

令和元年6月25日現在

機関番号：32689

研究種目：基盤研究(B)（一般）

研究期間：2015～2018

課題番号：15H02729

研究課題名（和文）バイリンガル音声合成を用いた英語における日本語アクセントの強度と発話理解度の検証

研究課題名（英文）Comparing the comprehensiveness of English speech with the degree of Japanese accent using synthesized Japanese-English bilingual speech

研究代表者

近藤 真理子（Kondo, Mariko）

早稲田大学・国際大学院・教授

研究者番号：00329054

交付決定額（研究期間全体）：（直接経費） 8,500,000円

研究成果の概要（和文）：日本語話者の英語の発音を音声コミュニケーションの観点から考察し、日本語訛の英語の特徴と、英語母語話者と非母語話者が日本語訛の英語を理解できるか、どう評価するかを検証した。また、同一話者の音声に基づき日英バイリンガル音声合成し、外国語訛の度合いを変化させた日本語訛の英語と英語訛の日本語音声を使い、分節音の正確さと外国語訛判定の知覚実験を行った。日本語訛の英語の判定においても、バイリンガル合成音を使った知覚実験においても、従来から指摘されている特定の分節音の日本語訛度が高くなると、外国語訛の発音と判定された。結果を受けて、特定の分節音の習得に特化した発音教育法の開発が求められる。

研究成果の学術的意義や社会的意義

本研究は、これから益々一般的となる非母語話者間の音声コミュニケーションにおいて、すでに数多くの研究が行われている学習者としての発話理解・意思の疎通だけではなく、“情報を伝達する側”という視点で、英語母語話者・非母語話者を問わず誰にでも通じる日本語訛の英語とは何か、どのような発音であれば、母語を問わず誰にとっても理解されるのかを客観的に示し、英語音声習得教育の指針を示すことを目指した研究である。外国語訛があっても必ずしも理解度が低くなるわけではなく、母語話者のように話せるようになる教育よりも、必要な情報を伝えられる音声コミュニケーションを想定した外国語音声教育が大切なことが分かった。

研究成果の概要（英文）：Japanese accented English pronunciation was studied in terms of speech communication by native and non-native English speakers. The focus of the study is to examine if Japanese accented English is easily understood by other language speakers and whether the necessary information is conveyed to listeners. First, acoustic details of Japanese accented English were examined using a large-scale speech corpus. Based on these findings, Japanese accented English was synthesized by changing the degree of Japanese accentedness of some English phonemes using the Cyborg speech synthesis system, developed based on deep-learning-based speech synthesis. Then perception tests were performed to investigate segmental accuracy and Japanese accentedness. We found that assessors' first language had little effect on the comprehension of the Japanese accented English. The results suggest that new pronunciation teaching systems need to be developed that focus on improving segmental accuracy.

研究分野：音声学、第二言語音声習得、音声コーパス

キーワード：第二言語音声習得 バイリンガル音声 英語音声習得 第二言語音声評価 音声知覚

様式 C-19、F-19-1、Z-19、CK-19（共通）

1. 研究開始当初の背景

非母語話者間での言語コミュニケーションがどの言語でも一般的になっている昨今、母語話者と同じように話すのではなく、いかに様々な母語(L1)の聞き手に伝えたいことを伝えるか、伝わるかが益々重要となってくる。本研究は、これから益々一般的となる英語非母語話者間の音声コミュニケーションにおいて、すでに数多くの研究が行われている学習者としての発話理解・意思の疎通だけではなく、“情報の伝達”という言語の重要な役割を想定し、母語話者・非母語話者を問わず誰にでも通じる外国語(L2)としての英語の音声習得の実態を把握し、コミュニケーションのツールとしての英語音声とは何かを探るものである。例えば、現実に外国語が使われている、観光案内所や交通機関、駅・空港等の公共の場、医療現場、災害時などの緊急アナウンス等、非母語話者（日本であれば日本語母語話者）が外国語（例えば英語）を使って非母語話者（英語圏以外の出身者）に情報を伝えなければならない状況は少なくない。この現状を鑑みると、母語話者のように話せるようになる教育よりも、非母語話者同士が必要な情報を得られる音声コミュニケーションを想定した外国語音声教育が大切なことが分かる。現実に外国語が使われる状況を念頭に、限られた教育環境、限られた時間で、効率的に“誰にでも理解しやすい発音”の習得・教育のために、学習者の母語別にコミュニケーションの妨げとなる重要なエラーを抽出し、そのエラーの修正に特化した、効率的な音声指導法及び教材の開発が必要となる。

このような、外国語特に英語を取り巻く環境の変化に応じた、日本語話者を対象とした英語音声習得の現状と、それに基づく英語音声教育のあるべき方向を見極める必要があるとの考えから、この研究を行った。

2. 研究の目的

外国語学習者は、学習歴に比例して、ある程度知覚・生成とも母語話者のレベルに近づくが、母語話者とは同じレベルにはならない。発音が悪くても、最低限のコミュニケーションは成り立つが、一定の基準を下回ると理解不能と判定される。L2 音声習得において、母語話者による流暢さの知覚は、音素レベルよりも韻律レベルの逸脱のほうがより強く外国語訛の印象を与えるが、分節音レベルの発音の正確さは、聞き手の発話内容の理解度に深く関係しており、コミュニケーションにおいて、伝えたい内容を正確に聞き手に伝えるという点においては、重要な要素となる。

本研究は、日本語話者の英語の発音を音声コミュニケーションの観点から検証し、意思の疎通に必要な英語の発音とはどのようなものか、日本語訛の英語は L1 に関わらず、一般的にどの程度通じるのかを、科学的・客観的に示すことを目的とする。英語母語話者の発音との比較ではなく、非英語母語話者にも”通じる”英語の発音とはなにかを検証し、英語母語話者・非英語母語話者とのコミュニケーションに必要な英語の発音の到達点を見出すことを目的とする。また日本語母語話者以外の話者による、日本語訛の英語発話の理解度を検証した。母語を問わず、L1, L2 両方の一般の言語使用者にとって聞きやすく分かりやすい、非母語話者による発話とはどのようなものか。理解される発音と理解されない発音の分岐点は何かを解明する。

L2 を使った非母語話者間の音声コミュニケーションにおいて、非母語話者同士が効果的に情報伝達をするために、最低限達成しなければいけない外国語音声の習得目標は何か、日本語母語話者の英語習得を例として、L2 分節音の習得と、その問題点、また外国語訛の英語が母語話者と非母語話者にとって、どの程度分かりやすいか、分かりにくいかを考察する。

3. 研究の方法

本研究は、大規模音声発話コーパスを使った日本語訛の英語の音声特性の解析と、日英語のバ

イリンガル合成音を使った知覚実験を柱として行う。具体的な方法としては、

(1)は日英語バイリンガル音声合成のための(a)日本語話者の英語音声の分析と、(b)日英語バイリンガル話者の日本語および英語の音声発話収録を行った。先行研究の調査と、大規模 L2 英語発話コーパスを使い、日本語話者の英語発話を英語話者の発話データと調査比較。間違いの頻度の高いもの、コミュニケーションに差し障るものを、英語評定レベル別に検証し、中・上級レベルでも繰り返される間違いを特定した。

(2)では、(1)の結果を受け、(a)語強勢と母音の弱化、(b)母音の綴りのアルファベットと発音、(c)母音の挿入について考察した。ここでは他項目との関連から、(a)と(b)の詳細を述べる。

(3)では日本語母語話者の英語音声を、(a)英語母語話者、(b)日本語母語話者、(c)その他の言語話者を対象に、(i)分節音の正確さ、(ii)韻律の正しさ、(iii)流暢さ、(iv)母語訛の少なさ、について評価してもらい、L1による日本語訛の英語の発音の判定に違いがあるかどうか調査した。

(4)では、日本語とアメリカ英語のバイリンガル話者が発話した、同一話者の発話による純粋な日本語と純粋な英語の音声を基に作成した、英語訛の非常に強い日本語音声から純粋な日本語音声と、日本語訛の非常に強い英語音声から純粋な英語音声の、音素の逸脱度を段階的に変化させた合成発話音を使い、日英母語話者を対象とした母語アクセントの強度を測る知覚実験を行い、外国語訛度と発話理解度を検証した。合成音をつかった知覚実験は、ある程度の被験者数が必要となるので、クラウドソースを使って行った。

4. 研究成果

(1) 日本語話者の英語発話の特徴

大規模音声コーパスを用いて日本語話者の英語発話の分析を行った。その結果、従来から指摘されている/l/-/r/, /b/-/v/, /s/-/θ/などの子音の発音の間違いは必ずしも双方向に間違いが起きているのではなく、/l/の代わりに/r/, /v/の代わりに/b/と一定の方向性があることが分かった。また、日本語訛の研究は子音の間違いについての研究が多いが、間違いの2/3は母音の音質と挿入であり、今後の日本語訛の英語の分節音の研究の方向を見直す必要があることが分かった。

(2) 日本語話者の英語音声発話分析

(a) 語強勢と母音の弱化

日本語話者は初級を含めて英語母語話者グループと同様、強勢音節と弱音節で音圧(3グループとも $p < .001$)、F0(3グループとも $p < .001$)、母音長(duration, 3グループとも $p < .001$)で有意な差があったが、音質に関しては、英語母語話者グループは強勢の有無で母音の音質に有意な差が見られたが ($p < .001$)、日本語話者は上級者グループを含む3グループとも有意な差が見られなかつた($ps > .9$, n.s.) (図1)。つまり、日本語母語話者は英語の強勢の有無を、音圧、F0、母音長で制御するが、母音の音質の変化は強勢リズムに関連した音響特性にはなっていない。

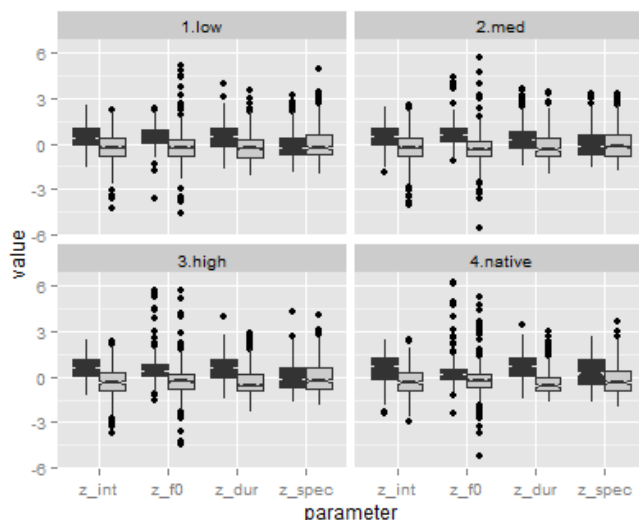


図1. 初級(low)、中級(med)、上級(high)の3つの異なる英語習熟度の日本人英語学習者と英語母語話者の英語の強勢音節(黒)と弱音節(灰色)の母音の音響特性

(b) 母音の綴りのアルファベットと発音

綴りのアルファベットの文字の発音への影響は、英語母語話者でも多少はあるが、英語母語話者の母音はF1またはF2のどちらかが他の母音の対応するフォルマントと有意に異なっているが、F1とF2の両方が有意に異なることはなかった。しかし、日本人の上級者は/a/と/e/に関してはF1とF2の両方の値が有意に異なっていたが (F1: $p < .001$, F2: $p < .003$)、/a/と/o/はF2のみ有意差があり (F2: $p < .001$)、初級グループのF1は/a/:e/, /a/:o/, /e/:o/すべて $p < .001$ で、F2は/a/:o/と/e/:o/間で有意差があった ($p < .001$) (図2)。

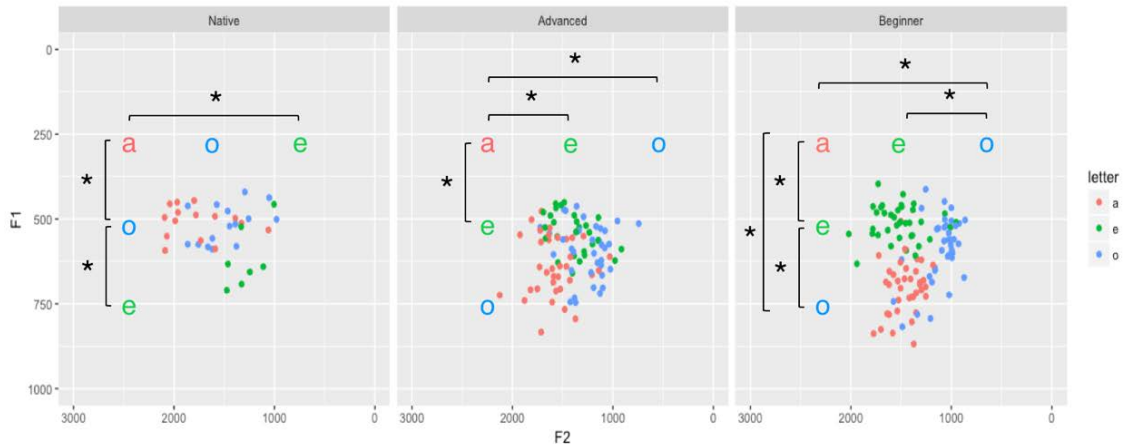


図2. 英語話者 (左) と日本人上級 (中) と初級 (右) の英語母音のF1(Y軸)とF2(X軸)値

つまり、上級グループのフォルマント分布は、英語母語話者のフォルマント分布に近く、/a/:e/のように綴りのフォルマント値への影響は若干みられるものの、明らかに異なる母音となっているとは言い切れない。一方、初級グループのフォルマント値は/a/:e/に関してはF1の値のみ有意な差あったが、/e/:o/も /a/:o/もF1とF2の両方の値が有意に異なっており、異なる音として産出しようとしているだろうことが推測される。

(c) 音節構造に起因する母音の挿入

日本語の音節構造との関係では、英語習熟度が上がるにつれて、挿入母音の数は少なくなったが、英語レベルがかなり高い学習者でも、ごく少数を除き完全に母音の挿入がなくなることはなく、誤挿入されたの長さを”挿入母音を含む単語の長さを音素数で割った長さ”に対する比率に換算して正規化した時間長を習熟度別に比較した結果も、習熟度による挿入母音の時間長に有意な差は見られなかった ($t(58.911) = -1.4793, p > .1$)。つまり、英語の習熟度が高くなっても、母語である日本語の音節構造の影響が強くみられ、上級話者であっても母音を挿入するときはしっかりと挿入していることが分かる。

(3) 日本語訛の英語の評価

日本語母語話者 183 名が”The North Wind and the Sun” を音読したものに、英語母語話者 25 名(方言は異なる)の音読データを加えた計 208 名分の音声データを、母語の異なる英語音声の専門家 16 名(アメリカ英語話者 4 名、日本語話者 4 名、北京方言話者、広東方言話者、ドイツ話者、ポーランド話者、スペイン語、フランス語、韓国語、パンジャブ話者各一名)が、(a)分節音の正確さ、(b)リズム、イントネーション等の韻律、(c)流暢さ、(d)母語訛の少なさの四項目について、1-10 の十段階で評定した。内、日本語母語話者のみの発話の評定値について、評定者の母語の違いにより日本語訛の発音の評定値がどの程度異なるかを検証した。結果としては、全体的に見て比較的相関は高く、0.7 以上のものが多く、0.6 以下は非常に少ない。特に、母語訛の少なさは評定値の一致度が他の項目よりも高い傾向がみられる。特にアメリカ人評定者間の相関は比較的高く、母語訛の少なさの相関は非常に高く、全員 0.81 以上であった。

評定者間で個人差はあるが、母語を共有する評定者同士の評価の相関は、母語を共有しない評定者との相関よりも若干高めであったが、母語や類型的な言語の音韻の違いによる差はそれほど大きいとは見受けられなかった。おそらく、評定者は全員（英語）音声学の専門家であったために、同じ基準で評価をしていた可能性が高い。今後は、英語の非専門家、一般の英語話者が日本語訛の英語をどう評価するのか、検証する必要があると思われる。

(4) 日英語のバイリンガル合成音を使った知覚実験

これまでの第二言語音声習得とその評価の研究を踏まえ、deep-learning-based speech synthesisを応用した音声合成システム（Cyborg speech synthesis system, 図3）を使い、実際に外国語訛の強さの度合いを制御した多言語音声合成を使って、まずアメリカ英語訛の度合いを制御した日本語音声を作成し、合成音で知覚実験を行った。

日本語とアメリカ英語のバイリンガル話者一人が、双方の言語の訛のない日本語と英語を発話したものを使い、

DBLSTM-base の音響モデルを訓練した。このバイリンガル話者の録音音声を使い、分節音の長さや

ピッチ曲線を測定し、それを基に自然な韻律の合成発話を生成し、自然音声と合成音声の特性を併せ持った音声を生成した。

上記(1)-(3)の結果を基に、日本語と英語の分節音で調音が明らかに異なるもの、つまり異なる音の音質が異なるもの、かつ、英語母語話者の日本語学習者の発話エラー

分析から、発音の間違が多いと報告されている子音を抽出し、日英バイリンガル一人の話者の発話音声をもとに、分節音を置き換えた。

分節音は、ラ行子音（J 歯茎弾き音→E 後部歯茎接近音）、「シ」の子音（J 無声歯茎硬口蓋摩擦音→E 後部歯茎摩擦音）、ザ行子音（J 歯茎破擦音→E 歯茎摩擦音）、「ジ」の子音（J 有声歯茎硬口蓋破擦音→E 硬口蓋歯茎破擦音）を用いた（J=日本語、E=英語）（表1）。

合成音した日本語音声を、クラウドソースを使い131人の日本語話者が、分節音の正確さ（5段階のリッカート尺度）と英語訛の度合い（7段階のリッカート尺度）を判定した。日本語の発話の刺激音は77%が外国語訛がないと判断された。音素に関しては、/r/の音質が外国語訛の判断に与える影響が非常に大きく、後部歯茎接近音の度合いが23%を超えると「アメリカ人の発話」と判断される割合が圧倒的に高かった。つまり日本語話者は、自分たちが英語の分節音の発音で苦手とされる音素の発音に気づいていて、音質の違いに敏感に反応していることが分かった。

同じ方法で、日本語訛の程度を変化させた英語音声を使った実験を行い、結果を学会に投稿中である。

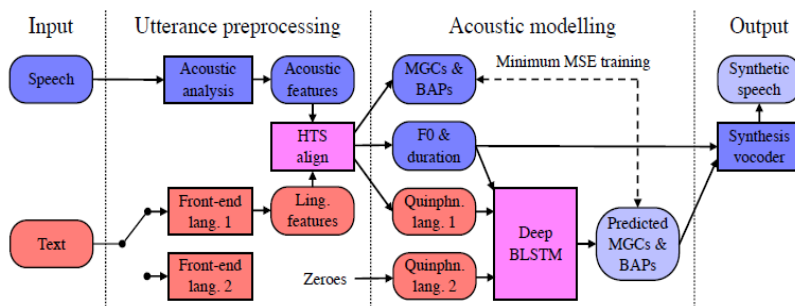


図3. Cyborg speech synthesis system.

音声データ(楕円)、プロセス(四角)、音声(薄青)、テキスト(赤)、音声とテキストの両方(ピンク)。

Japanese		English		Substitutions made	
IPA	Open JTalk	IPA	Combilex GAM	Max	No. prompts
r	r	ɹ	r	9	19
ç	sh	ʃ	s	8	13
dz	z	z	z	5	7
dz	j	dʒ	dʒ	3	8
tç	ch	tʃ	tʃ	2	11

表1. 日英語の L2 発話で問題となる子音と合成音での使用回数

5. 主な発表論文等

[雑誌論文] (計 18 件)

- ① Yazawa, K., Whang, J., Kondo, M., & Escudero, P. (2019). Language-dependent cue weighting: An investigation of perception modes in L2 learning. *Second Language Research*. SAGE Journals. <https://doi.org/10.1177/0267658319832645> (査読有)
- ② Kondo, M. and Konishi, T. (2018) “Acquiring L2 phonemes and recognition of their allophonic variances”, Botinis, Antonis (ed.), Proceedings of 9th Tutorial and Research Workshop on Experimental Linguistics, pp. 69-72, 28-30 August, 2018. (Ebook ISSN: 2529-1092) (査読有)
- ③ Henter, G. E., Lorenzo-Trueba, J., Wang, X., & Yamagishi, J. (2018) “Deep Encoder-Decoder Models for Unsupervised Learning of Controllable Speech Synthesis”, arXiv preprint arXiv:1807.11470, 17 pages, 2018. <https://arxiv.org/abs/1807.11470> (査読有)
- ④ Henter, G. E., Lorenzo-Trueba, J., Wang, X., Kondo, M. & Yamagishi, J. (2018) "Cyborg Speech: Deep Multilingual Speech Synthesis for Generating Segmental Foreign Accent with Natural Prosody," 2018 ICASSP, Calgary, Canada, April, 2018, pp. 4799-4803. (査読有)
- ⑤ Konishi, T., Yun, Jihyeon and Kondo, M. (2018) “Acoustic correlates of L2 English Stress-Comparison of Japanese English and Korean English”, *Phonetics and Speech Sciences*, Vol. 10, No. 1, pp. 9-14. (pISSN 2005-8063/eISSN 2586-5854; March, 2018. (査読有)
- ⑥ Yazawa, K., Ozaki, Y., Short G., Kondo, M. & Sagisaka, Y. (2015), “A Study of the Production of Unstressed Vowels by Japanese Speakers of English Using the J-AESOP Corpus”, *Proc. of Oriental COCODA*, IEEE. pp.96-100. Shanghai, China. 28-30 October, 2015. (査読有)

[学会発表] (計 21 件)

- ① Kondo, M. “Teaching English pronunciation: features of Japanese learners and corpus-based research”, International Workshop on “Pronunciation teaching, corpora and automatic tools: English, French and Japanese”, Chambre de Commerce et d’Industrie, Montauban, France, 11 March, 2019. (Invited talk)
- ② Henter, G.E., Lorenzo-Trueba, J., Wang, X., Kondo, M., Yamagishi, J. “Generating segment-level foreign-accented synthetic speech with natural speech prosody”, 情報処理学会 第 120 回 音声言語情報処理研究会, 2018-MUS-118(8), pp.1-3, 2018 年 2 月 20 日 .
- ③ Kondo, M., “Asymmetrical interpretation of English liquid consonants by Japanese speakers”, *The Journal of the Acoustical Society of America*. November, 2016. (Invited talk)
- ④ Kondo, M. & Short, G., “Perception of English liquid consonants by Japanese speakers: Evidence from mimicry speech”, New Sounds 2016, Aarhus, Denmark, 10-12 June, 2016.

[図書] (計 2 件)

- ① 近藤真理子『第 2 章 第一・第二言語の音声特性と音声習得』, 西原哲雄・福田稔・早瀬尚子・谷口一美 (監修), 言語研究と言語学の進展シリーズ①, 言語の構造と分析- 統語論、音声学・音韻論、形態論 (西原哲雄 編), 第 II 部『最新の音声学・音韻論研究の進展』(都田青子・近藤真理子・西原哲雄著), 開拓社:東京. 2018 年 ISBN: 9784758913713 (査読無)
- ② 近藤真理子「日本語話者の英語流音の認識」, 西原哲雄・田中真一・早瀬尚子・小野隆啓(編) 開拓社叢書 29『現代言語理論の最前線』所収, pp. 147-163, 開拓社:東京, 2017 年 ISBN : 978-4-7589-1824-4 (査読無)

[その他]

6. 研究組織

(1) 研究分担者

研究分担者氏名 : 山岸順一

ローマ字氏名 : YAMAGISHI Junichi

所属研究機関名 : 国立情報学研究所

部局名 : コンテンツ科学研究系

職名 : 教授

研究者番号 (8 桁) : 70709352

(2) 研究協力者 (他 6 名)

研究分担者氏名 : ヘンター・グスタフ E.

ローマ字氏名 : Henter・Gustav Eje