

## 科学研究費助成事業 研究成果報告書

令和 5 年 6 月 12 日現在

機関番号：34419

研究種目：基盤研究(C) (一般)

研究期間：2017～2022

課題番号：17K02957

研究課題名(和文)外国語学習におけるTTS合成音声の有効性と可能性

研究課題名(英文)Effectiveness and Potential of Synthetic Text-to-Speech in Foreign Language Learning

研究代表者

松田 紀子(Matsuda, Noriko)

近畿大学・総合社会学部・准教授

研究者番号：40759007

交付決定額(研究期間全体)：(直接経費) 2,700,000円

研究成果の概要(和文)：本研究では、日本人英語学習者を対象に複数の心理言語学実験を用い、外国語学習におけるTTS(Text-to-Speech)合成音声の有効性と可能性を、特に知覚学習という観点から明らかにしようと試みた。その結果、学習者にとってパラ言語情報と言語情報の独立した処理は難しい可能性が高いこと、TTS合成音声を使用したリスニング活動は、特に知覚処理に問題を抱えやすい学習初期段階に有効である可能性が高いことが示された。加えて、知覚学習の効果を高めるには、意味よりも音声に注意を向けたリスニング活動が望ましいこと、十分に見聞きしたことがある、つまり親密度が高いと学習者が感じる教材の使用が望ましいことが示唆された。

研究成果の学術的意義や社会的意義

本研究では、近年、目覚ましい進歩を遂げているTTS(Text-to-Speech)合成音声の外国語学習における活用、特に学習初期段階の知覚学習における活用の有効性と可能性を、聴覚性プライミング効果と単純接触効果の実験パラダイムを使用して探り、一定の知見を提供できたものとする。また、その知見に基づいたTTS合成音声の使用方法についても、一定の知見を示した。日本のように日常生活で外国語に触れる機会が少ない環境において、外国語学習へのTTS合成音声の活用は、音声インプットの絶対量の不足を補う方法の1つとして今後ますます期待されるだろう。

研究成果の概要(英文)：This study aimed to examine the effectiveness and potential of synthetic Text-to-Speech (TTS) in foreign language learning, particularly for perceptual learning, using multiple psycholinguistic experiments with Japanese learners of English. The results demonstrated that the learners had difficulty processing paralinguistic and linguistic information independently. Moreover, the findings indicated that listening activities utilizing synthetic TTS were effective, particularly in the early stages of learning, when learners had problems with perceptual processing. In addition, the results suggest that learners should focus on the speech during listening activities and use familiar materials to increase the effectiveness of perceptual learning.

研究分野：応用言語学

キーワード：Text-to-Speech 合成音声 知覚学習 Talker variability プライミング効果 定型表現 単純接触効果 オンライン実験

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属します。

### 1. 研究開始当初の背景

外国語学習者が目標言語を正確に聞き取って内容を理解できるようになるためには、音声言語の処理過程の中で基礎となる音声知覚段階を自動化すること(意識的な処理が、反復練習によって速く、無意識の処理に移行すること)が重要であると言われている。特に学習初期段階では、リスニングにおける問題の多くが知覚の誤りに基づいている(Field, 2019)ため、リスニング力の伸長には、知覚処理を促すためのリスニング活動が欠かせない。しかし、日本のように日常生活で外国語に触れる機会が少ない環境では、知覚処理を促すために必要な音声インプットの絶対量が不足しやすい。近年、目覚ましい進歩を遂げているTTS(Text-to-Speech)合成音声の活用は、その不足を補うための解決策の1つと考えられる。しかし、TTS合成音声を語学学習に利用した際に得られる学習効果を検証した研究は数少なく、知覚学習効果について検討した研究はほとんどなされていない。

### 2. 研究の目的

本研究の目的は、日本人英語学習者を対象に、複数の心理言語学実験を用いて外国語学習におけるTTS合成音声の有効性と可能性を、特に知覚学習という観点から明らかにすることである。その成果からTTS合成音声を用いて知覚学習を効率的におこなうための知見と方法を外国語学習者に提供したいと考えている。

具体的には、心理言語学実験を用いてTTS合成音声を音声インプットとして活用した場合の英語の単語または句単位以上(定型表現)の知覚学習効果を測る。その際、各種実験を用いて、知覚学習を促進または抑制する可能性のある、学習内容に関連する要因である①音声の特異性(声・性別・なまり等の違い)と②音声提示の際に組み合わせるタスクの複雑さ、学習者に関連する要因である③学習者の習熟度と④学習期間の影響を探る。

### 3. 研究の方法

学習者の知覚学習効果を測るため、単語単位と句単位以上(定型表現)に分けて実験を実施した。単語単位では聴覚性プライミング実験を実施し、学習内容に関連する要因である①②と、学習者に関連する要因である③の影響を探った。①の音声の特異性、特に知覚処理を考える上で重要なパラ言語情報である「話者の変化(Talker variability)」に関しては、先んじて実施したTTS合成音声を使用した実験結果との比較を念頭に置き、英語母語話者の肉声を録音して使用(発話速度、音の大きさについて可能な限り統制)した。心理学実験ソフトウェアを使用してこれらの音声を提示する際、②の組み合わせるタスクの複雑さを変えるために、音声に注意を向ける押韻判断課題(単語のペアが韻を踏んでいるかどうかを判断する課題)と、意味に注意を向ける類義性判断課題(単語のペアの意味が似ているかどうかを判断する課題)の2つの課題を使用した。参加者は海外生活経験や留学経験のない日本人英語学習者40名で、前述の各課題を終了した後(順番はカウンターバランス)に、次の課題として、出現した単語(そのうちの半分は話者を変化させる)と新出の単語を混ぜて復唱した。各単語の音声提示(ランダム)から復唱までの反応速度を測り、新出の単語の平均値から既出の単語の平均値の差分を用いてプライミング効果を測った。

句単位以上(定型表現)では、聴覚性プライミング実験を実施するのは困難であったため、知覚学習効果を間接的に測定できる可能性が高い単純接触効果(Zajonc, 1968)の実験パラダイムを採用した。単純接触効果の実験パラダイムでは、ある対象への反復接触により、知覚的流動性が向上(=知覚学習の促進)することで対象への好意度が高まると考えられている。先行研究では英語の定型表現(書きことば)を用いた実験で親密度(「どの程度見聞きしたことがあるか」という心理的接触回数を評価したもの)と好意度に正の相関がみられ、累積的な単純接触効果がみられることが分かっている(松田・村田, 2022)。そのため、話しことばにおける累積的な単純接触効果の有無の確認も兼ねてこの実験パラダイムを採用し、ここ数年で人の肉声に近い自然な声を生成できるようになった人工知能(AI)の技術を使用したTTS合成音声で諸条件を統制した定型表現のグループを作成し、知覚学習効果を検証した。新型コロナウイルス感染症への対策によって、対面での実験の実施が長期にわたって困難となったことを鑑み、オフライン(対面)実験だけではなく、心理学実験ソフトウェアとサーバーサービスを併用し、オンライン実験も実施した。

1つ目の実験では、各条件を統制した英語の定型表現を既存のリストに基づいて親密度が高いものと低いもの、各2グループずつ、計4グループ作成し、学習内容の要因である①音声の特異性、特に「話者の変化(Talker variability)」に関して、単語単位で実施した実験と同じ結果が出るかどうかをオフライン実験で検証した。参加者は日本人英語学習者25名で、8種類のTTS合成音声で作成したフレーズが2回ずつランダムに音声提示された。参加者は1回目の提示では親密度評価、2回目の提示では好意度評価を求められた。

2つ目の実験では、1つ目の実験を踏まえ、反復接触を経て好意度が上昇するか(=知覚学習が促されるか)をオンライン実験で検証した。参加者は1つ目の実験に参加した15名で、①の影響を検証するために同じ実験方法を用いた。ただし、定型表現の音声については、1つ目の実

験と同じ声で2回提示されるグループと、1つ目の実験と異なるように4回話者が変化するグループを作成して用いた。

3つ目の実験では、研究代表者の肉声から合成音声を作成し、それを用いて各条件を統制した英語の定型表現を既存のリスト（1つ目と2つ目の実験とは異なる）に基づいて、親密度が高く語順が正確なもの、親密度が低くて語順が正確なもの、親密度が高く語順が不正確なもの、各3グループを作成し、学習者に関連する要因である③学習者の習熟度と④学習期間の影響を検証した。参加者は日本人英語学習者で、研究代表者の肉声を授業で聞いた長さが異なるグループから成る。参加者は定型表現を2回ずつランダムに音声提示されたが、1回目の提示では親密度評価、2回目の提示では好意度評価を求められた。

#### 4. 研究成果

##### (1) 単語単位の聴覚性プライミング実験

図1はプライミング効果を量的に示したものである。実験の結果、音声に注意を向けた場合、話者が同じ場合(図1の Same)はプライミング効果が見られ、話者が変化する(図1の Different)とプライミング効果が見られなくなった。そして意味に注意を向けた場合(図1の Synonym)、話者の変化による影響はなくなり、プライミング効果はわずかしか見られなかった。意味処理に注意が偏り、トップダウン中心の音声処理をしたことで音声への注意が低下したことが理由と考えられる。これらの結果から、学習者は話者によって異なる音響特徴と言語情報を独立的に処理することが難しい可能性が高いことが分かる。母語の場合は話者の変化を含むパラ言語情報と言語情報とは独立的に処理されていると言われており、音声もしくは意味に注意を向けてもプライミング効果は変わらない。しかし、学習者の場合は限られた認知資源をどの言語情報(本実験では音声か意味か)に割り当てるかでパラ言語情報を含めた音声処理の内容が大きく変わる可能性が高いことがわかった。これらの結果から、学習者の知覚学習を促すには、特に音声に注意を向けたリスニング活動をする必要があると指摘できる。

なお、③に関連して、TTS 合成音声を使用して音声に注意を向けた場合、特に習熟度が低い学習者において高い知覚学習効果が認められた(Matsuda, 2017a)。TTS 合成音声は様々な知覚的な情報が統一されているために、習熟度が低い学習者にとっては肉声よりも音声処理をしやすい可能性が高いと考えられる。学習者にとって多様な知覚的情報を持つ外国語の肉声に触れることは習熟度に関係なく重要であるが、Goh (2000) が指摘しているように、習熟度が低いほど知覚処理に関する問題をより多く抱えていることを踏まえると、学習初期段階の学習者にとって、肉声の多様さは聞き取りづらさを感じさせ、リスニング不安や苦手意識を感じさせるきっかけになる可能性がある。そのため、学習初期段階においてTTS 合成音声を使用し、音声に注意を向けたリスニング活動を介在させることで、知覚の誤りによるつまづきを減らす工夫ができるのではないかと考える。

##### (2) 句単位以上(定型表現)の単純接触効果の実験

1つ目の実験の結果、親密度が高い定型表現で、1回目と2回目に話者が同じ場合に好意度が高くなることがわかった。親密度(「どの程度見聞きしたことがあるか」)の評価の際、学習者は意味に注意を向ける必要がなく、主に音声に注意を向けていると考えられる。本実験の結果は、単語単位の実験において音声に注意を向けた場合、話者が同じ場合プライミング効果が見られ、話者が変化するるとプライミング効果が見られなくなった実験結果と合致する。なお、単語単位のプライミング実験では、実験方法のリミテーションによって親密度が高い単語のみを使用していた。そのため、TTS 合成音声の使用によって知覚学習が促進される可能性がより高いのは、十分な接触回数を経た(と学習者が評価する)定型表現であることは新たな知見となる。

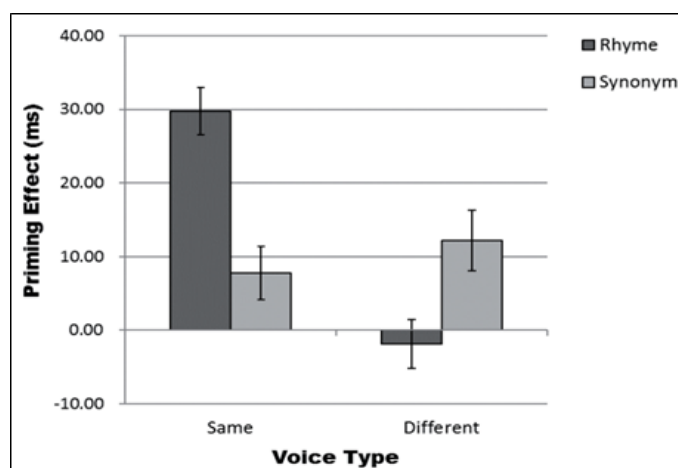


図1 プライミング効果 (Matsuda, 2017b, p. 14 より抜粋)

2つ目の実験の結果、1つ目の実験と同様に、親密度が高いフレーズにおいて、同じ声で提示された場合に好意度が上昇することがわかった。TTS 合成音声を用いたオンラインの実験で好意度が上昇したことは、新型コロナウイルス感染症への対策のために定着したオンラインでの学習における可能性を広げるものと考えられる。1つ目の実験の結果を合わせると、TTS 合成音声を使用して、知覚学習の促進のために音声に注意を向けたリスニング活動をする際には、十分に見聞きしたことがあると学習者が感じる教材、もしくは感じさせるまで繰り返し提示した教材を使用することが肝要であることを示している。

3つ目の実験の結果、参加者の習熟度が高い場合と、研究代表者の声による学習期間が長い場合に3グループ、つまり親密度が高く語順が正確なもの、親密度が低く語順が正確なもの、親密度が高く語順が不正確なもの順で好意度評定値が安定し、3グループの評定値の差が開くことがわかった。今後は、回数を重ねて好意度が上昇するかどうかを検討する必要があると考える。以上の実験1-3は定型表現の知覚学習効果について見たものだが、今後は文章単位での知覚学習効果を見ていく必要があると考える。

本研究は、日本人英語学習者を対象に、複数の心理言語学実験を用いて外国語学習におけるTTS 合成音声の有効性と可能性を、特に知覚学習という観点から明らかにしようと試みた。その成果からTTS 合成音声を用いて知覚学習を効率的におこなうための知見と方法を示したが、同時に課題も生まれており、新たな研究課題として発展させる必要があると考える。聴覚性プライミング実験及び単純接触効果の実験から得られた成果や知見は、関連学会のシンポジウムや学会等で発表し、論文等として成果を公開した。

#### 引用文献

- Field, J. (2019). Second language listening: Current ideas, current issues. In J. Schwieter & A. Benati (Eds.), *The Cambridge handbook of language learning* (pp. 283-319). Cambridge University Press. <https://doi.org/10.1017/9781108333603.013>
- Goh, C. C. M. (2000). A cognitive perspective on language learners' listening comprehension problems. *System*, 28(1), 55-75.
- Matsuda, N. (2017a). Evidence of the effects of text-to-speech synthetic speech to improve second language learning, *JACET Journal*, 61, 149-164. [https://doi.org/10.32234/jacetjournal.61.0\\_149](https://doi.org/10.32234/jacetjournal.61.0_149)
- Matsuda, N. (2017b). Impact of talker variability on L2 word recognition among Japanese EFL learners, *Vocabulary Learning and Instruction*, 6 (2), 8-22. <https://doi.org/10.7820/vli.v06.2.Matsuda>
- 松田紀子・村田健太郎 (2022). 「会えば会うほど好きになる」—英語の定型表現の習得における単純接触効果—. 『近畿大学総合社会学部紀要：総社る』 11(1), 39-49. <https://doi.org/10.15100/00022888>
- Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of personality and social psychology*, 9, 1-27.

5. 主な発表論文等

〔雑誌論文〕 計1件（うち査読付論文 1件/うち国際共著 0件/うちオープンアクセス 1件）

1. 著者名 Noriko Matsuda	4. 巻 6(2)
2. 論文標題 Impact of talker variability on L2 word recognition among Japanese EFL learners	5. 発行年 2017年
3. 雑誌名 Vocabulary Learning and Instruction	6. 最初と最後の頁 8-22
掲載論文のDOI（デジタルオブジェクト識別子） 10.7820/vli.v06.2.Matsuda	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

〔学会発表〕 計5件（うち招待講演 2件/うち国際学会 0件）

1. 発表者名 Noriko Matsuda
2. 発表標題 Familiarity and frequency effects on L2 multiword sequence processing online: An empirical study using TTS voices
3. 学会等名 第241回JACET関西支部リーディング研究会7月研究例会
4. 発表年 2019年

1. 発表者名 Noriko Matsuda
2. 発表標題 Familiarity and frequency effects on on-line processing of L2 formulaic sequences
3. 学会等名 外国語教育メディア学会（LET）関西支部基礎理論研究部会10月例会
4. 発表年 2018年

1. 発表者名 松田紀子
2. 発表標題 外国語学習におけるTTS合成音声の有効性と可能性
3. 学会等名 2017年度 藍野大学中央研究施設シンポジウム 藍野大学
4. 発表年 2017年

1. 発表者名 Noriko Matsuda
2. 発表標題 Impact of talker variability on L2 word recognition among Japanese EFL learners
3. 学会等名 2017 Vocab SIG Symposium by JALT Vocabulary SIG (招待講演)
4. 発表年 2017年

1. 発表者名 松田紀子
2. 発表標題 第二言語の音声処理における潜在記憶の役割について TTS合成音声を使用した実験的アプローチによる研究から
3. 学会等名 外国語教育メディア学会 (LET) 関東支部 音声映像研究研修部会講演会 (招待講演)
4. 発表年 2018年

〔図書〕 計1件

1. 著者名 丸橋良雄 湯谷和女 坂元敦子 松田紀子	4. 発行年 2018年
2. 出版社 英光社	5. 総ページ数 179
3. 書名 比較文化の語らい	

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究協力者	中田 ひとみ  (Nakata Hitomi)		

6. 研究組織（つづき）

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究協力者	梶浦 眞由美  (Kajiura Mayumi)		

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関