

アニメの部分会話の難易度の推定による日本語のリスニング練習支援

Junjie Shan¹⁾(非会員) 西原陽子²⁾(非会員) 山西良典³⁾(正会員) 前田亮²⁾(非会員)

1) 立命館大学大学院情報理工学研究科 2) 立命館大学情報理工学部 3) 関西大学総合情報学部

Japanese Language Listening Practice Support Using Dialogue Scenes of Anime with Estimated Levels

Junjie Shan¹⁾ Yoko Nishihara²⁾ Ryosuke Yamanishi³⁾ Akira Maeda²⁾

1) Graduate School of Information Science and Engineering, Ritsumeikan University

2) College of Information Science and Engineering, Ritsumeikan University

3) Faculty of Informatics, Kansai University

{gr0260hr @ ed, nisihara @ fc} .ritsumeai.ac.jp

概要

日本のアニメはジャパニメーションと呼ばれ、その芸術性が高く評価されている。アニメを通じて日本に興味を持ち、日本語の学習やリスニング練習を始める人も多い。我々はリスニング練習の教材としてアニメが利用できる可能性に着目した。1つのアニメのエピソードの中には1つの会話だけでなく、部分会話が多数含まれている。各部分会話に含まれる単語や表現が異なることから、部分会話の難易度も異なると考えられる。本研究ではアニメの会話のスク립トが与えられると、部分会話へ分割し、それぞれの難易度を自動的に推定する手法を提案する。練習者に適した難易度の部分会話を提供することにより、リスニング練習の支援を図る。評価実験では被験者に事前テストの受験後に、部分会話を視聴することによるリスニング練習をさせ、その後事後テストを受験させた。実験の結果、日本語能力検定試験の難易度がL2のグループにおいて、単にアニメの部分会話を視聴する群ではテストの正解率の伸びは10.6%であったが、難易度が推定された部分会話を視聴する群では伸びが26.4%であり、提案手法の有用性を確認できた。

Abstract

Japanese animation called “Japanimation” is highly regarded for its artistry. Many people abroad are interested in Japan through *animes* and start learning Japanese and listening practice. We focused on the possibility of using *animes* as teaching materials for listening practice. An episode of *anime* contains not only one conversation but also several partial conversations. Since the words and expressions in each partial conversation are different, the degree of difficulty might also be different. This paper proposes a method to estimate the degree of difficulty of the partial conversation of *anime*. By providing partial conversations with estimated levels, we support learners in their listening practices. In evaluation experiments, participants were asked to take a pre-test, practice their listening with the degree assigned partial conversations, and take a post-test. The participants were divided into two groups. An experimental group used the degree set partial conversations while a control group used just partial conversations for their listening exercises. The experimental results showed that the increase in correct rate between pre-test and post-test was 26.4% in the experimental group of L2 of JLPT while that was 10.6% in the control group.

1 まえがき

アニメや漫画といったポップカルチャーは日本を代表する文化となりつつある。従来は子供用の娯楽としての側面が強かったが、近年その芸術性が高く評価されるようになり、国内国外において大人でも楽しむ人が増えてきている [1]。建造物の構築や映画製作など、アニメ以外の他の分野にも大きな影響を与えてきており、日本のアニメは特にジャパニメーションと呼ばれるようにもなった。海外の国際映画祭（ベルリン映画祭、ヴェネツィア映画祭、アカデミー賞など）にも出展、ノミネートされ、受賞もされており、その高い芸術性が国際的にも認められつつある。

日本のアニメや漫画は翻訳されて海外でも放映、発売されている。アニメや漫画を通じて日本語に興味を持ち、学習を始める人も多い [2]。2016 年には 360 万人以上の方が第二言語として日本語を学んでおり*¹、年々その数は増加している。学習者は自らの日本語のレベルを確認するために検定試験を受験する。英語では TOEIC や TOEFL などが検定試験としてあり、日本語では日本語能力検定試験*²がこれに相当する。日本語能力検定試験では受験者に対し言語知識、リーディング、リスニングの 3 種類の問題が出題される。学習教材は言語知識とリーディングの練習に焦点を当てたものが大半であり、リスニングの練習に焦点を当てたものは少ない。リスニング練習をするには音声が使われた教材が必要である。

我々はリスニング練習の教材としてアニメが利用できる可能性があると考えた。日本に留学する中国からの留学生はしばしば日本のアニメをリスニング練習に用いることが知られている [3]。アニメで話される音声はテレビドラマと比較すると発音が明瞭であり、ニュースと比較するとゆったりとした速度で話されている。日本のアニメは登場人物の声を当てる際に声を専門とする声優がそれを担当することが多い。アニメの音声は必ず「ただしく」「明瞭に」聞き取れることを前提としており、声優はそのための特殊な訓練を受けている [4]。特殊な訓練を受けた声優により音声 that 当てられた日本のアニメを用いることは、正しく、明瞭な音声を用いてリスニング

練習をしていくことができるという有用性があり、リスニングの教材としても有効と考えられる。日本のアニメは Netflix や Amazon Prime などの海外の動画配信サービスでも提供されており、アニメを日本語のリスニング練習に用いることは可能である。芸術作品であるアニメを言語の学習にも応用することができれば、芸術作品を通じた学習の機会を生み出すことにもつながると考えられる。

アニメは同一話者達による同一トピックについての会話だけではなく、話者のペアを変えた様々なトピックでの会話（部分会話）を含んでいる。トピックが変わると使用する単語や表現が変化するため、リスニングの練習者にとって部分会話の難易度はそれぞれ異なると考えられる。1 つのアニメのエピソードを視聴すると複数の難易度の部分会話 that 次々に流れてくることになる。練習者の練習のしやすさを考慮すると、各アニメのエピソードは部分会話に分割され、難易度も事前に推定されるべきと考えられる。現状ではアニメのエピソードを部分会話に自動的に分割し、難易度を推定した上で練習者に提供する手法やシステムは少ない。

そこで本研究では日本語のリスニング練習のために、アニメの部分会話の難易度を推定する手法を提案する*³。提案手法はアニメの会話のスク립トが与えられると、スク립トを部分会話へ分割し、各部分会話の開始時刻と終了時刻を取得する。その後、各部分会話に対し難易度を推定する。リスニングの練習者が練習したい難易度を指定すると、推定された難易度の部分会話の開始時刻と終了時刻が与えられる。練習者は難易度が推定された部分会話を視聴することで、リスニング練習が可能となる。

本論文では外国語の一つとして日本語、動画の一つとしてアニメをとりあげ、アニメを用い日本語のリスニング練習をする人を支援する。本研究で提案するシステムは他の言語や他の動画においても適用可能と考えている。

*¹ <https://www.jpfl.go.jp/j/about/press/2016/dl/2016-057-1.pdf>

*² <https://www.jlpt.jp/>

*³ 本研究ではアニメの字幕テキストデータを作成し、字幕テキストデータを分析するが、これらの行為は著作権法に触れないことを確認している（著作権法第 30 条の 4 および著作権法第 47 条の 5）

2 関連研究

第二言語の獲得にコンテキストが付与されたマルチメディアの教材を用いる例は多い [5, 6, 7]. マルチメディア教材の候補としては映画やドラマ, アニメなどが考えられる. 日本では世界的に見てもアニメや漫画といったポップカルチャーが普及している国である. アニメがストーリー性を有することから, コンテキストを意識した外国語練習に活かしやすく, 実際にアニメを用いた日本語の練習が行われている [8]. また, アニメや漫画を用い, 日本語の表現を練習するための Web サイトなども構築されている [9, 10].

アニメがもたらすコンテキストの中で練習をすると, 単語の認識やリスニング, 発音のリテラシーが発達することが確認されている [11]. アニメを視聴することによりリスニングの練習が効果的に行える可能性がある. アニメは1つ1つのキャラクターの発話により作られる部分的な会話が, 線形に接続され作られている. 異なる部分会話は異なる会話のトピックを持つと考えられる. トピックが異なれば使用される単語や表現が変わるため, 部分会話の難易度はそれぞれ異なると考えられる. 既存研究において教師が学習者のレベルに応じ, アニメを手動で選り分け提供したものがあ [12, 13]. 手動で行うには, アニメを視聴し, 部分会話へ分割し, それぞれの難易度を推定する必要があるが, これは時間と労力のかかる作業である. そこで本研究ではアニメを部分会話へと自動分割し, それぞれの難易度を自動推定し, 外国語の練習者に提供する手法を提案する. 練習者は自分のレベルにあった難易度の部分会話を用い, リスニング練習をしていくことが可能となる.

文書の難易度を推定する際に, 含まれる単語や表現を用いることが多い. ここで表現とは複数の単語からなる言い回し (例えば, 「…にあって」や「…かのようだ」など) を指す. ユーザの語彙力に応じ, 文書の難易度を評価する方法や [14, 15], 教科書コーパスを用い文書の読みやすさを評価する方法 [16] など, 様々な手法が提案されている. 本研究では部分会話の難易度を推定するために難易度の分類として, 日本語能力検定試験の分類を採用する. これは, 日本語を第二言語として学ぶ人の多くは, 自らの達成度を確認するために本試験を受験することが多く, 難易度の分類として一般的であることによる.

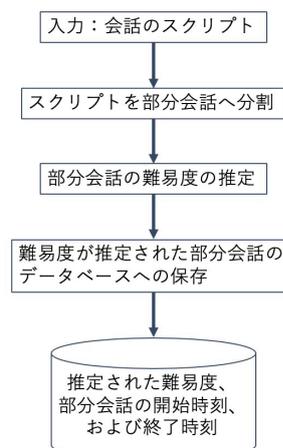


図1 提案手法の概要.

日本語能力検定試験には4つの難易度 (L1 から L4, L1 が最も難易度が高く, L4 が最も難易度が低い) があり, 各難易度の試験の受験に際し習得すべき単語と表現のリストが公開されている*4. リスニングテストは習得すべき単語と表現が考慮された上で作られていると考えられる. そこで本研究では過去のリスニングテストに含まれた単語と表現と, 部分会話に含まれる単語と表現を比較することにより, 部分会話の難易度の推定を行う. 試験の難易度に応じ習得すべき単語や表現のリストは他の言語でも存在する. 提案手法により他言語でのリスニング練習も可能である.

3 提案手法

図1に提案手法の概要を示す. 提案手法はアニメの会話のスク립トが与えられると, スクリプトを部分会話へ分割する. 続いて, 部分会話に含まれる単語と表現を用い, 部分会話の難易度を推定する. 難易度が推定された部分会話は開始時刻, 終了時刻の情報と共にデータベースに保存される. 日本語のリスニング練習をしたい人は, 推定された難易度を用いることにより, データベースから自らの練習に適した部分会話を選択し, 練習することができる.

*4 例えば, <https://www.ask-books.com/jp/hajimete-1j1pt/>がある.

表1 提案手法に与えられるアニメの会話のスキプトの例。各行は話者の発話の開始時刻，終了時刻，発話テキストにより構成される。アニメ「氷菓」の第2話より引用。

| 開始時刻 | 終了時刻 | 発話テキスト |
|---------|---------|------------------------|
| 0:09:38 | 0:09:40 | 普通の学校史ですね |
| 0:09:42 | 0:09:45 | これを読むのはかなり大変そうですね |
| 0:09:47 | 0:09:51 | これは毎週借りるやつがいてもおかしくないな。 |
| 0:09:51 | 0:09:54 | あんた ここで本借りたことないでしょ。 |
| 0:09:54 | 0:09:57 | うちの図書室の貸出期間は二週間なの。 |
| 0:09:58 | 0:10:00 | だから毎週借りる必要はないのよ。 |

3.1 アニメの会話のスキプト

提案手法にはアニメの会話のスキプトが与えられる。表1にアニメの会話のスキプトの例を示す。スキプトには3種類の情報があり，話者の発話の開始時刻，発話の終了時刻，発話のテキストである。

3.2 スキプトの部分会話への分割

難易度の推定を適切に行うために，スキプトを部分会話へ分割する。仮にあるトピックについての会話が継続しているならば，話者たちは話者交代の区間を大きくは取らず，連続して発話を出すと考えられる。したがって，大きな無音区間があるならば，それは部分会話終了し，新しい部分会話が始まることを示す可能性が高い。そこで，提案手法はスキプト内の発話の終了時刻と次の発話の開始時刻の時間の差を用い，スキプトの部分会話への分割を行う。分割に用いる無音区間の閾値は3.4節で行う予備的な調査により決定する。

3.3 部分会話の難易度の推定

部分会話の難易度を推定するために，各難易度で習得すべき単語と表現のリスト，および過去の日本語能力検定試験のリスニングテストを用いる。リスニングテストは習得すべき単語と表現を念頭におき作成されていると考えられる。したがって，過去のリスニングテストのスキプトを用いることで，部分会話の難易度を推定でき

表2 日本語能力検定試験の各難易度の習得すべき単語と表現の数

| | L1 | L2 | L3 | L4 |
|----|------|------|-----|-----|
| 単語 | 2962 | 3636 | 685 | 738 |
| 表現 | 105 | 177 | 108 | なし |

ると考えられる。スキプトの数が十分にあるならば，機械学習の手法を用い難易度を推定することができるが，公開されている過去のリスニングテストの数は十分であるとは言えず，機械学習を用いることは難しい。そこで過去のリスニングテストが含む各難易度の単語と表現の割合と，部分会話のその割合の類似度を評価し，類似度が最も高い難易度を部分会話の難易度とする。

3.3.1 部分会話からの単語と表現の抽出

はじめに部分会話のスキプトを形態素解析し，単語と表現を抽出する。提案手法では形態素解析器としてIgo^{*5}を用いる。単語と表現の難易度別のリストを用い，形態素解析により分かち書きされた中からマッチングにより単語と表現を抽出する。表2に各リストに含まれる単語と表現の数を示す。表3と表4に，リストに含まれる単語と表現の例を示す。

続いて抽出された単語と表現の難易度を，難易度別のリストを用い評価する。表2に示した習得すべき単語の数はL2, L1が多く，L3, L4は少ない。習得すべき表現の数は難易度別で大きな差はない。L4の表現のリストは公開されていない。これはL4では日本語のひらがなや基本的な単語の習得が目標とされており，表現の習得は含まれていないためと考えられる。提案手法では4つの単語リストと，3つの表現のリストの合計7つのリストを用いる。

部分会話のスキプトにはリストに存在しない単語が含まれている。話者の名前やアニメに特有の単語がこれに相当する。リストに含まれない単語は部分会話の中で占める割合は低く，それらの単語を飛ばしても発話の意味の理解に影響はないと考えられる。そこで，リストに存在しない単語は除外し，リストに存在する単語と表現のみを用い，部分会話の難易度の推定を行う。

3.3.2 部分会話とリスニングテストのベクトル表現

類似度評価のために部分会話をベクトルにより表現する。ベクトルは，部分会話が含んでいる各難易度の単語

*5 <https://igo.osdn.jp/>

表3 各難易度の習得すべき単語の例

| 難易度 | 単語の例 |
|-----|-----------------------------------|
| L1 | 愛想, 遺跡, 団扇, 英雄, 疎か, 介護, 禁じる, 空腹 |
| L2 | 愛情, 医師, 失う, 液体, 応用, 片道, 記憶, 苦情 |
| L3 | 挨拶, 意見, 美しい, 選ぶ, 贈り物, 会場, 季節, 比べる |
| L4 | 青い, 行く, 後ろ, 映画館, 同じ, 学校, 可愛い, 果物 |

表4 各難易度の習得すべき表現の例

| 難易度 | 表現の例 |
|-----|--------------------------------|
| L1 | ... にあって, ... ながらも, ... かたわら |
| L2 | ... かのようだ, ... 恐れがある, ... にしても |
| L3 | ... になる, ... だろう, ... ている |
| L4 | (なし) |

と表現の割合を要素とする。

部分会話を i とする。各難易度 j の単語の割合を $w_{i,j}$ とする。同様に、各難易度 j の表現の割合を $e_{i,j}$ とする。 j は L1 から L4 までとする。部分会話はベクトル d_i により表現される。

$$d_i = (w_{i,L1}, w_{i,L2}, w_{i,L3}, w_{i,L4}), (e_{i,L1}, e_{i,L2}, e_{i,L3}) \quad (1)$$

ベクトル d_i は7つの要素を持ち、4つが難易度別の単語の割合に関するもの、3つが難易度別の表現の割合に関するものである。

続いて、過去のリスニングテストを同様のベクトルで表現する。このベクトルを基準ベクトルと呼ぶ。基準ベクトルの作成には、2003年から2007年までの公開されているリスニングテストのスク립トを用いる。

基準ベクトルを s_m とする。 m は L1 から L4 までとする。基準ベクトルは以下の式 (2) で表す。

$$s_m = (w_{m,L1}, w_{m,L2}, w_{m,L3}, w_{m,L4}), (e_{m,L1}, e_{m,L2}, e_{m,L3}) \quad (2)$$

基準ベクトルは4本ある。表5に提案手法で用いる基準ベクトルの詳細を示す。リスニングテストの難易度がL3やL4では、含まれる単語の多くはL4のリストに含まれるものであり、含まれる表現の多くはL3のリストに含まれるものとなっている。難易度が上がり、L1やL2になると、含まれる単語の割合がL2で増加し、含まれる表現の割合もL2で増加する。

表5 難易度ごとの基準ベクトル

| 難易度 | $w_{m,L1}$ | $w_{m,L2}$ | $w_{m,L3}$ | $w_{m,L4}$ | $e_{m,L1}$ | $e_{m,L2}$ | $e_{m,L3}$ |
|-----|------------|------------|------------|------------|------------|------------|------------|
| L1 | 4.11% | 16.61% | 14.83% | 64.45% | 5.14% | 19.25% | 75.61% |
| L2 | 0.58% | 12.62% | 13.94% | 72.86% | 4.26% | 18.08% | 77.66% |
| L3 | 0.16% | 1.9% | 13.39% | 84.55% | 4.86% | 12.35% | 82.78% |
| L4 | 0.0% | 0.68% | 1.47% | 97.86% | 3.49% | 10.93% | 85.58% |

3.3.3 難易度推定のためのベクトルの類似度計算

部分会話のベクトル d_i と、4本の基準ベクトルのそれぞれとの類似度を計算し、最も類似度が高い基準ベクトルの難易度を部分会話の難易度として推定する。部分会話のベクトルの d_i は単語の割合のベクトル d_{wi} と、表現の割合のベクトル d_{ei} から構成されるとする。基準ベクトルの s_m も同様に二つのベクトル s_{wm} と s_{em} から構成されるとする。二つのベクトル d_i と s_m の類似度をコサイン類似度の重み付き和として、式 (3) により算出する。

$$sim(d_i, s_m) = \alpha \times \cos(d_{wi}, s_{wm}) + (1 - \alpha) \times \cos(d_{ei}, s_{em}) \quad (3)$$

類似度の範囲は0から+1の範囲になる。これはベクトル内の全ての値が負ではなく、正となるためである。単語の割合についてのコサイン類似度と、表現の割合についてのコサイン類似度は重みがつけられた上で加算される。重み α の値は3.5節の予備調査により決定する。提案手法では全ての部分会話にL1からL4までのいずれかの難易度が付与される。

3.4 予備調査1: 無音区間の閾値

部分会話へと分割する際の無音区間の閾値を決定する予備調査を行った。調査は以下の手順で行った。

1. 複数の閾値を設け、スク립トを部分会話へと分割する。
2. 各閾値での部分会話の平均時間を算出する。
3. 適切な閾値を決定する。

閾値の範囲は5.0秒から9.0秒とし、0.5秒刻みで閾値を設けた。合計9種類の閾値に対し検証した。

日本語能力検定試験でのリスニングテストの各問題の音声の平均時間は概ね120秒である。そこで実験手順の2.で算出する平均時間の中で120秒に最も近くなるものを探し、実験手順の3.で適切な閾値として採用した。

本実験ではアニメ「氷菓」を用いた。このアニメは主に高校生4人が学校生活に隠された謎を推理する物語を描いており、登場人物による謎解きや考察をするシーンが多数含まれている。このアニメは一般的な日本のアニメ作品の一つとして選定した。エピソード数は23、エピソードの合計時間は563.5分(=23エピソード×24.5分)であった。キャラクターによる発話数の合計は9,186回であった。

調査の結果、閾値が6.5秒のときに部分会話の平均継続時間がおおよそ120秒程度となった。この結果から、提案システムでは閾値として6.5秒を採用することを決定した。得られた部分会話は合計347個あったが、この中には一言のセリフのみを含むものが63個あった。一言のセリフしかない部分会話はリスニング練習には適さないと考えられるため63個の部分会話を削除した。3.5節の予備調査2と4章の評価実験では、合計284個の部分会話をを用いる*6。

3.5 予備調査2: 式(3)の重み α

難易度を決定する式の重みを決定する予備調査を行った。調査は以下の手順で行った。

1. 提案手法を用い、部分会話の難易度を推定する。
2. 難易度が推定された部分会話を被験者に提示する。
3. 被験者は推定された難易度が適切かどうかを判定する。

実験に用いた部分会話は3.4節で得られたものと同じで、合計284個あった。重みの値は0.1から0.9までとし、0.1刻みで変化させた。重みの値が0.1と0.2、0.3と0.4、0.6と0.7、0.8と0.9では、全く同じ難易度推定の結果が得られたため、5種類の重みの結果を用いた。

被験者は中国で日本語を教える2名の教師であった。1つの難易度に対し1つの部分会話を提示した。5種類の重みと、L1からL3までの3種類の難易度に対し、合計15件の異なる部分会話を提示した。部分会話はランダムに選定された。難易度L4の部分会話は評価に含めなかった。L4に分類された部分会話の数は他の難易度に比べて少なかった。L4は基礎的な内容であるため、アニメの部分会話の中にL4の難易度に相当するものが少ないと考えられる。また後述する評価実験においても、

*6 実験で用いた284個の部分会話は会話のコンテキストやまとまりの整合性がとれなくなったものはないことを確認した。

表6 重み α ごとの各難易度に分類された部分会話の数

| 重み α | L1 | L2 | L3 | L4 |
|-------------|-----|----|----|----|
| 0.1/0.2 | 180 | 52 | 39 | 13 |
| 0.3/0.4 | 177 | 64 | 31 | 12 |
| 0.5 | 164 | 85 | 24 | 11 |
| 0.6/0.7 | 162 | 90 | 23 | 9 |
| 0.8/0.9 | 160 | 91 | 24 | 9 |

表7 過去のリスニングテストの問題に対する難易度推定の正解率

| 重み α | L1 | L2 | L3 | L4 | 合計 |
|-------------|-----|-----|-----|-----|-------|
| 0.7 | 4/5 | 5/5 | 5/5 | 5/5 | 19/20 |
| 0.8 | 3/5 | 5/5 | 5/5 | 5/5 | 18/20 |

L4を受験予定とする被験者はいなかった。そこで本予備調査ではL1からL3の難易度と推定された部分会話をを用いて重みの適切値を調査した。

表6に提案手法により分類された各難易度の部分会話の数を示す。難易度が高いL1に分類された部分会話の数が最も多く、難易度が低いL4に分類された部分会話の数が最も少なくなった。

被験者からは重みが0.7または0.8のときに両方からの同意が得られた。重みが高いということは単語の割合を表現の割合よりも重視するという意味になる。各難易度で習得すべき単語数は表現数よりも多く、このことが難易度推定にも影響を与えていると考えられる。

難易度推定の妥当性を分析するために、過去の日本語能力検定試験のリスニングテストの問題に対し、重み0.7と0.8の場合で難易度の推定を行い、実際の問題の難易度との一致を調べた。L1からL4までのリスニングテストから、それぞれ5問ずつをランダムに抽出し、合計20問の難易度推定を行った。表7に難易度推定の結果を示す。重みが0.7の場合は20問中19問、重みが0.8の場合は20問中18問の難易度を正しく推定することができた。この結果からも重み0.7で推定された難易度は、実際のテストの難易度にある程度対応し、提案手法の難易度推定は妥当と示された。本手法では重み0.7を採用し、リスニング練習支援の評価実験に用いる。3.4節で得られた284個の部分会話の難易度を推定した結果、L1が162個、L2が90個、L3が23個、L4が9個となった。

表8 予備調査により得られた無音区間の閾値と重み α を用いた他の日本アニメの部分会話の難易度推定の結果

| アニメ | L1 | L2 | L3 | L4 |
|-------------|----|-----|----|----|
| けいおん! | 51 | 145 | 36 | 10 |
| ご注文はうさぎですか? | 59 | 78 | 19 | 8 |

3.4 節と 3.5 節の予備調査で得られた無音区間の閾値と重み α を用い、他のアニメについても部分会話への分割と難易度の推定を行った。アニメは「けいおん!」と「ご注文はうさぎですか?」の2つであり、それぞれのエピソード数は24個と12個であった。表8に各難易度の部分会話の数を示す。「氷菓」ではL1が最も多く次いでL2であったが、2つのアニメではL2が最も多く次いでL1となった。ただし、いずれのアニメでもL1とL2に分類される部分会話の数は、L3とL4に分類される部分会話の数よりも多くなり、類似した傾向が見られた。無音区間の閾値と重み α はアニメの種類によらず有効と考えられる。

4 評価実験

提案手法により難易度が付与された部分会話が、リスニング練習に与えた効果を評価する実験を行った。

4.1 実験手順

次の手順に従い実験を行った。

1. 実験者は被験者を募集する。被験者が次に受験予定の日本語能力検定試験の難易度 (L1 から L3 まで) の3つのグループに分類する*7。各グループをさらに半分にし、一つは実験群でもう一つは統制群とする。合計で6つのグループを得る。
2. 実験者は被験者に対し、日本語能力検定試験の模擬試験のリスニングテストを解答することを依頼する。リスニングテストの平均正解率をグループごとに算出する。この平均正解率を事前正解率と呼ぶ。
3. 実験者は被験者に対し、リスニング練習用の部分会話を提供する。部分会話の難易度は提案手法により推定されたものとする。被験者は部分会話を視聴

*7 現在の日本語能力検定試験は N1 から N5 までの分類であるが、N1 と L1, N2 と L2, L3 と N4, と対応づけられているため、N の分類を L に直し受験予定の難易度とした。
<http://www.jlpt.jp/sp/about/comparison.html>

表9 使用したリスニングテストの例 (難易度 L2)

| |
|--|
| 女の人二人が話しています。 二人はこの後、どうすると言っていますか。 女1: あっ、また前の日からごみ出している人がいる。 女2: 本当! ごみは当日の朝、出すことになっているのに。 女1: これじゃ、また猫や鳥が散らかしちゃう。 女2: 今度、管理人さんに言ってみようかしら。 女1: それより、注意を紙に書いて、貼っておくのはどう。 女2: そうね、それで、やめてくれるといいけど。 女1: 本当はよかったら直接注意したほうがいいけど。 女2: でも、いつ来るか分からないし。 女1: 本当に困るわね。まったく。 |
| 二人はこの後、どうすると言っていますか。 1. 猫や鳥を捕まえます。 2. 管理人さんに言います。 3. 注意を紙に書いて掲示します。 4. 前の日、ごみを捨てる人に直接注意します。 |

し、リスニング練習を行う。

4. リスニング練習後、実験者は被験者に対しリスニングテストを解答することを依頼する。2. で用いたテストとは別のテストである。2. と同様にリスニングテストの平均正解率をグループごとに算出する。この平均正解率を事後正解率と呼ぶ。
5. 実験者はグループごとに事後正解率と事前正解率の差を算出する。

4.2 使用したリスニングテスト

評価実験で用いたリスニングテストは中国で日本語学習を支援するサイトにある模擬試験のテストを用いた*8 表9に使用したリスニングテストの例を書き起こしたものを示す。リスニングテストの問題数は被験者1名につき5問であった。リスニングテストは解答を選択肢から1つ選ぶことにより行う。5つの問題のうち4つは図や表を参照しながら、日本語の会話や読み上げを聞き、解答をする。残りの1つの問題は図や表などの補助的な情報はなく、日本語の会話や読み上げのみを聞き、解答をする。前4つの問題と比較すると、最後の1問は補助的な情報がない分、被験者にとって難しい問題となっている。

4.3 被験者

被験者は日本語学習と議論が行われている中国の Web サイトを通じて募集した。Web サイトでは学習者が集まり、学習レベルや興味に応じてグループに分かれ、交流を行なっている。本実験では次に受験予定の日本語能

*8 <https://language.koolearn.com/jp/tingli/>

表 10 評価実験に最後まで参加した各グループの被験者数

| グループ | L1 | L2 | L3 | 合計 |
|------|----|----|----|----|
| 実験群 | 16 | 25 | 17 | 58 |
| 統制群 | 20 | 21 | 21 | 62 |

表 11 各グループに提供された部分会話の数

| グループ | L1 | L2 | L3 |
|------|---------------------------------------|-----------|-----------|
| 実験群 | L1 を 10 個 | L2 を 10 個 | L3 を 12 個 |
| 統制群 | L1 を 4 個, L2 を 4 個, L3 を 4 個, 合計 12 個 | | |

力検定試験の難易度を質問し、難易度ごとに被験者を集めた上で 2 つのグループに分けた。表 10 に各グループの被験者の数を示す。被験者を募集した際に L4 の難易度を受験予定の人はいなかった。L4 は基礎的な内容であるため、あえて試験を受け合格認定をとる必要がないと考える人が多いためと考えられる。そこで、L1 から L3 までの 3 つの難易度で評価を行った。実験に最後まで参加しない人もいたため、最終的に被験者の人数は各グループで 16 名から 25 名となった。

4.4 提供された部分会話

表 11 に被験者に提供された部分会話の数を示す。部分会話の数は各グループが 10 から 12 個となった。部分会話は実験者がランダムに選定した。数が異なっているのは、部分会話の視聴時間を均等にするためである。実験手順 3. で、被験者は割り当てられた部分会話を視聴する。実験群には次回受験予定の難易度と同じ難易度の部分会話を 10 個から 12 個提供する。統制群には L1, L2, L3 と推定された難易度の部分会話を 4 つずつ、合計 12 個提供する。実験群が視聴した部分会話と統制群が視聴した部分会話は重複があり、実験群が視聴した 10 個または 12 個の部分会話の中から 4 個を選択し、統制群に提供した。部分会話を提供する際、部分会話の動画を提供することはできないため、被験者にはアニメのタイトル（今回は「水菓」を伝えた）、エピソード番号、視聴開始時刻、視聴終了時刻の 4 つの情報を提供した。被験者は有料の動画共有サイト（bilibili など）を用いて部分会話を視聴した。

4.5 評価方法

算出された事前正解率と事後正解率をグループごとに比較する。仮に同じ難易度を受験予定で、実験群の方が統制群よりも事前正解率と事後正解率の差が大きくなれば、提案手法により難易度を付与した上で部分会話を提供することが有用と評価する。

4.6 実験結果

表 12 に実験群の実験結果を示す。表 13 に統制群の実験結果を示す。表 12 と表 13 は実験手順 2. と 4. で得られたリスニングテストの平均の正解率を示す。事後正解率と事前正解率の差は、実験群の L1 は 11.3%, L2 は 26.4%, L3 は 15.4% であった。統制群の正解率の差は L1 が 10.0%, L2 は 10.6%, L3 は 6.4% であった。L1, L2 と L3 グループに対し、有意水準を 0.05 とし平均の正解率の差の検定を行った。L1 については $t=0.18$, $p=0.43 > 0.05$ となり、有意差が見られなかった。L2 については $t=1.93$, $p=0.045 < 0.05$ となり、有意差が見られた。L3 については $t=1.23$, $p=0.12 > 0.05$ となり、有意差が見られなかった。L2 グループについては実験群の方が統制群よりも大きな差が得られた。このことから、L2 グループについては提案手法により難易度が付与された部分会話がリスニング教材として有効であり、リスニング練習を支援できることが示された。

5 考察

実験群と統制群の平均正解率の伸びについて考察する。実験群は L1 グループは 11.3%, L2 グループは 26.4%, L3 グループは 15.4% の伸びがあった。3 つのグループの伸びの平均値は 17.7% であった。これに対し統制群は、L1 グループは 10.0%, L2 グループは 10.6%, L3 グループは 6.4% の伸びがあった。3 つのグループの伸びの平均値は 9% であった。実験群と統制群のいずれのグループにおいても伸びが正の値となっていることから、事前正解率よりも事後正解率の方が点数が高くなったことになる。このことからいずれの難易度であっても、部分会話を視聴することにより、リスニングテストの練習ができると考えられる。そして、L2 においては実験群の方が統制群よりも事前テストから事後テストへかけての伸びが大きく有意差が得られたことから、受験予定の難易度と同じ部分会話を視聴することが、リスニングテストへより有効に働くと考えられる。

表 12 実験群のグループごとの事前テストと事後テストの平均正解率

| | | 問 1 から 4 までの正解率 | 問 5 の正解率 | 平均正解率 |
|----|---------|-----------------|--------------|--------------|
| L1 | 事前 | 59.4% | 43.8% | 56.3% |
| | 事後 | 70.3% | 56.3% | 67.6% |
| | 事後 – 事前 | 10.9% | 12.5% | 11.3% |
| L2 | 事前 | 50.0% | 20.0% | 44.0% |
| | 事後 | 79.0% | 36.0% | 70.4% |
| | 事後 – 事前 | 29.0% | 16.0% | 26.4% |
| L3 | 事前 | 50.0% | 5.9% | 41.2% |
| | 事後 | 66.2% | 17.6% | 56.6% |
| | 事後 – 事前 | 16.2% | 11.8% | 15.4% |

表 13 統制群のグループごとの事前テストと事後テストの平均正解率

| | | 問 1 から 4 までの正解率 | 問 5 の正解率 | 平均正解率 |
|----|---------|-----------------|---------------|--------------|
| L1 | 事前 | 48.8% | 30.0% | 45% |
| | 事後 | 58.8% | 40.0% | 55% |
| | 事後 – 事前 | 10.0% | 10.0% | 10.0% |
| L2 | 事前 | 48.8% | 23.8% | 43.8% |
| | 事後 | 61.9% | 23.8% | 54.4% |
| | 事後 – 事前 | 13.1% | 0.0% | 10.6% |
| L3 | 事前 | 48.8% | 28.6% | 45% |
| | 事後 | 60.7% | 14.3% | 51.4% |
| | 事後 – 事前 | 11.9% | -14.3% | 6.4% |

実験群の L2 グループは統制群と比較すると、もっとも大きな正解率の差が得られた (26.4% - 10.6% = 15.8%) (t=1.93, p=0.045 < 0.05)。L2 のリスニングテストにおいて、L1 と L2 の単語の割合は L3 と L4 に比較すると大きく増加する。表 5 に示す L1 と L2 の単語の割合は、L2 のリスニングテストにおいて 0.58% と 12.6% であるが、L3 のリスニングテストではそれぞれ 0.16% と 1.9% である。提案手法では基準ベクトルを用いてこの差を表現できている。このために L2 グループに適切な部分会話をより多く提供でき、結果として大きな正解率の差が得られたと考えられる。

実験群の L3 グループでは L2 グループについて大きな正解率の差が得られた (15.4% - 6.4% = 9%) が、有意差は見られなかった。L3 の難易度の部分会話は、会話の流れが類似していることが多かった。会話の流れとし

ては、挨拶をし合うものや、一人の登場人物が強調して話すために同じ単語を繰り返し使うものがあった。視聴する部分会話の数が多きとも、同じような流れのものであると練習効果が低く、このために実験群と統制群で正解率に有意差がなかったと考えられる。一つアニメだけを用いると L3 グループには練習に十分な部分会話を提供することが難しいと考えられる。異なるアニメの部分会話も使い、会話の流れが類似するものは除去することにより、L3 のリスニング練習に効果をもたらす可能性があることがわかった。

実験群の L1 グループと統制群の L1 グループのスコアの伸びの差は 1.3% (=11.3%-10.0%) で、大きな差はなかった。このことから L1 グループに対しては、いずれの難易度であっても同じ時間の部分会話を提供すると、同じ正解率の伸びの効果が得られるということにな

る。これは L1 を受験する際には習得すべき単語と表現の合計数が多くなるためと考えられる。習得すべき単語と表現の数が多くなると、それをういて作成可能なリスニング文章の種類数が増える。L1 を受験する際には、多くの種類の文が聞き取れる必要があるが、これを 12 回の部分会話では練習しきれなかったと考えられる。このため実験群と統制群で伸びがほぼ同じになったと考えられる。実験終了後に得られた L1 グループの実験群の被験者からのコメントには、より多くの数の部分会話を提供してくれると練習に役立つとあった。より多くの部分会話を提供することが L1 のリスニング練習には効果をもたらす可能性があることがわかった。

リスニングテストの問 1 から 4 までと、問 5 の正解率の伸びを比較する。L2 の実験群において、問 1 から 4 までの正解率の伸びは、問 5 の正解率の伸びより大きくなった ($t=3.81$, $p=0.015 < 0.05$)。問 1 から 4 までは図や表を見ながら音声を読み、解答をするテストである。図表の内容と音声の内容を対応づけながら、音声の内容を理解し、設問を解答していくため、対応づけの練習が重要となる。部分会話の視聴ではキャラクターたちの会話の様子が動画で表現され、その内容と対応づける音声が出力される。被験者たちは部分会話を視聴することで、動画と音声の内容の対応づけの練習が行うことができた。このために問 1 から 4 までの正解率の伸びが大きくなったと考えられる。

問 5 の正解率の伸びは、L2 の実験群において、問 1 から 4 までの正解率の伸びと比較すると、伸びが小さかった。問 5 は図や表などの視覚的な補助情報はなく、音声のみで内容を理解し、設問に解答をしていく問題である。部分会話を視聴することにより、動画と音声の内容の対応づけは練習しやすくなるが、音声のみで内容を理解することの練習は難しい。このため問 5 の正解率の伸びは小さくなったと考えられる。

6 あとがき

本論文では、日本語のリスニング練習を支援するために、アニメの部分会話の難易度を推定する手法を提案した。提案手法はアニメの会話の SCRIPT を部分会話へ分割し、部分会話に含まれる単語や表現の割合を用い、部分会話の難易度を推定する。これにより、日本語のリスニング練習をする人に適切な難易度の部分会話を提供

することができる。

評価実験では中国人の日本語学習者を被験者とし、実験群と統制群に分け、実験群には被験者が次に受験予定の難易度の部分会話を提供し、統制群には次に受験予定の難易度とそれ以外の難易度の部分会話を混ぜたものを提供した。各群は事前のリスニングテストを受験し、その後部分会話をういてリスニング練習をし、最後に事後のリスニングテストを受験した。L2 グループについて、実験群はリスニングテストの正解率が 26.4% 伸び、統制群は 10.6% 伸び、有意差が見られた。このことから単に部分会話を視聴するだけではなく、難易度を推定した上で部分会話を提供することが、リスニング練習の支援に効果があることが示された。

今後は、難易度が高い L1 テストの受験者にとって適切なリスニング練習の部分会話の提供方法について検討する。

参考文献

- [1] 中村伊知哉, 小野打恵. 日本のポップパワー 世界を変えるコンテンツの実像. 日本経済新聞出版, 2006.
- [2] 須藤健太郎. 中国「80 后」世代にとっての日本アニメ. 大阪産業大学経済論集, Vol. 10, No. 1, pp. 43–62, 2008.
- [3] Jin Cai. The influence of Japanese anime language to Chinese network buzzwords. In *MATEC Web of Conferences*, 2016.
- [4] 植野敏昭, 堀川りょう. アニメ声優の演技力のひみつ. 日本音響学会研究発表会講演論文集 (2019 年春季), 2019.
- [5] Ismail cakir. The use of video as an audio-visual material in foreign language teaching classroom. *Turkish Online Journal of Educational Technology*, Vol. 5, No. 4, pp. 67–72, 2006.
- [6] Yanqing Sun and Qi Dong. An experiment on supporting children's English vocabulary learning in multimedia context. *Computer Assisted Language Learning*, Vol. 17, No. 2, pp. 131–147, 2004.
- [7] Chih cheng Lin. Learning action verbs with animation. *The JALT CALL Journal*, Vol. 5, No. 3, pp. 23–40, 2009.
- [8] Hiroko Furo. Using anime as a teaching tool in

us undergraduate courses. In *WorldCALL 2008*, 2008.

- [9] 熊野七絵, 廣利正代. 「アニメ・マンガ」調査研究—地域事情と日本語教材. 国際交流基金日本語教育紀要, Vol. 4, pp. 55–69, 2008.
- [10] Kohei Matsuoka, Yoko Nishihara, and Ryosuke Yamanishi. Learning support system of role words by using classified lines of comic scene. In *International MultiConference of Engineers and Computer Scientists*, 2018.
- [11] Natsuki Fukunaga. “those anime students”: Foreign language literacy development through japanese popular culture. *Journal of Adolescent and Adult Literacy*, Vol. 50, No. 3, pp. 206–222, 2006.
- [12] Chan Yee Han and Wong Ngan Ling. The use of anime in teaching japanese as a foreign language. *Malaysian Online Journal of Educational Technology*, Vol. 5, No. 2, pp. 68–78, 2017.
- [13] Chan Yee Han, Wong Ngan Ling, and Ng Lee-Luan. Japanese language students’ perception of using anime as a teaching tool. *Indonesian Journal of Applied Linguistics*, Vol. 7, No. 1, pp. 93–104, 2017.
- [14] Yoshinori Miyazaki and Ken Norizuki. Developing a computerized readability estimation program with a web-searching function to match text difficulty with individual learners’ reading ability. In *WorldCALL 2008*, 2008.
- [15] Zhiwei Jiang, Gang Sun, Qing Gu, Tao Bai, and Daoxu Chen. A graph-based readability assessment method using word coupling. In *the 2015 Conference on Empirical Methods in Natural Language Processing*, pp. 411–420, 2015.
- [16] Satoshi Sato, Suguru Matsuyoshi, and Yohsuke Kondoh. Automatic assessment of japanese text readability based on a textbook corpus. In *the Language Resources and Evaluation Conference 2008*, 2008.

Junjie Shan



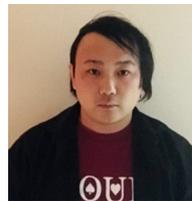
2012年南開大学計算機科学学部卒業。2018年立命館大学情報理工学研究科博士前期課程修了, 博士後期課程入学, 現在に至る。エデュテインメント, ヒューマンコンピュータインタラクション, ニューラルネットワークに興味を持つ。

西原 陽子



2003年大阪大学基礎工学部卒業。2005年同大学院基礎工学研究科博士前期課程修了。2007年同研究科博士後期課程修了。博士(工学)。日本学術振興会特別研究員(DC1, PD)を経て, 2008年東京大学大学院工学系研究科助教, 2009年同講師, 2012年立命館大学情報理工学部准教授, 2021年同教授, 現在に至る。立命館大学食総合研究センター兼務。ヒューマンコンピュータインタラクション研究, 自然言語処理研究に興味を持つ。情報処理学会, 人工知能学会, 各会員。

山西 良典



2007年名古屋工業大学工学部知能情報システム学科卒業。2009年同大学院工学研究科情報工学専攻博士前期課程修了。2012年同博士後期課程修了。博士(工学)。2012年立命館大学情報理工学部助手, 2013年同特任助教。2014年同助教, この間, UBC(カナダ)客員助教。2018年同講師。2020年関西大学総合情報学部准教授, 現在に至る。感性情報処理, Webインテリジェンス, マ

ルチメディア情報処理，音楽情報処理，言語情報処理に関する研究に従事．電子情報通信学会，情報処理学会，人工知能学会，日本感性工学会，日本音響学会，ACM，ACL 各会員．

前田 亮



1995 年図書館情報大学図書館情報学部卒業．1997 年同大学大学院図書館情報学研究科修士課程修了．2000 年奈良先端科学技術大学院大学情報科学研究科博士後期課程修了．博士（工学）．日本学術振興会特別研究員，科学技術振興事業団 CREST 研究員を経て，2002 年立命館大学理工学部助教授，2004 年同大学情報理工学部助教授，2007 年より同准教授，2011 年同教授，現在に至る．デジタル図書館，情報アクセス技術に関する研究に従事．ACM，電子情報通信学会，情報処理学会，日本データベース学会，情報知識学会，情報メディア学会，日本デジタル・ヒューマニティーズ学会，言語処理学会各会員．