

英語音声認識アプリを用いた自律的発音学習

Using an English Speech Recognition App for Autonomous Pronunciation Learning

中西のりこ

(要約)

近年の英語教育では「話すこと」に重点を置いた能力の向上が求められるが、学習者ごとに異なった発話をするアウトプット活動においては、一斉にモデル音声を復唱するという指導法が通用しない。そこで本研究では、個々の学習者が自律的に自分の発音矯正を行うことができるようになることを目指し、現在 ver.03 が web 上で一般公開されている「音素カウンター」の分析機能と音声認識アプリの機能を併用した実践を行った。本稿で紹介する2つの実践で用いた分析手法を自動化してシステムに組み込み、今後、「音素カウンター」ver.04 の開発に取り組みたい。

(Abstract)

Recently, English education pays a lot of attention to ways to improve learners' speaking skills. However, tasks such as choral repetition in the classroom are at times inappropriate, given every learner's utterance may be different. The present study aims to help learners correct their own English pronunciation by using a speech recognition app, together with the Phoneme Counter, of which the third version is currently accessible online. The analytical procedures used for the activities reported in this paper will be integrated into the system of the Phoneme Counter (version four) in the near future.

キーワード：英語発音，自律的学習，伝わりやすさ，音声認識，音素数算システム

Key words: English pronunciation, autonomous learning, intelligibility, speech recognition, phoneme counting system

1. はじめに

1-1. 研究の背景

外国語学習において音声面の指導は不可欠であるが、実際の教育現場で発音をはじめとする音声指導を実践するには、多くの制約がある。文法や読解の指導に時間を割かれ音声指導にまで手が回らないという時間的制約や、大規模クラスでは個々の学習者の発話に一人の教員が対応できないという教室環境の制約、学習者が恥ずかしがって声を出さないという情意的な制約、教員自身が英語音声について十分な知識を備えていないという指導者側の問題などが挙げられるが、それ以前に、「何を目標にして指導をするのか」という到達目標が定まっていないという問題がある。

近年の英語教育現場では、英語母語話者のような発音にこだわらず「通じやすい発音 (intelligible pronunciation)」を英語発音習得の目標とするという考え方 (Celce-Murcia, Brinton, Goodwin, & Griner, 2010; Jenkins, 2000, Kachru, 1982 など) が主流となっている。Munro and Derwing (1995) では「通じやすさ (intelligibility)」を “the extent to which a listener actually understands an utterance” と定義し, comprehensibility (a listener’s perception of how difficult it is to understand an utterance) や accentedness (a listener’s perception of how different a speaker’s accent is from that of the L1 community) と区別している。ここで問題となるのが, intelligibility を定義するにあたって listener (聞き手) の存在が必要となるということである。つまり「通じやすさ」を測る尺度に聞き手の理解力という要素が含まれるため, 例えば教室内で日本語を母語とする学習者が話し手と聞き手となって英語で会話をする場合, カタカナ発音が最も「通じやすい」発音になってしまう。一方で, 教室外の現実世界での会話では, 話し手・聞き手の両者が日本語母語話者であるケースは稀であるだけでなく, 聞き手が英語母語話者であるとも限らない。非英語母語話者間で, お互いの母語が異なると「通じやすさ」の度合いは下がる (Jenkins, 2000)。そのため, 教室内では難なく通じていた発音が現実世界では全く通用しないということが起こり得る。つまり, 「通じやすさ」を基準とした英語発音習得を目指すのであれば, まずは, 聞き手がどのような言語的背景を持った人であるのかを想定する必要がある。

そこで本研究では, 「通じやすさ」に影響を及ぼす聞き手の存在を音声認識アプリに固定することにより, 話し手である学習者の英語発音がどれほど認識されやすいかの尺度とすることを試みた。人と人とのコミュニケーション活動における「通じやすさ」の判断基準を音声認識アプリのような機械に委ねることに対しては批判的な意見もある。例えば Derwing & Munro (2005) は, 音声認識によって得られた結果を解釈して指導に役立てるための音声学的知識が十分でない指導者が多い上に, 音声認識システムが人間の聞き手と同様の反応を示すとは言い難いとして懐疑的な立場を取っている。一方で, 音声認識の精度は日進月歩で確実に上がっている (Chen & Chen, 2018; Ehsani & Knodt, 1998; Liaw, 2018; 峯松・鎌田・朝川・鈴木・牧野・西村 & 広瀬, 2011)。さらに, 音声認識による文字入力の方がキーボード入力よりも速いという実験結果を示した研究 (Ruan, Wobbrock,

Liou, Ng, & Landay, 2016) もあることから, 将来的には, 自分の音声を機械に認識させるスキルがスピーキング活動時のみならずライティング活動時にも必要とされる時代が来ることが予想される。本研究ではこのような背景から, 2016年から開発中であった「音素カウンター」というシステムに音声認識機能を追加し, 学習者が自分自身の英語発音がどのような誤解を生みやすいかを認識し「通じやすい発音」を目指すシステムの構築を試みた。この試みは, 学習者ごとに異なった発話をするアウトプット活動において, 一斉にモデル音声を復唱するという指導法が通用しないという問題を解決する自律的発音学習の一助ともなり得るものである。

1-2. 音素カウンター ver.03 までの軌跡

音素カウンターとは, web 上で入力された英文を簡略化した発音記号に変換し, 英文に含まれる音素数を算出し出力するシステムである (中西, 2016a)。2016年に ver.01 が web 上で無料一般公開されて以来, いくつかの機能が追加され, 2018年2月現在, ver.03 が公開されている (Nakanishi, 2017; 中西, 2017; Nakanishi, 2018)。図1に, ver.03 時点で利用可能な機能を示す。

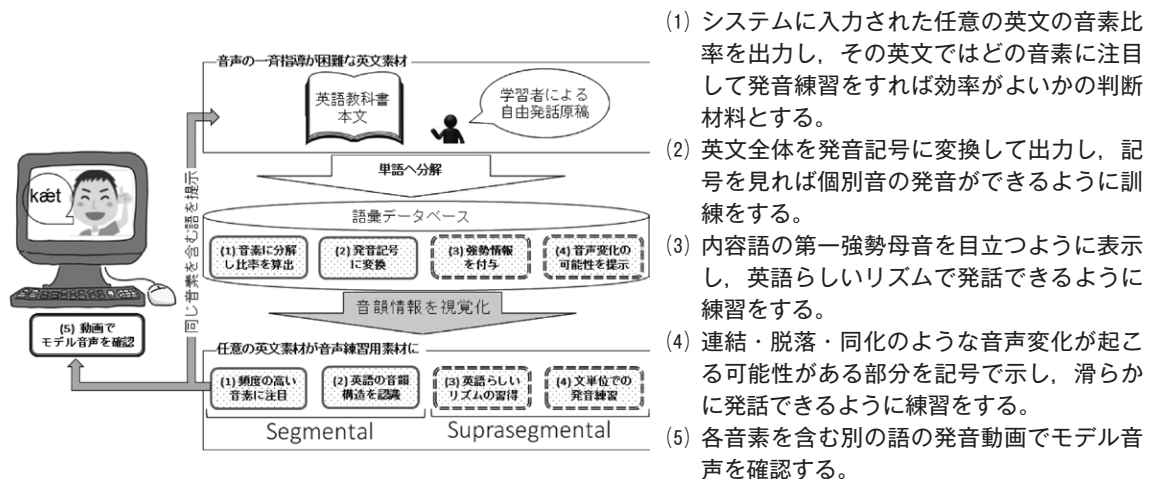


図1. 音素カウンター ver.03 に搭載されている機能 (2018年2月現在)

2016年の ver.01 公開段階では図1の(1), (2)の機能からスタートし, 学習者が発音記号に慣れ親しむ手助けをすることを目標とした (中西, 2015)。2017年前半のバージョンアップでは, (1), (2)の分節音 (segmental) に加えて, リズムや音声変化といった超分節音 (suprasegmental) にも着目し, (3)強勢記号や(4)連結・脱落・同化の記号を出力画面で提示する機能を追加した。さらに2017年後半には, 標準的な北米英語話者による発音モデルを示す動画へジャンプする機能(5)を搭載し, 音素カウンター ver.03 とした。これらの機能の利用マニュアルと, このシステムで使用する発音記号表記の凡例は, システム上からダウンロード可能である (Nakanishi, 2017)。

1-3. 研究の目的

現在、音素カウンターは ver.04 へのバージョンアップに向けて準備を進めている（中西，2016b；中西，2018a；中西，2018b）。上記(1)では、キーボード入力された英文テキスト中に含まれる音素のうち比率の高いものに重点をおいて練習するよう学習者に促すことを想定していた。しかし実際には、各音素の出現比率だけでなく、個々の学習者がどの音素の発音を苦手としているのかを考慮し、より効果的な練習のポイントを提示することが望ましい。そこで ver.04 では音素カウンターを2画面構成とし、iOS に搭載されている音声認識アプリを用いて学習者が入力した英文テキストを、従来通りキーボード入力した同じテキストと比較する仕組みの構築に着手している。

本研究ではこのような仕組みの構築に先駆けて試みた2つの事例を紹介する。どちらの事例も、日本語母語話者のみが受講する一斉授業内で起こりがちな問題の解決につながるものである。これらの事例のために利用した分析手法をシステムに組み込み自動化すれば、学習者が自分自身の英語発音の問題点に気づく手助けをする自律的発音練習システムとして機能することが期待される。なお、以下で紹介する事例の分析対象となる音声や音声認識原稿は、筆者が担当する英語科目を履修する大学生が授業外課題として提出したものである。個人情報は一切公表しないことを前提に、教育・研究のために画像・映像・音声を公表する旨、書面による承諾書を得た。さらに研究の目的と概要を説明し、研究に協力しなくても成績評価などで不利益をこうむることはないことを明確にするなど、倫理的配慮を行った。

2. 事例1

2-1. 背景

筆者が担当する「英語会話 I, II」では年間60回×90分の授業を通して、「聞くこと」「話すこと」に重点を置いた活動を行っている。受講生は4回もしくは6回の授業ごとに与えられたトピックについてインタラクションや研究発表を行い、英語音声によるコミュニケーション活動を行う。図2は、この授業で実施した Versant[®] Speaking Test のスキルごとの平均点の推移である（2015年度生 $n = 78$ ，2016年度生 $n = 90$ ，2017年度生 $n = 130$ 。2015・2016年度生1年春および2015年度生2年春はテスト未実施のためデータなし）。Versant[®] Speaking Test はリスニング能力とスピーキング能力を測定するテストで、総合スコアに加えて文章構文・語彙、流暢さ・発音のスキルごとの得点が20点から80点の幅で採点される。CEFR 換算の目安は、A1 レベル（26～35点）、A2 レベル（36～46点）である（Nikkei, 2017）。図2の4つのグラフを比較すると、3学年ともに文章構文・語彙の平均点と流暢さ・発音の平均点との間に開きが見られる。文章構文と語彙については、1年冬の受験時にはどの年度も平均点がA2レベルに達しているが、流暢さと発音については、A1レベルに留まっている。

この3学年を1年次に担当した年間60コマ×90分の授業実践では、各トピックについ

での説明や活動の指示を聞き取り理解し、関連する語彙を駆使して文章を組み立て、自分の考えを受講者同士で伝え合うという力が回を重ねるごとに伸びていることが実感された。つまり「何を聞き取り何を伝えるか」という面では授業活動の効果が文章構文および語彙の得点に表れたと考えられる。一方で、発音及び流暢さの得点の伸びが思わしくなかったことは、「どのように伝えるか」という面において大きな課題が残されていることを示している。このことの一因は、授業内で音声面の指導が行き届かなかったことだけではなく、教室内では話し手・聞き手の両者が同じ日本語の母語話者であるため、カタカナ発音が「通じやすい」発音として化石化 (fossilization) していることにあると考えられる。このような問題が生じた背景を以下に挙げる。

- 受講者間のインタラクションでは、日本語の影響を受けた英語発音の方が理解されやすい。
- 受講者ごとに発話内容が異なるため、モデル音声を用意できない。
- PC 教室の都合上、2クラス合同の50～60人での一斉授業のため、個々の学習者の発話への対応に教員の手が回らない。
- 受講者間フィードバックの際に「流暢さ・発音」という項目を設けたが、英語の音韻体系そのものの基礎知識が学生にないため、あいまいな指摘しかできない。

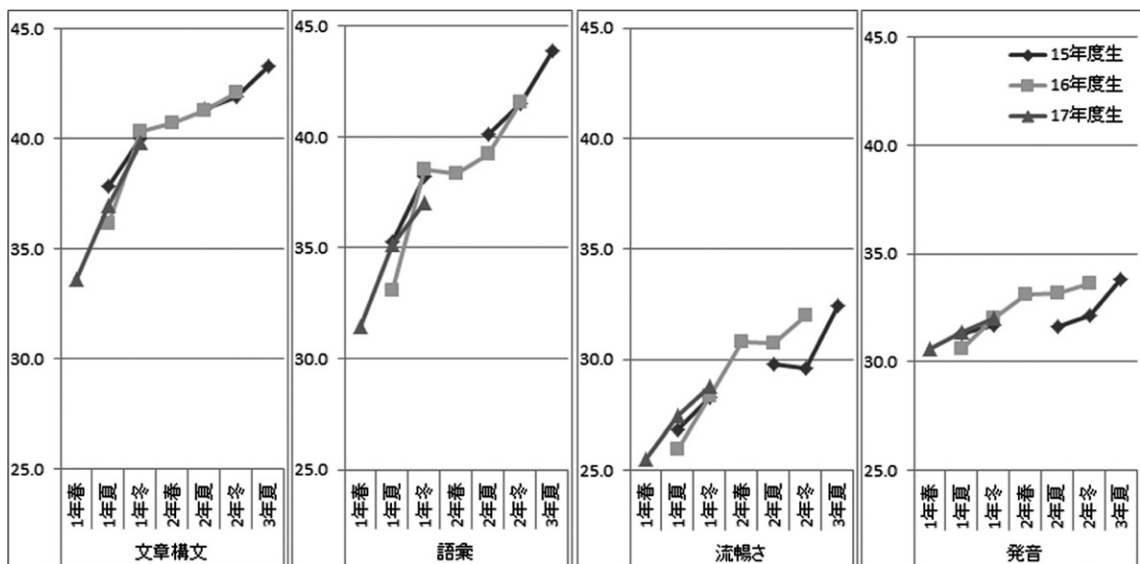


図2. Versant[®] Speaking Test 平均点の推移

2-2. 方法

2017年度冬のVersant[®] Speaking Test終了後、1年生は1年間のスコアの伸びを確認し授業活動の振り返りを行った。その後、自分が伸ばしたいと考えるスキルによって、eラーニングプログラム (L-Interface, 2017) のコンテンツもしくは音素カウンターを利用した発音練習のどちらかに取り組んだ。eラーニングプログラムには語彙・文法・リーディ

ング・リスニングのトレーニング用の教材が用意されているが、大半の受講生が音素カウンターを利用した発音練習に取り組むことを選んだ。前述のように、自分自身の Versant[®] Speaking Test の結果を見て、発音や流暢さに問題があると反省した受講生が多かったと考えられる。音素カウンターを利用した発音練習は、以下の手順で行われた。なお、この日には音素カウンター ver.04 に向けての試行版として、英文入力画面を2画面構成にしたwebサイトを用意した。

- (1) 教材配信・課題提出用の web サイトに予め用意していたイソップ寓話 *The North Wind and the Sun* (International Phonetic Association, 1999) のスクリプトをダウンロードする。
- (2) 自分のスマートフォンから音素カウンターにアクセスし、左側の画面に(1)をコピーペーストする。
- (3) 右側の画面をタップし、上記のスクリプトを音声認識アプリを利用して入力する。
- (4) (2)と(3)それぞれの英文に含まれる音素数を比較する画面を開き、どの音素の認識率が低いかを把握する。
- (5) (4)で気になった音素の発音記号を音素数分析画面でタップし、モデル音声を聴き練習する。

2-3. 結果

上記の活動中、教員は教室内を回り受講生のサポートを行った。ここでは、その際に質問があった受講生のうち1人の事例を紹介する。この受講生が受験した Versant[®] Speaking Test スコアは、(春・夏・冬)の順に、文章構文 (27・43・47)、語彙 (35・38・44)、流暢さ (28・30・34)、発音 (34・32・31) であった。文章構文と語彙のスコアは順調に伸びたが、流暢さと発音のスコアの伸びが思わしくなかったという典型的な例である。

表1は、上記(3)の作業で入力した音声認識スクリプト(左側)と(2)で用いた英文スクリプト(右側)である。(3)で入力されるスクリプトには改行が含まれないが、表1では見やすさのために改行の処理を施した。大文字や句読点も音声認識では反映されないが、音素数には影響を及ぼさないためそのままの形で分析対象とした。また、語中で /ɔ:r/ と認識されるべき部分(6か所)を網かけ、語末で /θ/ と認識されるべき部分(4か所)を下線で示した。これについては後述する。

表1. *The North Wind and the Sun* を音声認識入力したスクリプトとコピーペーストしたスクリプト

音声認識入力スクリプト	コピーペーストスクリプト
Dad knows window and Gisele what are disputing with you it's a stronger when I turn off the router came along wrapped it in Tai wan o'clock	The North Wind and the Sun were disputing which was the stronger, when a traveler came along wrapped in a warm cloak.
they agreed that the one who pass succeed in making it's a problem to take his crock all the stupid calls you today to get stronger down to the other	They agreed that the one who first succeeded in making the traveler take his cloak off should be considered stronger than the other.
damn the nose window put you at had ads he could but it's up more he proved some more grocery diesel trouble report heat crock around him and that I trust di agnose window Gabe after attempt	Then the North Wind blew as hard as he could, but the more he blew the more closely did the traveler fold his cloak around him; and at last the North Wind gave up the attempt.
Jan does sound shined out will Morty and image get to the counter that Robin Road took off kids crock and the soul do you know switch window was over to conquests up the song what's a stronger office at the two	Then the Sun shined out warmly , and immediately the traveler took off his cloak. And so the North Wind was obliged to confess that the Sun was the stronger of the two.

表2は、上記の2種類のスクリプト中に出現した音素のうち、個別音の発音に問題があると考えられる母音 /oo/, /ɔ:r/ と子音 /θ/ (表1の網掛けおよび下線部分) の数である。これは、表1左側の音声認識入力スクリプト中の音素数を分子とし、右側のコピーペーストスクリプト中の音素数を分母とした、音素認識比率に等しい。表中最右列の強勢とは、/ɔ:r/の場合、内容語で語強勢を受ける母音として出現した数、/θ/の場合、内容語で語強勢を受ける母音の直前の子音として出現した数である。例えば /ɔ:r/ という母音はコピーペーストのスクリプトでは全体で8回出現したが、音声認識原稿では3回しか認識されず、それら3回の内訳は語中1回、語末2回であり、3回ともに強勢が置かれる /ɔ:r/ であった。また、母音 total, 子音 total の認識比率を見ると、この受講生が音声認識入力したスクリプトはコピーペーストしたスクリプトより多くの音素を含んでいたことが分かる。

表2. *The North Wind and the Sun* の音声認識スクリプトとコピーペーストスクリプト中の音素数 (n)

音素	全体	語頭	語中	語末	強勢
/oo/	12/7	1/0	6/6	5/1	8/7
/ɔ:r/	3/8	0/0	1/6	2/2	3/8
母音 total	165/144	35/23	89/79	41/42	81/65
/θ/	0/4	0/0	0/0	0/4	0/0
子音 total	266/232	96/90	86/73	84/69	68/59

受講生は、音素カウンターの分析画面に表示されるこの比率を参考に、自分の発音のどこに問題があるのかを確認した。この事例で紹介した受講生の場合、本来のスクリプトでは6回認識されるべき語中の /ɔ:r/ が1回 (1/6 = 17%)、4回認識されるべき語末の /θ/ が0回 (0/4 = 0%) と低い認識率となっている。音素カウンターの分析画面では、注目したい音素の出現数の数字をクリックもしくはタップすると、元のスクリプトのどこにその音素が含まれていたかがハイライトされる。受講生はこの機能を利用して、North という語の発音に問題があったことに気付くことができた。表1で示されるように上記(3)の作業で入力した音声認識スクリプト側では、North は knows, nose, diagnose, know + s(witch) と認識されていた。つまり、/ɔ:r/ を /oo/, /θ/ を /s/ や /z/ と置き換えて発音していたと考えられる。この受講生は該当する音素の発音記号をクリックし、/ɔ:r/, /θ/ をそれぞれ含む別の語のモデル発音動画を視聴しながら発音練習をし、再度(3)の手順に戻りスクリプトを音声認識で入力したところ、North という語の認識率が向上した。ここではこの受講生が最初に気付いた /ɔ:r/ と /θ/ の例のみを取り上げたが、2-2.で紹介した手順を繰り返すことにより、別の音素についても1つずつ問題を解決していくことができると考えられる。

この事例は、受講者間の通常のインタラクションではカタカナ発音でも通じてしまうという教室環境において、スマートフォンの音声認識機能と音素カウンターを使い学習者が自分の発音がどう認識されるかを確認し、特定の音素の発音の仕方を見直すことができたという事例である。

3. 事例2

3-1. 背景

兵庫県が主催する英語プレゼンテーションコンテストに、数名の学生が参加した。2017年10月末に提出した発表概要や資料を元に11月半ばに事前審査の結果が発表され、審査を通過した学生が12月初旬のコンテストに向けて準備を進めるというスケジュールに沿って、出場者は「より多くの留学生を兵庫県に惹きつけるための施策の提案」を行った。この事例において英語音声指導をする際に問題となった背景を以下に挙げる。

- 授業外での活動のため、対面での指導の時間を取ることができない。
- 発表の構成や内容の準備に大半の時間を費やし、発表方法の練習にまで手が回らない。
- 発表原稿が完成するまで発音練習に着手できない。

3-2. 方法

コンテスト出場者からの発表準備に関する相談は基本的にすべて web 上のファイル保存サーバー及びメールのやり取りを通して行った。ここで問題となったのが、発表直前の音声指導の方法である。出場者が各自行った予行演習の様子を撮影した動画ファイルや発表

原稿を読み上げた音声ファイルは容量が大きく送受信が容易ではない上に、全ての出場者の動画や音声を筆者のPCで再生し問題点を指摘する時間的余裕も限られていた。

そこで、各出場者に、以下の手順で発表原稿の音声認識結果を筆者へ送信するよう伝えた。

- (1) 自分のスマートフォンでメール送信画面を開く。
- (2) 音声認識アプリを用いて、発表原稿をメール本文に音声認識入力し、筆者へ送信する。
- (3) その際に、キーボード入力した発表原稿のファイルも添付して送信する。

表3は、出場者の一人から送信されてきた音声認識入力原稿（左側）とキーボード入力原稿（右側）の一部抜粋である。実際に送信されてきた音声認識結果は改行が含まれないひとかたまりの文書であったが、表3では見やすさのために改行の処理を施した。大文字や句読点も音声認識では反映されないが、音素数には影響を及ぼさないためそのままの形で分析対象とした。また、「Tatsuno」「Hyogo」「Usukuchi」という固有名詞や日本語は当然のことながら認識されていない。これは発表者の日本語発音に起因するものではなく、音声認識アプリの英語辞書にこれらの語が含まれていないためである。このような語はオーディエンスにとって聞き取りにくい語であることを伝え、パワーポイント画面を指差しながらゆっくりはっきりと発音するよう指導した。

表3. コンテスト出場者の発表原稿（抜粋）

音声認識入力原稿	キーボード入力原稿
This brought it combines friend and food that is so so so deLisa why with two so souls in his 80s traditional Japanese seasoning and do you know that's not equal go is famous for swizzles it has about 400 year history And one of the biggest so so's producers in Japan especially Sue could you Toys "R" Us	This project combines "Friends" and "Food". That is through soy sauce. The reason why we chose soy sauce is: It is traditional Japanese seasoning. And do you know Tatsuno in Hyogo is famous for 'soy-sauce'? It has about 400-year history. And, one of the biggest 'soy sauce' producers in Japan. Especially, "Usukuchi soy-sauce"

筆者の側では、こうして受信した音声認識原稿（総語数 360 語）と、同時に受信したキーボード入力原稿（総語数 327 語）それぞれに含まれる音素比率を音素カウンターを用いて算出・比較し、留意すべき発音のポイントを整理し、出場者に伝えた。

3-3. 結果

表4は、上記の2種類の原稿中に出現した音素の数をいくつか抜粋したものである。各音素の出現数の表記の方法は、表2と同様である。

表4. 音声認識入力とキーボード入力の発表原稿中の音素数 (n)

音素	全体	語頭	語中	語末	強勢
/i/	29/26	6/6	6/7	17/13	0/0
/ə/	98/100	32/33	46/51	20/16	0/0
/ɔɪ/	4/13	0/0	3/0	1/13	4/13
/oo/	31/11	0/0	8/5	23/6	28/10
母音 total	451/433	95/85	249/262	107/86	216/201
/ð/	27/27	23/23	1/3	3/1	9/8
/z/	32/32	0/0	4/6	28/26	1/3
/r/	19/28	0/3	19/25	0/0	12/19
子音 total	688/694	244/226	231/255	213/213	197/188

まず、母音 total の認識比率に注目すると、語末母音の比率が $107/86 \div 124\%$ と高い数値を示している。キーボード入力原稿では6回しか出現しない /oo/ が23回も音声認識されていることが原因の一つと考えられるため、音素カウンターで確認したところ、音声認識原稿には“so”という語が何度も出現していた。このことについては後述する。次に頻度が高かった語末母音は /i/ ($17/13 \div 131\%$) と /ə/ ($20/16 \div 125\%$) であった。語末の母音挿入の可能性は否めないものの、これらはどちらも弱形の母音であることから、強く耳に残るほどの挿入ではないと判断した。また、語中の母音、子音 total とともに90%から110%の範囲に収まっていることから、この発表者については、母音挿入という問題を指摘することは保留とした。

次に、母音全体、子音全体の認識率のうち数値が低いものに注目し、母音では /ɔɪ/ ($4/13 \div 31\%$)、子音では /r/ ($19/28 \div 68\%$) の音を含む語を音素カウンターで確認した。/ɔɪ/ はキーボード入力原稿中12回出現した“soy sauce”の“soy”が壊滅的で、ほとんどが“so so”と認識されていた。上述の /oo/ の認識率が極端に高かったのはこのためである。さらに、/r/ は特に語頭での認識率がゼロとなっている。/r/ を語頭に持つ語をキーボード入力原稿中で検索したところ“result”, “reason”, “right”という語が1つも認識されていないことが分かった。

最後に、 $1/3 \div 33\%$ という低い数値を示した語中の /ð/ と強勢が置かれる音節の頭の /z/ について確認した。その結果、どちらも本来のスク립ト中の出現頻度は3回で、語中の /ð/ はすべて“together”, 強勢音節頭の /z/ は“result”1回と“example”2回であった。“result”は上述の /r/ についての指摘で改善が期待でき、“together”, “example”については1回ずつは認識されていたため、これらの発音については指摘を保留とした。

このような分析結果を受け、この出場者には /ɔɪ/ と /r/ を含む語の発音を指摘するメッセージを送信することにした。図3は、その際のメッセージのやり取りの抜粋である。23時43分に送信した指摘に対して、10分後の23時53分には発音の修正ができた旨の返信が届いたことが分かる。

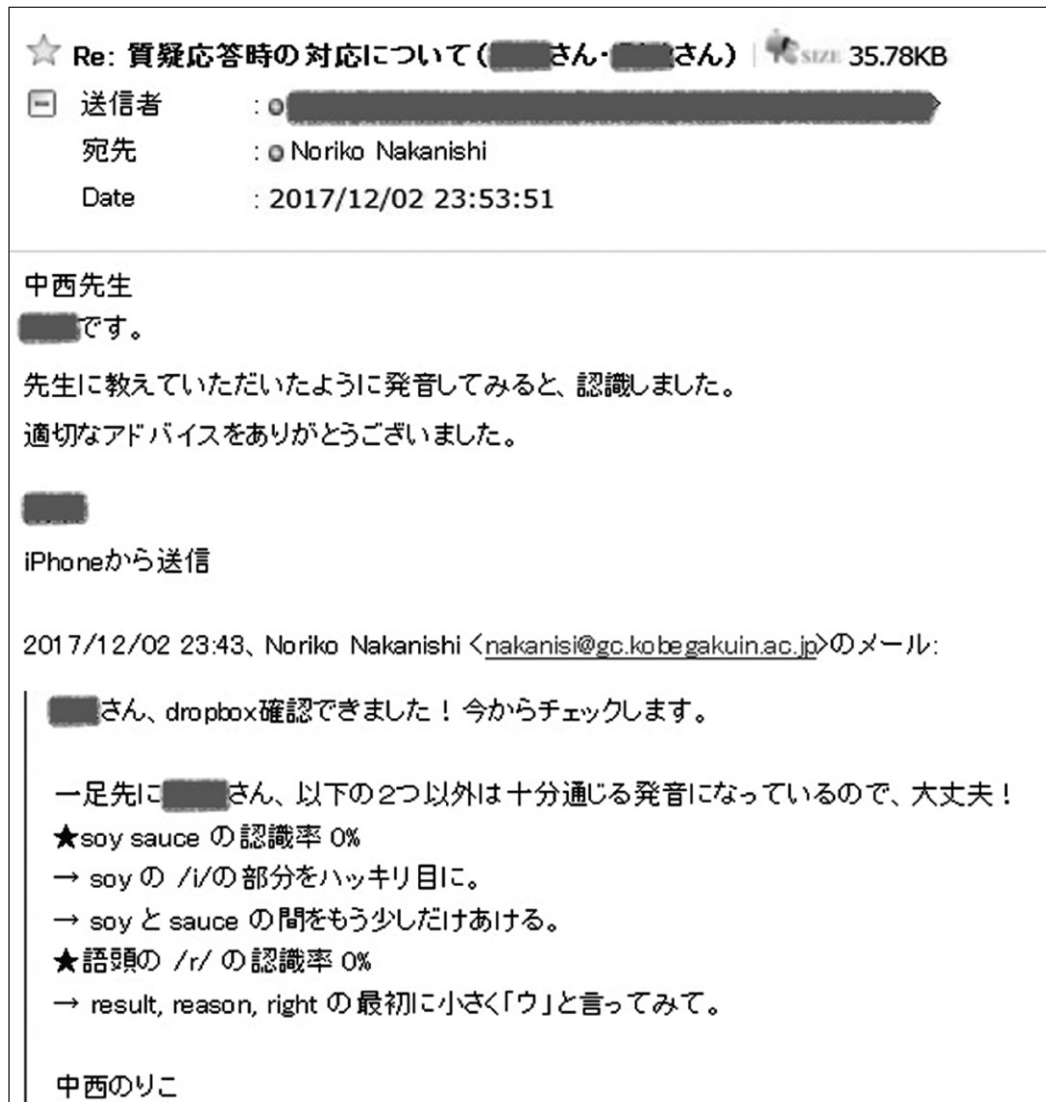


図3. 音声認識アプリ及び音素カウンターを利用した音声指導のやり取り（抜粋）

この事例は、教員が学習者と対面で直接指導をする時間的余裕がない場面で、学習者の発音についてのアドバイスをするために音素カウンターによる分析を試みたものである。音声認識アプリを用いて入力した英文に含まれる音素比率の数値を手がかりに、個別音の発音の癖を見つけることを試みたパイロット事例といえる。

ただし、この事例において、わずか5行のアドバイスを元にして学習者が自分で発音修正をすることができたのは、授業を通して発音記号を読むことができるようになっていたことや、音素カウンターの出力画面を見て文強勢の置き方や音声変化が起こる場所を把握するコツをすでにつかんでいたことが幸いしたと考えられる。音素カウンター ver.04 へのバージョンアップでは、英語音声に関する知識が乏しい状態の学習者にも感覚的に理解できるようなコンテンツやインタフェースを心がけたい。

4. 音素カウンター ver.04 構築に向けて

本研究における2つの事例では、音声認識アプリと音素カウンターを併用することにより、学習者が自分の英語発音の問題点を見つけ、自律的に発音矯正を行うことができるようになる可能性が見出された。指導者が学習者と向き合って発音指導を行うことができない場面でも、音声認識結果や発音記号で音声情報が視覚化されれば、学習者が自分自身の発音を客観的に捉えることができるという点で有効な手法であると考えられる。また、ここで試みた分析は、単に2つの英文テキストに含まれる音素数を比較するだけのことであったことから、システム中に分析のパターンを組み込んでしまえば自動化が可能である。

一方、一般的な日本語母語話者の英語発音の問題点は、分析画面で音素比率を見るだけで発見することができるものばかりではない。スムーズにリズムよく話すためには、本研究で取り上げたような比較的単純な音素置き換えだけではなく、子音クラスターのような特定の音の連続、連結・脱落などの音声変化や、文強勢の置き方に問題がないか確認する必要がある。音素カウンター ver.04 へのバージョンアップでは、音声変化や強勢の置き方の面で問題がある発話がどのように音声認識されるかを分析し、発音矯正のポイントを提案するシステム構築につなげたい。どのような音声の特徴がどのような音素認識率の偏りとなって表れやすいかを特定するために、今後、このような事例を十分に蓄積していくことが望まれる。

また、本研究ではiOSに標準装備されているアプリによる音声認識結果を学習者の英語発音の伝わりやすさを測る目安とした。しかし、これは人間の聴覚ではなく機械による自動認識アルゴリズムを用いて測定したものである。単に入力された音声のみではなく語連鎖などの文脈を考慮し認識する仕組みにはなっているものの、実際の人間同士のコミュニケーションにおける発話意図の伝わりやすさを100%再現しているとは言い難い。また、音声認識の技術は日進月歩で向上しているため、例えば同じ学習者が同じスクリプトを音声認識させた結果が数年後に向上していたとしても、この学習者の英語発音が向上したのか、アプリ側の音声認識の技術が向上したのか判断することはできない。英語発音だけでなく言語能力自体、本来は人と人とのコミュニケーション活動によって培われていくべきものであるということを念頭に置いた上で、現状の一斉指導の教室環境では不可能に近い指導を補うツールの1つとして、今後も音素カウンターの機能充実に取り組みたい。

謝辞

本研究は、以下の研究助成を受けたものです。

- 科研費基盤研究（C）26330321 言語音もたらすイメージ：その普遍性と文化的特性
2014年度～2016年度
- 神戸学院大学教育改革助成金英文テキスト音素比率算出システム「音素カウンター」の構築
2015年度
- 科研費基盤研究（C）17K02914 英語の音韻体系を意識させる理論的・自律的発音学習の効果実証研究
2017年度～2021年度

参考文献

- [1] Celce-Murcia, M., Brinton, D. M., Goodwin, J. M., & Griner, B. (2010). *Teaching pronunciation paperback with audio CDs (2): A course book and reference guide*. Cambridge: Cambridge University Press.
- [2] Chen, H. H. J., & Chen, L. W. C. (2018). Automated Speech Assessment. *The TESOL Encyclopedia of English Language Teaching*, 1-6.
- [3] Derwing, T. M., & Munro, M. J. (2005). Second language accent and pronunciation teaching: A research-based approach. *TESOL Quarterly*, 379-397. doi. 10.2307/3588486
- [4] Ehsani, F., & Knodt, E. (1998). Speech technology in computer-aided language learning: Strengths and limitations of a new CALL paradigm. *Language Learning & Technology (1)*, 54-73.
- [5] International Phonetic Association. (1999). *Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet*. Cambridge University Press.
- [6] Jenkins, J. (2000). *The Phonology of English as an International Language* (1st edition.). Oxford, England: Oxford University Press.
- [7] Kachru, B. B. (1982). *The other tongue. English across cultures*. Urbana, Illinois: University of Illinois Press.
- [8] L-Interface. (2017). Academic Express3. <http://www.supereigo.com/academic3/>
- [9] Liaw, M. L. (2018). Speech Recognition. *The TESOL Encyclopedia of English Language Teaching*, 1-7.
- [10] 峯松信明・鎌田圭・朝川智・鈴木雅之・牧野武彦・西村多寿子 & 広瀬啓吉. (2011). 音声の構造的表象に基づく学習者分類の検証と発音矯正度推定の高精度化. 情報処理学会論文誌, 52 (12), 3671-3681.
- [11] Munro, M. J., & Derwing, T. M. (1995). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language learning*, 45 (1), 73-97. doi. 10.1111/j.1467-1770.1995.tb00963.x
- [12] 中西のりこ. (2015). 「発音記号を使用した英語発音指導の取組み」. 外国語教育メディア学会第55回全国研究大会発表要項集, 22-22.
- [13] 中西のりこ. (2016a). 「英文テキスト音素比率算出システム「音素カウンター」の開発：一般英語教材を発音練習用教材として利用する方法」. 大学英語教育学会第55回国際大会発表要項集, 18-18.
- [14] 中西のりこ. (2016b). 「大規模クラスにおける英語音声指導：音声認識アプリ Dragon Dictation を用いた自己分析の試み」. 外国語教育メディア学会関西支部2016年度春季研究大会発表要項集.
- [15] Nakanishi, N. (2017). Phoneme Counter (Ver3.1) [computer Software]. Kobe, Japan: Kobe Gakuin University. <http://noriko-nakanishi.com/phoneme/counters>
- [16] 中西のりこ. (2017). 「「音素カウンター」を用いた授業実践」. JASELE 第43回島根大会. 2017年08月20日.
- [17] 中西のりこ. (2018a). 「合成音声 (Text-to-Speech) ・音声認識 (Speech-to-Text) ソフトの英語教育への応用：「音素カウンター」×音声認識」. ATEM 西日本支部大会. 2018年03月03日.
- [18] 中西のりこ. (2018b). 「日本語母語話者の英語発音の特徴：音声認識アプリを用いた分析」. KELES 第23回研究大会. 2018年06月09日.
- [19] Nakanishi, N. (2018). Development of Phoneme Counter: Open Software for Calculating Phoneme Counts in English Texts. *Language Education and Technology*, 55. in press.
- [20] Nikkei. (2017). Versant speaking test. http://www.versant.jp/versant_test.html
- [21] Ruan, S., Wobbrock, J. O., Liou, K., Ng, A., & Landay, J. (2016). Speech is 3x faster than typing for english and mandarin text entry on mobile devices. *arXiv preprint arXiv:1608.07323*.