

表情に基づく再生速度自動調節機能を備えた外国語学習支援の提案

西田 直人* 野崎 陽奈子† 志築 文太郎‡

概要. 外国語を学習する際、映画、ドラマ、またはアニメなどの動画コンテンツを外国語音声のまま視聴することは、リスニング能力を高めることに有効である。しかし、多くの外国語学習者は、母語話者と同じように聞き取りを行えないために、動画の再生速度についていけない場面に多く直面する。そこで我々は、学習者個人の理解度に応じて動画の再生速度を自動調節することにより動画の聞き取りを支援するシステムを検討している。本システムにおいて、学習者の理解度は、学習者個人の表情、および、動画の内容から判断される。今回、学習者の表情、および、動画の笑いどころに基づき、動画コンテンツの再生速度を動的に自動調節するシステムを試作した。



図 1. システムの概略図. a) 動画の再生速度を遅くする手法. b) 動画の再生途中に一時停止時間を設ける手法.

1 はじめに

外国語の学習には、音楽、本、および動画など様々なコンテンツが用いられる。その中でも、映画、ドラマ、またはアニメなどの動画コンテンツを外国語音声のまま視聴することは、リスニング能力を高めることに有効である [2]。特に、母語話者向けの動画コンテンツは外国語学習者向けの動画コンテンツと比較して数多く存在しており、それらは娯楽性が高

く、かつ実践的な言語の言い回しが用いられているため、人気がある [6]。

しかし、多くの外国語学習者は、母語話者と同じように聞き取りを行えないために、動画の再生速度についていけない場面に多く直面する [3]。このため、母語話者向けの動画コンテンツを通常の再生速度のまま教材として扱うことは難しい。

そこで我々は、学習者個人の理解度に応じて再生速度を自動で調節することにより、学習に用いる動画コンテンツの難易度を再生速度の面から調節するシステムを提案する (図 1)。今回、予備実験の結果から、学習者個人の理解度を表情認識を用いて推定することとした。

2 関連研究

教材の再生速度について、教育学の分野にて研究が行われてきた。Blau [3] は、リスニング教材の再生速度を遅くする、または一時停止時間を設けることにより、学習者の内容把握に関するテストの点数が向上することを示した。したがって本研究においても、再生速度を遅くする、または一時停止時間を設ける機能を作る。

コンテンツに対する理解度の調査について、Fujii および Rekimoto [5] は、外国語の動画コンテンツを視聴している際の学習者の視線を用い、彼らの動画コンテンツに対する理解度を推測した。この研究においては、学習者の視線が字幕の方へ向いているほど、学習者の外国語の能力、およびコンテンツに対する理解度が低い傾向があるということが示されている。さらに Fujii および Rekimoto は、学習者の理解度に合わせて字幕中の単語の対訳を行う補助を行い、劇中に出現する単語についての語彙テストの結果が、動画の視聴前後の比較にて向上したことを示した。

我々の研究では、学習者が動画コンテンツを理解できている箇所および理解できていない箇所の比較

Copyright is held by the author(s). This paper is non-refereed and non-archival. Hence it may later appear in any journals, conferences, symposia, etc.

* 筑波大学情報メディア創成学類

† 筑波大学情報理工学位プログラム

‡ 筑波大学システム情報系

において、字幕への視線だけでなく、表情の変化または貧乏ゆすりの有無など、身体的特徴の違いを実験を通じて調査した。さらに、この結果をもとに表情に基づく再生速度調節機能を試作した。

3 予備実験

動画コンテンツの視聴時、理解ができている箇所およびできていない箇所において学習者の身体的特徴に変化が現れるか否かを調べるために予備実験を行った。筆者と同じ研究室の学生5名（M=23.6歳，SD=1.82，全員男性）を参加者とした。参加者はまず、ウェブカメラにて顔のフレーム画像を撮られながら5本の母語話者向け英語動画コンテンツを英語音声および英語字幕の条件にて視聴した。英語コンテンツは、英語学習書に掲載されているTED Talks¹，TED Education²，および洋画のクリップ³から選んだ。参加者の理解度を大まかに知るために、動画コンテンツ中に、動画の内容に関する4択問題を5題提示されるシーンを設けた。その後、もう一度動画コンテンツを視聴しつつ、フレーム画像に対し、理解ができている箇所および理解できていなかった箇所のラベリングを行って貰った。また参加者には実験後にアンケートに回答して貰った。

実験後、実験者が参加者の顔のフレーム画像を目視で確認した限りでは、参加者の表情や視線に顕著な違いは見られなかった。内容把握問題の正答率については、0問正解が1人、1問正解が1人、2問正解が2人、4問正解が1人であった。この内容把握問題の成績から、実験参加者は動画コンテンツの内容を部分的にしか理解していなかったと考えられる。その一方で、新たに分かったこととして、予備実験の中において全員が無表情を保ち、動画コンテンツの内容として笑いどころの箇所においても笑っていないかった。予備実験後のアンケートにおいて、無表情になる理由としては、集中しようとして顔がこわばる、理解ができないから話の笑いどころが分からない、および、分からない箇所があまりに多いと表情を変えている余裕がないという意見があった。以上の結果から、外国語の動画コンテンツへの理解ができている場合、参加者は動画コンテンツを十分に楽しめず、無表情を続ける傾向にあると分かった。

したがって本研究では、動画を視聴中の学習者の表情に着目し、話の笑いどころが含まれる箇所にて笑わなかった場合、話の理解ができているものと見なし、動画の再生速度を遅くする、あるいは一時停止時間を設ける支援を行うことにした。

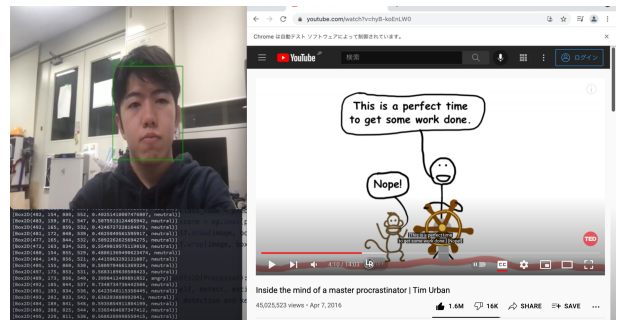


図 2. システムを使っている様子。顔面部分の緑のフレームは無表情を示している。

4 実装

我々は、ユーザの表情を検出し動画の再生速度を調整するシステムを実装した。

実装言語に Python を用いた。参加者の表情を検出するためにウェブカメラにて取得したフレーム画像および Perception for Autonomous Systems [1]，ブラウザ操作のために Selenium [4] を用いた。

また、予め動画中に笑い声が聞こえる箇所を Laughter-Detection [7] にて抽出し、その箇所を笑いどころと定義した。

学習者が動画コンテンツの笑いどころにて笑っていない場合に、再生速度を調節する2種類の機能を我々は実装した（図2）。1つ目は、再生速度を遅く調節する機能である（図1a）。再生速度は先行研究 [9] にならい、1倍速、0.8倍速、0.6倍速、および0.4倍速の4種類である。2つ目は、動画の再生途中に一時停止時間を設けることにより、学習者が理解しやすいように動画再生を調節する機能である（図1b）。一時停止時間については、0.5秒程度の一時停止時間を設けることにより英文の復唱の再現性が向上したという先行研究 [8] に基づき、0.5秒にした。これらの実装から、調整後の動画コンテンツの笑いどころにて学習者が笑った場合、学習者が理解可能な速度に動画が自動調整されたと解釈する。一時停止時間の挿入間隔については、人間の短期記憶がおおよそ3秒間持続するため [10]，3秒に設定した。なお、再生速度の変更および一時停止時間の挿入という2種類の機能は、理解がされていないと判断された時点から次の笑いどころまで持続する。

5 おわりに

本稿では、外国語の動画コンテンツを用いて外国語学習を行う際に、学習者が動画コンテンツの再生速度についていけない場合に備え、学習者の理解度に合わせて再生速度自動調節機能を備えたシステムを提案した。学習者の理解度については、予備実験の結果に基づき、表情および笑いに関するコンテンツの文脈から判断することにした。

¹ <https://www.youtube.com/watch?v=arj7oStGLkU>,
<https://www.youtube.com/watch?v=eIho2S0ZahI>

² <https://www.youtube.com/watch?v=eM2VWspRpfk>

³ <https://www.youtube.com/watch?v=Rnww09Zol6w>,
<https://www.youtube.com/watch?v=LZ0v9A5iEG4>

参考文献

- [1] O. Arriaga. Perception for Autonomous Systems, 2021. <https://github.com/oarriaga/paz> (最終閲覧日: 2021 年 11 月 19 日) .
- [2] B. Bal-Gezegin. An Investigation of Using Video vs. Audio for Teaching Vocabulary. *Procedia - Social and Behavioral Sciences*, 143(3):450–457, 2014.
- [3] E. K. Blau. The Effect of Syntax, Speed, and Pauses on Listening Comprehension. *TESOL Quarterly*, 24(4):746–753, 1990.
- [4] S. F. Conservancy. Selenium, 2021. <https://www.selenium.dev/ja/> (最終閲覧日: 2021 年 11 月 19 日) .
- [5] K. Fujii and J. Rekimoto. SubMe: An Interactive Subtitle System with English Skill Estimation Using Eye Tracking. In *Proceedings of the 10th Augmented Human International Conference 2019, AH '19*, pp. 1–9, New York, NY, USA, 2019. Association for Computing Machinery.
- [6] N. Fukunaga. “Those Anime Students”: Foreign Language Literacy Development Through Japanese Popular Culture. *Journal of Adolescent and Adult Literacy*, 50(3):206–222, 2006.
- [7] J. Gillick. Laughter-Detection, 2021. <https://github.com/jrgillick/laughter-detection> (最終閲覧日: 2021 年 11 月 19 日) .
- [8] X. Zhang, T. Miyaki, and J. Rekimoto. With-You: Automated Adaptive Speech Tutoring With Context-Dependent Speech Recognition. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, CHI '20*, pp. 1–12, New York, NY, USA, 2020. Association for Computing Machinery.
- [9] 小屋 多恵子. 初級英語学習者の聴解に与える発話速度調整の効果. 小金井論集, 13:11–30, 2017.
- [10] 高倉 祐樹, 大槻 美佳, 中川 賀嗣, 大澤 朋史, 谷川 緑野. 言語性短期記憶のメカニズムとその障害について. 高次脳機能研究, 31(4):411–421, 2011.

未来ビジョン

我々はこの研究を第一歩として、外国語学習の効率化および外国語学習に対する心理的負荷の軽減を目指す。

現状、翻訳機の性能が正確さおよび速さの面から不完全であることから、他国語を話す人々と不自由なく話し、正しい情報をより多く迅速に得るために必要最低限の語学力は必須である。現在においても、非英語圏の諸国において、多くの人間が義務教育の過程から英語学習を始め、多くの時間を割いてきている。さらに、グローバル化の流れにより、中国語またはスペイン語といった巨大市場における公用語を理解できることもビジネスの面から有用性が高くなっている。

しかし、これらの言語を自在に操るためには途方もない時間を学習時間に充てる必要がある。したがって、学習に際し心理的負荷がかからず、継続して行える学習方法にて学習する

ことが望ましい。この点において、娯楽性のある動画コンテンツの視聴であれば過剰な心理的負荷をかけず、無理なく行うことができ、その上口語的表現や背景文化など、より実践的な外国語知識をつけることができるため、余暇時間に行う外国語学習の方法としては最適である。

しかし、現時点においては動画コンテンツを用いた学習は外国語学習の一手段としてあまり有名でなく、一部の上級学習者向けの学習方法となっている。

我々はこの研究において、動画コンテンツ視聴を通じた外国語学習が知名度および難易度の側面から、より一般的になるようなアプリケーションを作りたいと考え、今回のような実装を行った。

この研究を通じ、動画コンテンツが全ての学習者層に対し、一般的な外国語学習の手段となることを願うばかりである。