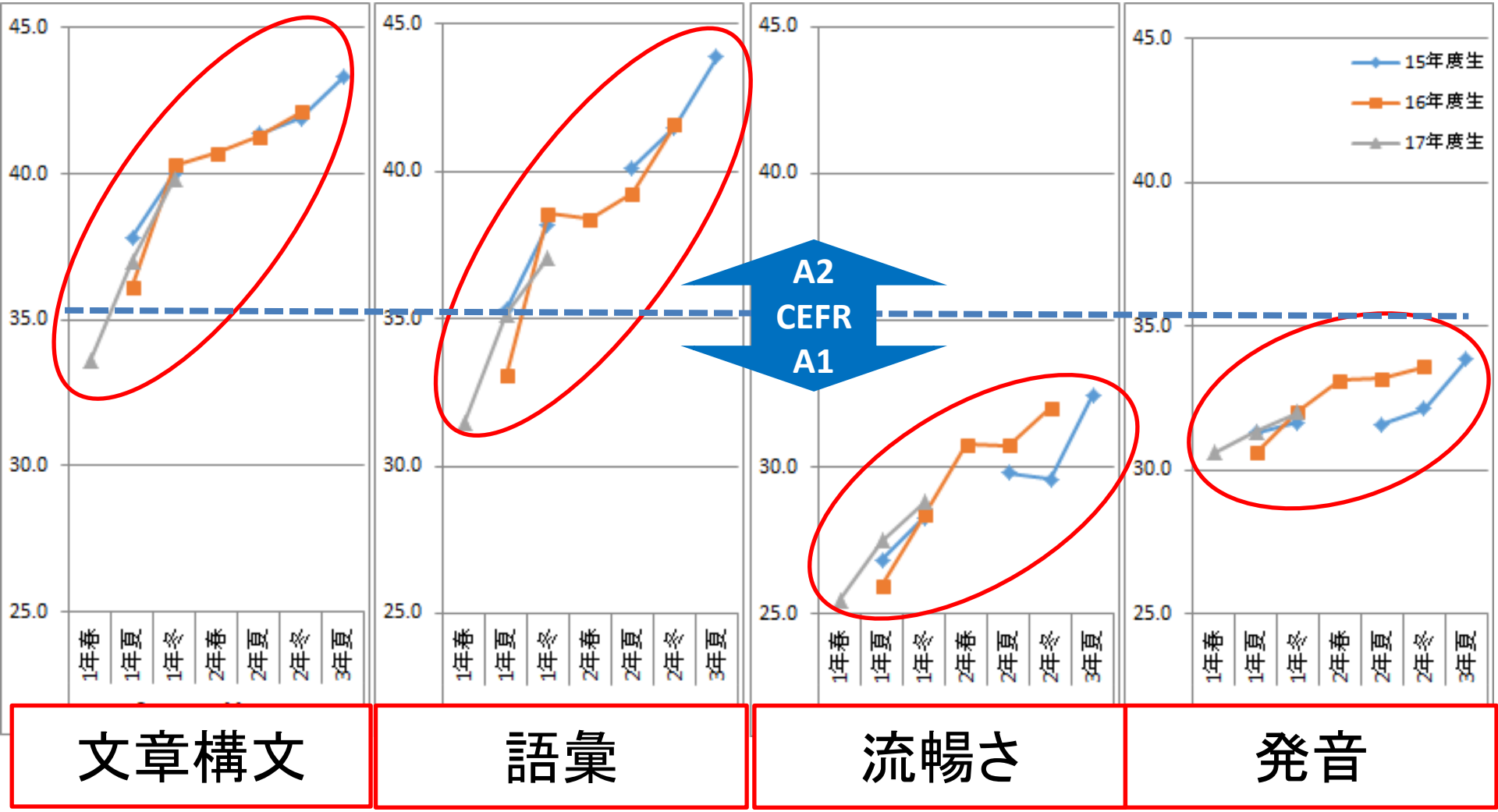
日本語母語話者の英語発音の特徴
音声認識アプリを用いた分析

神戸学院大学
中西のりこ

2018.06.09
KELES 第23回研究大会
@関西国際大学

1. 研究の背景

1-1. スピーキング要素のアンバランス



Versant Speaking test 2015-17年結果

1-2. 教室での発音指導の問題点

- 指導すべき内容が多くあるため発音は後回しになる。
- 個々の学習者の発話への対応に教員の手が回らない。
- 学習者ごとに発話内容が異なると、モデル音声を用意できない。
- 日本語を母語とする学習者間のインタラクションでは、カタカナ発音の方が通じやすい。
- 英語の音韻に関する知識が乏しいため、学習者間ではフィードバックの内容があいまい（そもそも発音記号さえ知らない。）

中西 (2016b).

1-3. 発音指導の目標

選択肢		英語専攻（n=46）	非英語専攻（n=283）
① NS 型	ネイティブスピーカーと同じような発音	34 (73.9%)	103 (36.4%)
② EIL 型	伝わるのであれば、発音にはこだわらない	9 (19.6%)	170 (60.1%)
③ Japalish 型	日本語の特徴を残しておきたい	3 (6.5%)	10 (3.5%)

（カッコ内）英語／非英語専攻グループ内での比率。

中西 (2008).

⇒「スマホで円滑なコミュニケーションがとれる程度の英語発音を目指そう」

1-4. 先行研究

- **Derwing & Munro (2005).**
  音声認識ソフトが人間の聞き手と同様の反応を示すとは言い難い。

- **峯松・鎌田・浅川・鈴木・牧野・西村・広瀬(2011).**
  母語干渉の影響を受けた外国語発音の自動分析
  発音矯正度推定

- **Ruan, Wobbrock, Liou, Ng, & Landay (2016).**
  音声認識による文字入力 vs. キーボード入力
  (n = 32 × 50 trials × 2 ways)
  →スピード: 英語 3.0倍、中国語 2.8倍
  →正確さ: 英語 20.4%軽減、中国語 63.4% 軽減

⇒本研究の目的:2018年1月時点で、日本語を母語とする英語学習者の英語発音がどのように認識されるか、記録する。

1-5. Research Questions

音声認識アプリで

1) 英語発音の能力を予測できる？

→Versant Speaking Test 結果と相関があるか？

2) 文強勢の正しさを予測できる？

→認識される音の数と関係があるか？

3) 音声変化（連結）の正しさを予測できる？

→認識される語頭子音の数と関係があるか？

4) 音声変化（脱落）の正しさを予測できる？

→認識されるあいまい母音の数と関係があるか？

中西 (2018).

2. 方法

2-1. 対象

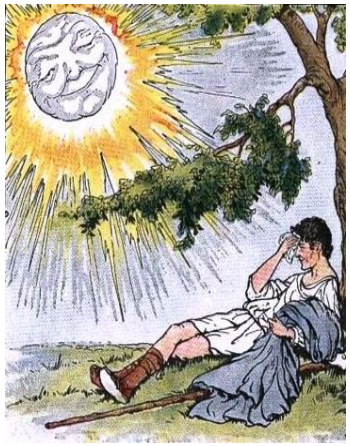
- 関西圏の大学で国際系の学部に所属する1年次生 (n = 94)
- TOEIC score (2017. 12. 12)

	Total	Listening	Reading
M	382.7	234.7	148.0
SD	102.9	64.1	51.0
max	675.0	360.0	315.0
min	195.0	85.0	70.0

- Versant speaking test score (2018. 01.19)

	Total	Sentence	Vocabulary	Fluency	Pronunciation
M	35.0	40.6	37.9	29.2	32.2
SD	4.1	4.1	6.6	6.3	3.6
max	46.0	54.0	56.0	47.0	45.0
min	26.0	31.0	22.0	20.0	26.0

2-2. 素材



The North Wind and the Sun

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 first succeeded in making the traveler
take his cloak off should be considered stronger than the other.
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.
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.

(113 words, 144 vowels, 232 consonants, 336 segmentals)

International Phonetic Association, 1999

2-3. 音声認識アプリ・音素カウンター

- iPad Air 2 Wi-Fi 64GB 標準搭載 音声認識アプリ(2018年1月時点)
- 音素カウンター　Ver4.0　（最終更新日：2018年3月2日）

データベースに含まれる語： 79,400 語（動詞・名詞の変化形を含む）

音素カウンターの使い方

分析結果 他の英文を分析

発音記号　データベースに存在しない語　今後の予定

分析１　分析２

※発音記号の下のアイコンをクリックすると動画で発音を確認できます。

※音素数をクリックすると分析結果の該当の音素がハイライトされます。

連_結　（脱落）　同_化

	iː	i	ɪ	e	æ	uː	u	ʊ	ɔː	ɔ	ɑː	ɑ	ʌ	əː	ə	eɪ	aɪ	ɔɪ	aʊ	oʊ	ər	əːr	ɪər	eər	ʊər	ɑːr	ɔːr
音素数	3	8	17	5	8	5	0	1	6	0	0	0	6	0	48	5	2	0	2	7	11	1	0	0	0	1	8
語頭			3						2				2		14				1		1						
語中	3	2	14	5	8	1		1	4				4		15	4	2		1	6	1	1				1	6
語末		6				4									19	1				1	9						2
強勢	3		6	5	8	5		1	6				6			4	2		2	7		1				1	8

	p	t	k	b	d	g	tʃ	dʒ	f	θ	s	ʃ	h	v	ð	z	ʒ	m	n	ŋ	l	w	r	j
音素数	4	23	17	5	24	2	1	1	5	4	15	2	8	6	25	8	0	9	26	6	18	13	9	1
語頭		8	9	4	2	1			2		8	2	8		24			3	4		1	13	1	

2-4. 音声認識の手順

入力画面

音素カウンター　Ver2.0　（最終更

データベースに含まれる語： 79,393 語（動詞・名詞の変化形を含む）

分析したい英文を入力してください　入力内容をクリア

① ここに英文を入力。
直接入力・コピーペースト可能。
iPad, iPhone では音声認識入力可能。

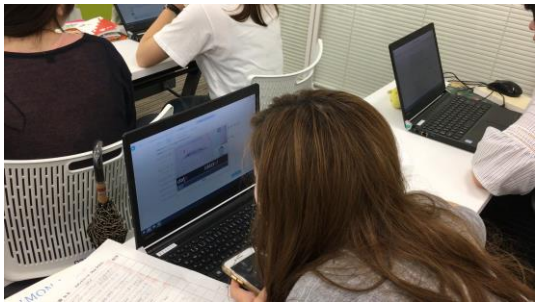
② 「同意する」にチェック。

入力された英文についての　情報を本システム作成者に提供することに同意します。

◉ 同意しない
○ 同意する

③ 「分析」をクリック。

分析



中西 (2017)

3. 結果の概要

3-1. 全体像

	①分節音認識率	②スペース音挿入率	③全体認識率
M	57.2%	15.8%	64.0%
SD	15.2%	6.6%	12.2%
max	97.0%	35.1%	97.8%
min	27.7%	0.0%	42.2%

① = (母音数+子音数+子音クラスター数) / 336

② = 単語間のスペースに音が認識されたケース数 /114

③ = モデル音と同じ音が認識された(されなかった)割合

3-2. 分節音別 認識率（母音）

	n	*M*	*sd*	%
íː	3	63.7	5.2	67.7
i̇	7	64.1	12.0	68.2
ɨ́	6	57.3	6.0	61.0
ɪ	11	68.0	12.0	72.3
é	5	51.6	10.9	54.9
ǽ	8	32.9	14.9	35.0
úː	5	44.4	11.6	47.2
ʊ́	1	31.0	---	33.0
ɔ́ː	6	58.8	13.6	62.6
ʌ́	6	41.0	19.2	43.6
ə	48	53.3	13.3	56.7

	n	*M*	*sd*	%
i̇ə	1	50.0	---	53.2
éɪ	4	71.3	13.6	75.8
eɪ	1	40.0	---	42.6
áɪ	2	35.0	17.0	37.2
áʊ	2	52.0	1.0	55.3
óʊ	7	33.1	20.8	35.3
ər	10	42.5	25.7	45.2
ə́ːr	1	17.0	---	18.1
ɑ́ːr	1	68.0	---	72.3
ɔ́ːr	8	34.8	16.8	37.0

3-3. 分節音別 認識率（子音）

	n	*M*	*sd*	%
p	1	71.0	---	75.5
t	11	66.7	9.8	71.0
k	11	72.7	10.3	77.4
b	2	79.5	0.5	84.6
d	12	58.6	10.9	62.3
g	1	77.0	---	81.9
ʧ	1	64.0	---	68.1
f	5	42.8	10.4	45.5
θ	4	28.5	2.1	30.3
s	8	69.3	9.4	73.7
ʃ	2	70.5	8.5	75.0

	n	*M*	*sd*	%
h	8	69.8	13.3	74.2
v	6	37.0	14.9	39.4
ð	25	50.1	8.6	53.3
z	8	60.3	13.3	64.1
m	8	64.6	21.2	68.8
n	16	62.0	14.3	66.0
ŋ	6	73.0	10.4	77.7
l	7	31.1	15.0	33.1
w	13	66.7	6.1	70.9
r	2	47.0	10.0	50.0

3-4. 分節音別 認識率（子音クラスター）

語頭 語中	*n*	*M*	*sd*	%
kl	5	44.2	7.6	47.0
tr	4	31.5	3.0	33.5
str	3	68.7	3.9	73.0
bl	3	18.0	8.6	19.1
gr	1	47.0	---	50.0
ks	1	78.0	---	83.0
pj	1	45.0	---	47.9
sl	1	29.0	---	30.9

語末	*n*	*M*	*sd*	%
ld	1	14.0	---	14.9
mpt	1	58.0	---	61.7
nd	10	49.1	6.7	52.2
pt	1	23.0	---	24.5
st	2	54.5	3.5	58.0
dʒd	1	7.0	---	7.4

4. 検証

4-1. 英語発音の能力を予測できる？

→③全体認識率とテスト結果の相関を調べる。

- Versant Speaking Test (2018.01.19)

総合	0.39	文脈においてネイティブ同様に子音や母音を発音し、強勢を置くことのできる能力
文章構文	0.35	発言を把握し、それを逐語的に発話する能力
語彙	0.38	構文中の日常的かつ一般的な単語を理解し、自在に表現する能力
流暢さ	0.25	文章の組み立て、読み、反復の際のリズム、区切り及びタイミングの取り方の能力
発音	0.30	文脈においてネイティブ同様に子音や母音を発音し、強勢を置くことのできる能力

- TOEIC IP (2017.12.12)

total	0.26
Listening	0.31
Reading	0.13

4-2. 文強勢の正しさを予測できる？

→認識される音の数と関係があるか？

→②スペース部分への音挿入率が高い音源の特徴を調べる。

- 音挿入率 00.0%　　(強勢母音 13か所)

ðə nóːrθ wínd ənd ðə sʌ́n ər dɪspjúːtɪŋ wíʧ wəz stróːŋər wén aɪ trǽvələr kéɪm əlóːŋ rǽpt ɪn wóːrm klóʊk

- 音挿入率 35.1%　　(強勢母音 21か所)

ðə nóːrθ wíndoʊ ǽntə ðə sóːŋ wʌ́t iz ðə θíŋ wíʧ wʌ́ts ðǽts θróʊɪŋ óːf wén aɪ téːrn óːf ənd aɪ kéɪm əlóːŋ lʊ́k ʌ́p tə ðə fíːld láːmə krǽk

- 音挿入率 33.3%　　(強勢母音 15か所)

ðə nóʊz wíndoʊ ənd dǽts ə sóːŋ váɪə féɪsbuk bíːɪŋ réd tə əs ə stróːŋər wéər ət ðə rǽbət aɪl kʌ́m əlóːŋ mǽm klóʊk

4-3. 音声変化（連結）の正しさを予測できる？

→認識される語頭子音の数と関係があるか？
→②スペース末の子音挿入率が高い音源の特徴を調べる。

- スペース末子音挿入率 01.0%
  wén ə trǽvələr kéɪm əlɔ́ːŋ rǽpt ɪn ə wɔ́ːrm klóʊk

- スペース末子音挿入率 17.2%
  wén nát dráɪv ɪt áʊt kéɪm ðə rɔ́ːŋ rǽpɪd ɪn ðǽt wʌ́n gə́ːrl

- スペース末子音挿入率 12.2%
  wén jʊ tráɪ ðéər kéɪm ə rɔ́ːŋ wáɪ bət ɪt ɪn jər ɑ́ːrtwəːrk lə́ːrn mɔ́ːr krák

4-4. 音声変化（脱落）の正しさを予測できる？

→認識されるあいまい母音の数と関係があるか？

→ /ə/の比率が高い音源の特徴を調べる。

/ ənd ə(t) lǽs(t) ðə nɔ́ːrθ wín(d) géɪv ʌ́(p) ði ətémp(t) /

- /ə/の比率 10.4%

  / ənd ət ðə lǽs(t) ðə nɔ́ːrθ wíŋ géɪv ʌ́ (p) ði ətémp (t) /

- /ə/の比率 10.1%

  / ənd ət ðǽt ǽs( ) náɪs ríŋ tə gét ɔ́ːf əv ətémp (t) /

- /ə/の比率 05.6%

  / ɑn ðə( ) lǽs(t) tə nóʊ hɪz wíndoʊ géɪv ʌ́(p) ðeɪ ətémp(t) /

- /ə/の比率 05.4%

  / aɪ lʌ́v jʊ ənd lǽs(t) nóʊz wáɪ géɪm ət ə téɪp(t) /

5. 考察

5-1. 検証結果からの示唆

音声認識アプリで

1) 英語発音の能力を予測できる？

→Versant Speaking Test 結果と弱い相関がある。

2) 文強勢の正しさを予測できる？

→強勢が置かれる母音の数に焦点を当てる？

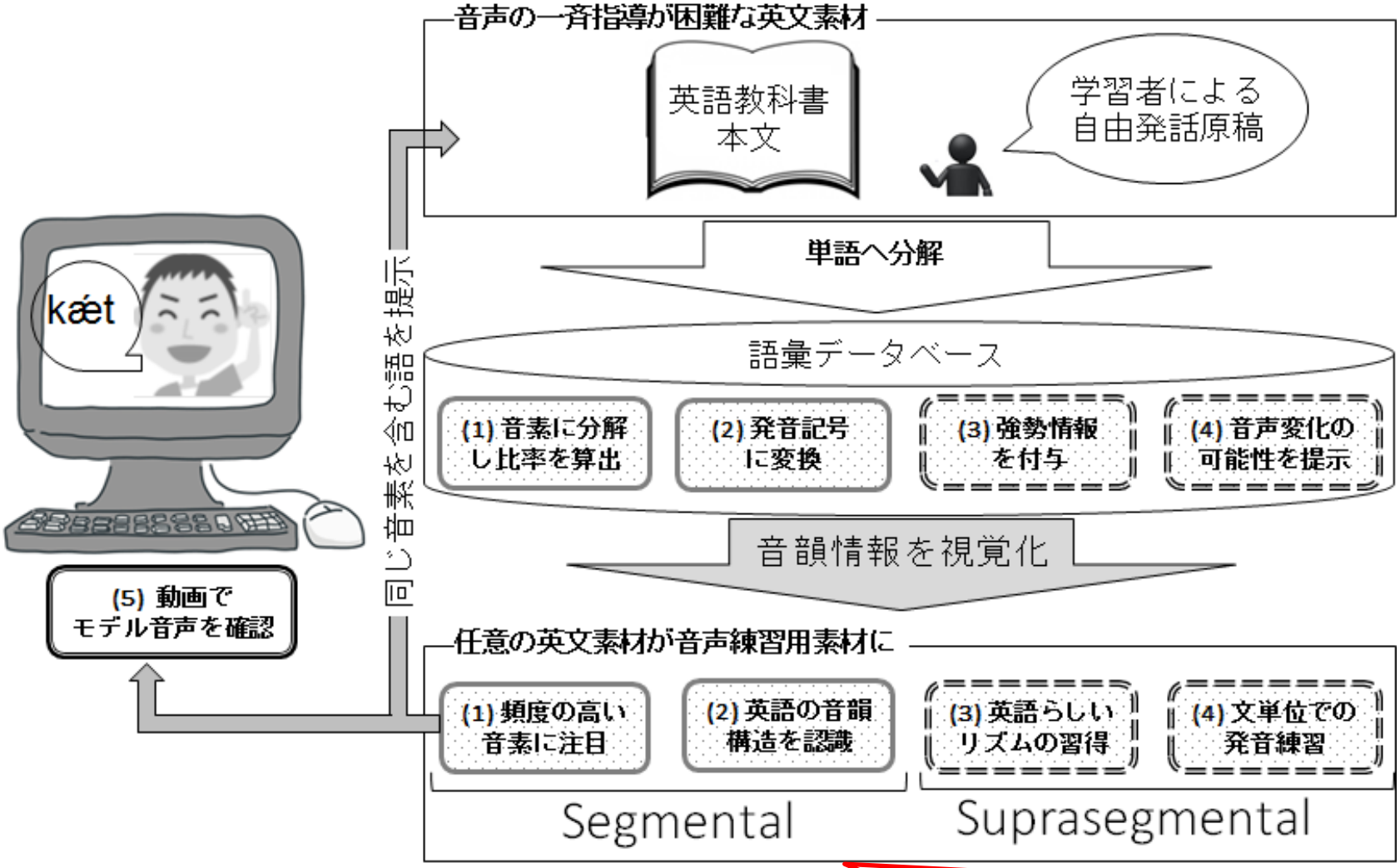
3) 音声変化（連結）の正しさを予測できる？

→認識される語頭子音の数とある程度の関係がある。

4) 音声変化（脱落）の正しさを予測できる？

→/ə/の比率をカウントするだけでは予測できない。

5-2.「音素カウンター」Suggestion機能について



eg.)「語末の母音挿入が多いようです。強勢が置かれる母音に注意して、リズムよく発音する練習をしましょう。」

5-3.音声認識アプリの進歩について

音声認識の精度は日進月歩で改良されている。
→カタカナ英語の認識率も上がってきている。
→ある時点の精度で認識した結果をSuggestionに反映させる必要がある。

6. 参考文献

Anthony, L. (2014). AntConc (Version 3.4.4w) [Computer Software]. Tokyo, Japan: Waseda University. Available from http://www.antlab.sci.waseda.ac.jp/

Derwing, T. M., & Munro, M. J. (2005). Second language accent and pronunciation teaching A research-based approach. Tesol Quarterly, 39(3), 379-397.

International Phonetic Association. (1999). Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet. Cambridge University Press.

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.

中西のりこ. (2008).「英語を専門としない学生の発音学習に対する意識-World Englishes時代に求められる英語発音」.神戸学院大学経営学論集第5巻1号, pp.1-15.

中西のりこ・阪上潤・山本祐太. (2015).「発音記号学習に対する大学生の意識の変化：KJ 法による分析」. 外国語教育メディア学会 関西支部2015年度春季研究大会発表要項集.

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. p. 22.

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. p. 18.

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.

中西のりこ. (2017).「音素カウンター (ver. 4.0)」 [computer Software]. 神戸学院大学.

中西のりこ. (2018).「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.

Nuance Communications. (2013). Dragon Dictation (Version 2.0.28) [iOS application]. Available from http://www.nuance.com.

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.

本研究はJSPS科研費

17K02914

の助成を受けたものです。