

部分字幕を用いたシャドーイング練習システムの検討

方 漱玉^{†1} 能勢 隆^{†1} 伊藤 彰則^{†1}

概要: シャドーイングは、聞こえてきた音声のすぐあとを追って同じ音声を発するという言語の勉強法であり、スピーキングとリスニング力を向上させるためのトレーニングとして注目されている。しかし、第二言語学習者にとって、元音声についていけず、復唱を続けられないという問題があり、練習は難しい。そこで本研究では、学習者の補助を目的とした、部分字幕を用いたシャドーイング練習システムを提案する。音声認識 API を利用して部分字幕を生成する日本語学習者向けのシステムを構築した。部分字幕と字幕無しの 2 つのグループで比較実験を行った結果、字幕無しグループより、部分字幕グループのシャドーイング遂行成績が高いことを示す。

キーワード: シャドーイング練習, 日本語学習者, 部分字幕, 音声認識

A Study on the Shadowing Practice System Using Partially Synchronized Caption

SHUYU FANG^{†1} TAKASHI NOSE^{†1} AKINORI ITO^{†1}

Abstract: Speech shadowing is a language learning technique in which subjects repeat speech at a delay to the onset of hearing the phrase. It has been attracting attention as training to improve speaking and listening skills. However, for second language learners, it is difficult to practice because they cannot keep up with the original speech and cannot continue to repeat it. This study proposes a shadowing practice system using partially synchronized captions to help learners. We built a system for Japanese learners which generates partial captions using the speech recognition API. The result of a comparative experiment between the two groups with partially synchronized caption (PSC) and no caption (NC) shows that the performance of PSC group is better than NC group.

Keywords: Speech shadowing, L2 learners of Japanese, partially synchronized caption, speech recognition

1. はじめに

近年、アニメを代表とする日本のポップカルチャーが世界中に発展するに従い、日本語を勉強する人が増えている。国際交流基金の調査によると、2018 年に海外日本語教育機関調査で日本語を勉強した人は約 384 万人で、142 カ国に分布している。リスニング、リーディング、スピーキング、ライティングは、言語学習の基本的な勉強方法である。しかし、リーディングと比較して、リスニングとスピーキン

グは練習が難しい。本研究は、日本語学習者のリスニングとスピーキング能力を向上させる方法の開発を目的とする。

「シャドーイング (Shadowing)」は、図 1 に示すように、言語学習時に、音声を聞きながらそれをまねて発音する通訳訓練法のことである。シャドーイングはスピーキング／リスニング力を向上させるためのトレーニングとして注目されている。音声を聞き終えてから繰り返す「リピート」とは異なり、シャドーイングは、聞こえてくる音声のすぐ後ろを影 (shadow) のように追いかけるのがポイントである。シャドーイングにはアクセント、イントネーションの改善、聴解力の向上、スピーキング力の向上、語彙、表現の習得などの利点があるが [1], 「聞く」、「発音する」を

^{†1} 現在、東北大学大学院工学研究科, 仙台市
Presently with Graduate School of Engineering, Tohoku University, 6-6-05 Aramaki aza Aoba, Aoba-ku, Sendai-shi, 980-8579 Japan



図 1 シャドーイング
Fig. 1 Speech Shadowing

同時に行うため、難易度の高い学習法である。学習者が元音声に追いつけない場合、復唱を続けられないことがよくある。経験豊富なシャドーイングチューターは、学習者がフォローできない時、元音声を一時停止したり、ヒントを与えたり、音声を繰り返したりすることによって、そのような問題を解決することができる。ただし、経験豊富な講師を見つけるのは難しい場合があり、費用が非常に高くなる可能性がある。したがって、シャドーイングを補助するシステムが有用である。既存のシャドーイング練習システムは、話すスキルの向上に焦点を当てている [2], [3], [4], [5]. 例えば、WithYou[2] は、学習者の即時の発音パフォーマンスに応じて音声再生の難易度を自動的に調整する練習システムである。ただし、このようなシステムは、リスニングの練習を適切にサポートしていない。発音、アクセント、イントネーションなどの発音能力の向上だけでなく、学習者のボトムアップのリスニングスキル、つまり音素から単語を認識する能力にも向上させる対象とするシステムも必要である [6].

本研究では、話すスキルの向上だけでなく、リスニングスキルを向上させることも重視する。そこで、部分字幕がリスニング練習に適している [7] ことを考慮し、部分字幕をシャドーイング練習に導入する。提案した部分字幕を用いたシャドーイング練習システムを構築して、システムの有効性を検討するために、小規模実験を行う。

2. 先行研究

2.1 部分字幕

映像に付与された字幕は、元は聴覚障害者の補助として用いられてきたが、字幕が普及し始めた 1980 年代から外国語学習教材としても用いられるようになる。一般的な字幕は、図 2 のような完全な字幕である。

しかし、完全な字幕をリスニング練習資料として使うと、学習者は字幕を読むことに集中するため、リスニング改善の効果が下がる [8]. この問題において、第二音声学習者の聞き取り能力を向上させるために、図 3 を示すよ

*1 <https://plus1world.com/netflix-english-japanese-subtitles>



図 2 字幕*1
Fig. 2 Caption

うに部分字幕システム (Partial and synchronized caption system[7]) が提案された。



図 3 英語部分字幕システム [7]
Fig. 3 PSC System[7]

このシステムは TEDtalks をリスニング練習資料として、最初は音声認識システムを使って完全な字幕を生成する。その上で、話速、単語の使用率 (頻度)、特異性 (学術用語) による語彙を選択して、部分字幕を生成する。

このシステムの有効性評価実験では、被験者は完全な字幕 (FC)、部分字幕 (PSC) と字幕無し (NC) の 3 つの条件で、ビデオの最初から 70% の内容を視聴した。その後、ビデオに関する理解度テストを実施する (PART1)。次に、全ての被験者は字幕無しの条件で最後まで 30% のビデオを見て、理解度テストを実施する (PART2)。

実験結果によると、PSC を使う被験者の PART1 のスコアは、NC グループより高い。FC グループと比較すると、PSC を使用する初級学習者のスコアが高く、中級学習者のスコアが低く、上級学習者のスコアがほぼ等しい。一方、PSC グループの PART2 のスコアは、学習のレベルに関係なく、FC と NC の両方より高い。つまり、リスニングを練習する場合は、しばらく部分字幕で練習してから字幕無しで練習すると、いつも字幕無しより良い結果が得られることを示す。

2.2 シャドーイング

シャドーイングの練習方法は、テキストを提示するか、テキスト提示のタイミング、どの段階で意味確認を行うか、フィードバックの仕方などにより、シンプルなものから、細かいステップに分けられたものまで、様々である。

音声的な要素に注意を向けて行うプロソディー・シャ

ドーイング (prosody shadowing) は、テキストを与えず耳からだけの情報により音声を復唱する。聞くことと発音するを同時に行うため、学習者の負担は重いと考えられる。パラレル・リーディングでは、テキストを見て音声と同時に発音する。岩下 [9] は、シャドーイング時にテキストを呈示すると、テキストが口頭産出の手掛かりとなり、通常のシャドーイングの場合より正確にモデル音声を産出できると述べている。岩下によると、テキスト提示は、上手にシャドーイングができない学習者の口頭産出を助け、課題遂行の負担を減らすという利点がある。しかしその一方、テキストは学習者を文字情報に頼らせて、モデル音声を注意して聞けなくしてしまう危険をはらんでいる。テキスト呈示のタイミング1つにしても様々なパターンがあり、その効果も異なるので、対象となる学習者にあった方法を検討する必要がある。また、学習者自身にいろいろな方法を試させ、ストレスを感じない程度の難しさと自分に合った方法を見つけることが必要かもしれない。そこで、本研究では、テキストを提示する一つの方法として、部分字幕の使用を提案する。

3. 提案法

3.1 従来のシャドーイング練習方法

教育現場でシャドーイングを導入する際は、以下の順に行うことが多い [10], [11], [12].

(1) リスニング

まずはテキストを見ずに全文を聞く。何の話題について話されているのか、ざっくりと全体の意味をつかんでみる。

(2) マンブリング

元音声を聞いて、小声でシャドーイングを行う。

(3) パラレル・リーディング

テキストを見ながら行うシャドーイングを行う。

(4) 意味確認

トランスクリプトを見て、文章全体の内容やメッセージを把握したり、聞き取れなかったり単語やわからない単語の意味を確認する。

(5) プロソディ・シャドーイング

テキストを見ずに耳からだけの情報により音声を復唱する。聞こえてきた音声と全く同じ発音・イントネーションでできるだけ正確かつ素早く復唱する。

(6) コンテンツ・シャドーイング

流れてきた音声の意味を理解しイメージを頭で思い浮かべるシャドーイングをする。

しかし、これが唯一の手順というわけではなく、李 [13] は、学習者が内容を 100 パーセント理解している教材を使うことが前提であるとしつつも、リスニング後に内容理解を口頭で行い、パラレル・リーディングをせずにシャドーイングを行うという手順を用いている。このように、シャドーイングの遂行手順については、教師の体験的知見によ

るものが多く、実証的に検討されたものは少ない。また、どちらか一方の手順が優れているというものでもなく、それぞれの目的によって異なるといえよう。

上述の2つの手順の違いとしては、大きくパラレル・リーディングの有無があげられる。この違いは、シャドーイングの遂行においてテキストを視覚呈示すべきか否か、あるいはシャドーイングの遂行時に視覚呈示のテキストをどのように使用すべきか、という重要な論点を示している。ここで、本研究では、部分字幕を一つのテキストの視覚呈示方法として提案する。

3.2 提案法のシャドーイング練習

(1) リスニング

(2) 字幕を見ないシャドーイング

(3) テキストを確認して意味を理解する

(4) 部分字幕を見るシャドーイング ((2) の字幕を見ないシャドーイング練習の時に、ついていけない、発音を間違った部分を字幕で示す)

(5) 字幕を見ないシャドーイング

4 と 5 の練習を繰り返す。

3.3 部分字幕の生成

提案システムでの部分字幕の生成方法を図 4 に示す。まず学習者の練習音声を録音する。AmiVoice API を使って、元音声と学習者の練習音声を音声認識する。音声認識結果を比較して、認識できなかった単語からかぞえて、最大 5 個の単語を字幕として示す。次の練習では学習者がこの字幕を見てシャドーイングをする。このようにした理由は、すべての誤認識単語を表示してしまうと、うまく発音できなかった箇所から後ろのほぼすべての単語が字幕として表示され、パラレル・リーディングと同様な条件になってしまうためである。

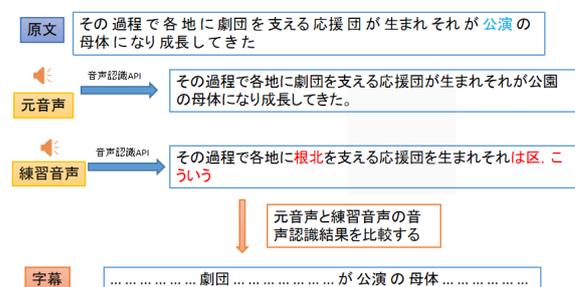


図 4 部分字幕の生成

Fig. 4 Generation of Partial Synchronized Caption

4. 部分字幕を用いたシャドーイング練習システムの有効性評価実験

4.1 練習システムの構築

提案するシャドーイング練習システムの構築を行った。字幕無しと部分字幕の練習画面を図 5 と図 6 に示す。



図 5 無字幕の練習画面

Fig. 5 Practice without Caption



図 6 部分字幕の練習画面

Fig. 6 Practice with Partial Synchronized Caption

4.2 実験手順

日本語能力試験 N1/N2 の日本語学習者 6 名が実験に参加した。6 名の被験者を、字幕無しシャドーイング練習 (グループ 1) と部分字幕を見るシャドーイング練習 (グループ 2) の 2 つのグループに分けた。

シャドーイングに用いた教材は新聞記事読み上げ音声コーパス (JNAS) の 186 文、一文の長さは 1~20 秒、10~99 拍で、平均 42 拍程度である。

練習手順は次のとおりである。

グループ 1 の練習手順:

- (1) リスニング。
- (2) 字幕を見ないシャドーイング (事前トレーニングレベルとして記録する)。
- (3) テキストを確認して意味を理解する。
- (4) 字幕を見ないシャドーイング。

- (5) 字幕を見ないシャドーイング。
- (6) 字幕を見ないシャドーイング。
- (7) 字幕を見ないシャドーイング。
- (8) 字幕を見ないシャドーイング。
- (9) 字幕を見ないシャドーイング (事後トレーニングレベルとして記録する)。

グループ 2 の練習手順:

- (1) リスニング。
- (2) 字幕を見ないシャドーイング (事前トレーニングレベルとして記録する)。
- (3) テキストを確認して意味を理解する。
- (4) 部分字幕を見るシャドーイング。
- (5) 字幕を見ないシャドーイング。
- (6) 部分字幕を見るシャドーイング。
- (7) 字幕を見ないシャドーイング。
- (8) 部分字幕を見るシャドーイング。
- (9) 字幕を見ないシャドーイング (事後トレーニングレベルとして記録する)。

トレーニングのレベル評価には WRR (Word Recognition Rate, 単語認識率) を利用した。WRR は、学習者発話を自動音声認識したときの認識率である。学習者が毎回シャドーイング練習で正しく復唱できる単語の割合によって、シャドーイング遂行成績を評価する。事前と事後のトレーニングレベルの差を計算して、字幕無しグループと部分字幕グループの練習結果を比較する。

4.3 今回の実験結果と検討

字幕無しグループの被験者 3 名のシャドーイング練習結果の平均 WRR を図 7 に示す。1 回目から 7 回目まで、WRR 値は全体的に上昇傾向を示した。WRR 値は毎回上昇すると予想したが、実際には 7 回目 (最終回) が前回より低くなった。

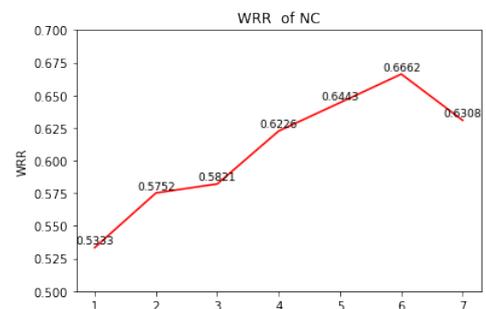


図 7 字幕無しグループのシャドーイング遂行成績

Fig. 7 the WRR of Group1

部分字幕グループ平均 WRR を図 8 に示す。1 回目から 7 回目まで、WRR 値は全体的に上昇傾向を示した。字幕を見る回の WRR は前回の字幕無し回より高いと予想したが、2 回目・3 回目の字幕ありでの WRR は直前の字幕な

しの WRR とあまり違いがなかった。

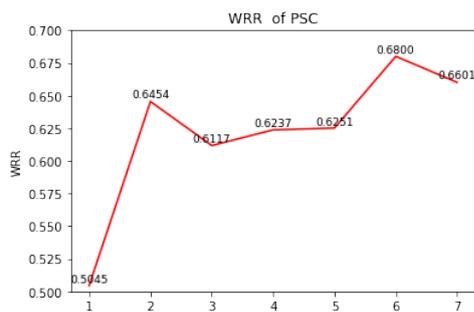


図 8 字幕無しグループのシャドーイング遂行成績
Fig. 8 the WRR of Group2

字幕無しグループと部分字幕グループで、最初の練習結果からの改善分を図 9 に示す。つまり、第 7 回の差は、事前と事後のトレーニングレベルの差である。今回の結果によると、字幕無しグループより、部分字幕グループのほうが最終的に正しく復唱できた単語が多いことが示された。

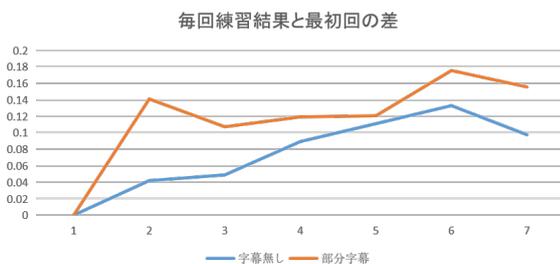


図 9 字幕無しグループと部分字幕グループの結果比較
Fig. 9 the Comparison between Group1 and Group2

本実験においては、実験協力者の人数が少なかったことに加え、完全な字幕を見てシャドーイングした場合との比較がない。また、長期練習を行わなかったため、学習者の日本語能力を向上させることができたかどうかは、現在もなお不明である。

したがって、部分字幕を用いたシャドーイング練習法と、完全な字幕や字幕無しの練習法とで、日本語学習者の能力にどのような変化が生じるかを探求することは今後の課題としたい。また、より多くの被験者による実験が必要である。さらに、今回用いた教材は新聞記事であったため、多くの被験者から練習資料が難し過ぎるという意見が寄せられた。より教材に適した、他のシャドーイング練習資料を探さなければならない。

5. おわりに

本研究では、第二言語学習者のシャドーイング練習を補助するために、部分字幕を用いたシャドーイング練習システムを提案した。リスニング練習における部分字幕の利点とシャドーイング練習でテキストを与える利点を考慮し、

部分字幕をシャドーイング練習に導入した。提案した部分字幕を用いたシャドーイング練習システムを構築して、システムの有効性を検討するために、小規模実験を行って、シャドーイング遂行成績を比較した。今回の結果によって、部分字幕システムを使用するグループは、字幕無しグループより、正しく復唱できる単語が多いことがわかった。しかし、今回の実験は、被験者が少ない、完全な字幕との比較がない、長期練習を行わないなどの不足があるため、今後はより多くの被験者を募集して、部分字幕を用いたシャドーイング練習法が、完全な字幕や字幕無しと比較して、日本語学習者の能力にどのような変化が生じるかを探求したい。

参考文献

- [1] 土居美有紀：初級授業にシャドーイングを取り入れる—先行研究から学ぶ—, 南山大学国際教育センター紀要, Vol.12, pp.77-91 (2011).
- [2] Zhang, Xinlei and Miyaki, Takashi and Rekimoto, Jun: WithYou: Automated Adaptive Speech Tutoring With Context-Dependent Speech Recognition, *Proc. 2020 CHI Conference on Human Factors in Computing Systems*, pp.1-12 (2020).
- [3] Elfz Media: *Shadowing - English Speaking Exercise*, Biltexsoftware (2019).
- [4] Miri Kim: *English Shadowing: TEDICT*, Apple App Store (2020).
- [5] Picup Inc.: *Shadowing-English Speaking Exercise*, Apple App Store (2017).
- [6] Hamada, Yo: Shadowing: What is It? How to Use It. Where Will It Go?, *RELC Journal*, Vol.50, pp.386-393 (2018).
- [7] Mirzaei, Maryam Sadat and Meshgi, Kouros and Akita, Yuya and Kawahara, Tatsuya: Partial and synchronized captioning: A new tool to assist learners in developing second language listening skill, *ReCALL*, Vol.29, No.2, pp.178-199 (2017).
- [8] Vandergrift, Larry: 1. Listening to learn or learning to listen?, *Annual Review of Applied Linguistics*, Vol.24, pp.3-25 (2004).
- [9] 岩下真澄：日本語文の視覚呈示がシャドーイングの遂行成績に及ぼす影響, 広島大学大学院教育学研究科紀要 第二部 文化教育開発関連領域, No.59, pp.219-227 (2010).
- [10] 門田修平：シャドーイングと音読の科学, コスモピア (2007).
- [11] 瀧澤正己：通訳訓練法の英語学習への応用 (1)—シャドーイング—, 北陸大学紀要, Vol.22, pp.217-232 (1998).
- [12] 玉井健, 門田修平：決定版 英語シャドーイング, コスモピア (2004).
- [13] 瀧澤正己：通訳訓練法から見直す聴解練習, 東呉日本語教育, Vol.16, pp.153-171 (1993).