録された she のエラー認定を話し手の he の記録に反映させる仕組みを追加した。この影響は Fig.11 の he の文字色が赤っぽくなっているところにも現れている。こうした改良の結果が反映されたことで、he と his が一気に浮上してきたことになった。

5. 考察

本研究について、音声教育、教師、生徒そしてデータマイニングの視点から考察する。

5.1 双方向スピーキング添削の効果

教育的に意義が大きいことは、自分の発声がどのように聞き取られたかを 3 秒で知った学習者が自力で修正し、通じる音声を試行錯誤することをうながせた(動機づけた)ことである。こうした振る舞いを可能にした仕掛けのひとつが双方向添削で、大きな役割を果たした。このように考察できた根拠には、発話修正力が視覚化がある。

従来のスピーキング自動添削アプリは片方向添削 (Fig.5) であった。今回採用したツールは双方向添削である。課題文と結果文の双方を赤ペン添削した (Fig.4, Fig.18) ことでデータが現実を反映しやすくなった。受験者の発話行動をかなり詳しく再現することができるようになったのだ。具体的な効果としては、学習者の発話修正力を視覚化 (Fig.3, Fig.6, Fig.8, Fig.10, Fig.11, Fig.12, Fig.13) できたことだ。

5.2 評価における時間情報の大切さ

従来のテストは主に優劣を点数化することを目的としていた学力テストであった。今回のツールでは学力の伸びに関わる修正力も測定可能になった。これを可能にしたのは時間情報を視覚化したデータマイニング (Fig.3, Fig.6 など) である。特に発声添削 1 サイクルごとに経過時間を把握できるようにしたことで、受験者の意図も推測できるようになった。

5.3 時間情報の埋め込み手法

今回の添削ツールでは複数の方法で時間情報を埋め込んでいる。

- (1) 日時をそのままの数値で順に表示 (Fig.2)
- (2) 左から右への視線移動で時系列を示唆 (Fig.3, Fig.10 の正方形アイコン)
- (3) 時系列アイコンの肩に経過時間を表示 (Fig.10 正方形などのアイコン)

(4) 話速 (wps) を棒グラフで正方形アイコンに埋め込んだ (Fig.10 など)

これらは受験者の行動とその原因や意図を推測する貴重な根拠となった。

5.4 文字や図形の大きさと色彩を活用

今回のツールには大きさや色彩で、利用者が自然に注目し、視線が移動するような工夫がある。主なものを列挙する。いずれも狭い画面内でスクロールしなくても一瞥して必要な情報に目が向くような工夫である。大きさや色彩はアフォーダンスの考え(説明されなくても意味を推定できる)に従っている。

- (1) 語彙クラウド (Fig.17) では文字の大きさでエラー頻度の多さを示している。配列はバラバラではなく、左上から右下に視線移動しながら閲覧できるようになっている。その上でこの流れに逆らうように大きなサイズの語彙を配置し気づきを促している。
- (2) 到達度チャート (Fig.13) では失敗事例のブロックを右端に赤系統の色彩とし、成功事例は1度目を緑、2回目以降を黄系統の彩色として左から右へと配置してある。
- (3) 学習履歴一覧表の課題文添削欄 (Fig.10 の青丸内) では失敗の多さを、アルファベット単位では青系統背景色の濃淡で、語彙単位では赤系統文字色の濃淡で表現している。

5.5 仮説「he/she エラーは母語干渉」

3 クラス 100 人ほどの大学生によるスピーキングテスト履歴から、少なくない受験者の HE 音/hi:/が ASR により SHE 音/ji:/と認識されていたことが判明した (4.1)。HE 音が英語と日本語で異なり「英語の調音点はより奥にあるが、日本語の場合は、硬口蓋にあるので、英語よりずっと前」[2]の/c/であることに依ると思われる。しかし HE 音が SHE 音と認識される現象を弱点とする学術文献は見かけなかった。he/she 転移はこれまで注目されてこなかったのかも知れない。

日本人の中には H 音が SH 音になってしまう人たちが少なからず存在する。例えば「日比谷 HIBIYA」と言ったつもりが相手には「渋谷 SHIBUYA」と聞こえてしまう現象である。「質屋」を「ひちや」と表記する地域もある。声門摩擦音/h/の調音点が英語よりも前方の硬口蓋にあって歯茎摩擦音/s/を生じていて、それを ASR が/j/と解釈しているのかもしれない。こうしたことから本人は he/she 転移に気付かないが、母語話者に he/she エラーと認識される発声(調音器官の動かし方や呼気圧の制御の癖)になっていると推測される。これは母語干渉[3]の一種と思慮された。

本研究 (Fig.15, Fig.16) は、仮説「英語発話において、HE 音に S 音が交じり SHE と認識される弱点を持つ日本人が一定数存在するのは母語干渉である」を支持する結果となった。

5.6 オンライン下のスピーキング指導で力

従来のクラス対面発音指導では、一人の教員が学生全員の発音を丁寧に聞いてまわって個別にアドバイスするだけでも無理があった。オンライン授業となつては聞いてまわれなくなり、なおさら困難な状態が予測された。ASR と自動添削による機械化でも、生徒音声の直接聴取に不都合であることに変わりがないとは言え、赤ペンで添削された生徒音声を詳細かつ瞬時に閲覧できる上、語彙表現の音声的弱点を素早く察知できることは、オンライン下のスピーキング授業で力を発揮した。

6. 課題

6.1 自力修正力を補助する学習法開発へ

従来のスピーキング教育が行き詰まっていることは、日本人の英語力の弱さやスピーキングテスト導入の延期などにあらわれている。生徒の自力修正力を補助する学習法と教授法の開発が求められている。

6.2 ビッグデータ収集と解析への道のり

今回の研究では、データマイニングの視覚図を示しながらそこから知見を得る手法を多用してきた。今後データが 100 万、1 億と増大してきたとき、人による観察と直観に頼り続けることには限界があるだろう。人工知能を使って解析し、たとえば he/she エラーのような知見を素早く見つける手法を取り入れていきたい。これが実現すれば、今まで知られていない弱点や教授法・学習法が見つかる期待できる。多くの教育者と研究者の協力を望んでいる。

参考文献

- [1] Benesse, 『進研ゼミ高 1 講座』が 7 月からスピーキングアプリを提供 1 回 1 分のレッスンで毎日楽しく「発話」トレーニング 2020 年度大学入試改革に向けて、英語スピーキング対策を強化、<https://blog.benesse.ne.jp/bh/ja/news/20180712release.pdf>, 2018
- [2] 続三義, 英語と日本語の音声対照研究—日本語発音教育の角度から—, 東洋大学「経済論集」40 巻 2 号 2015, <http://id.nii.ac.jp/1060/00006943/>
- [3] 丹羽牧代, 丹羽卓, 地蔵繁範, 母語干渉とうまく付き合おう, 彩流社, 2019