

アニメを用いた日本語の聴解スキルの訓練システムに関する検討

A Study on Japanese Listening Training Support System by Using Anime

Junjie Shan*¹ 西原陽子*² 山西良典*²
 Junjie Shan Yoko Nishihara Ryosuke Yamanishi

*¹立命館大学大学院情報理工学研究科

Graduate School of Information Science and Engineering, Ritsumeikan University

*²立命館大学情報理工学部

College of Information Science and Engineering, Ritsumeikan University

This paper reports a study on Japanese listening training support system by using Anime. The proposed system divides episodes of Anime into scenes by using silent sections of Anime. The system evaluates a level of a scene by using words and expressions in subtitles of the scene. Users of the proposed system are assigned scenes of Anime which level is suitable for the users. We experimented with the proposed system to verify the efficiency of the system. Participants of the experiment watched scenes given by the proposed system. They took tests prepared from previous JLPT listening tests before and after watching scenes. The scores after watching scenes were higher than those before.

1. はじめに

第二外国語として日本語を学ぶ人の数は年々増えている[国際交流基金 15]。日本の古典的な文化やモダンカルチャが海外で紹介されたのをきっかけとして、日本語に興味を持つ外国人も少なからずいると考えられる。モダンカルチャの一種として日本語のアニメやコミックがある。いずれも海外では高い評価を受け、現地での言語に翻訳され楽しまれていることも多い。我々は、単語や文法などを示した教科書に加えて、生きた言語が含まれるアニメやコミックを用いた言語学習の研究を行っている[Shan 17]。

特にアニメは聞き取りやすい発音の声優がキャラクターの声をあてており、耳で聞いて内容を理解するスキルを訓練するには最適の材料と考えられる。アニメのキャラクターたちは年齢、性別が多岐に渡っており、会話の中にも様々な種類の単語が含まれている。したがって、日本語を母語としない人々にとっては、その人の日本語のレベルに応じてわかりやすいセリフとそうでないセリフが存在すると思われる。各人のレベルにあったセリフが含まれるアニメを選択してみるにより、聞くスキルの訓練が効率よく行えると考えられる。

そこで本論文では、アニメを用いた日本語の聴解スキルの訓練システムを提案し、その実験結果を報告する。アニメやコミックを用いた日本語教育においては、教科書の事例としてアニメのシーンやコミックのコマを利用するものが多いが[Lammers 04]、聴解のスキルの訓練にアニメを用いた例は我々の知る限りでは少なく、利用するための手法が求められている。

2. アニメを用いた日本語聴解スキルの訓練システム

システムの処理手順の概要を説明する。

1. アニメの動画データを無音区間を用いてシーンに分割する
2. 分割されたシーン内のセリフのテキストを取得する

連絡先: 西原陽子, 立命館大学情報理工学部, 滋賀県草津市野路東 1-1-1

表 1: 日本語の単語と表現のレベルごとの数

Level	L4	L3	L2	L1
Word	738	685	3636	2962
Expression	none	108	177	105

3. テキストを形態素解析にかけ、単語を取得する
4. 取得した単語のレベルを評価する
5. シーン全体の日本語のレベルを評価する

それぞれの手順について本論文での詳細を説明する。3. の形態素解析では Igo を用いた。4. の単語のレベルは、日本語能力検定試験の参考書で分類されていたものを用いた。日本語能力検定試験は英語における TOEIC や TOEFL のようなもので、日本語を母語としない人が日本語の学習の程度を評価するために受験する試験である。旧基準では L1 から L4 までの 4 つのレベルがあり、L1 がもっとも難しく、L4 がもっとも平易となっている。各レベルで主に出题される単語と表現が定まっており、表 1 に示す分布となっている。本論文では各単語のレベルを L1 から L4 のいずれかとして評価し、各表現のレベルは L1 から L3 のいずれかとして評価する。

5. のシーン全体の日本語レベルは、シーンに含まれる単語と表現のレベルを用いて評価する。はじめに評価のための基準となるデータを日本語能力検定試験の聴解問題から用意する。聴解問題は L1 から L4 までである。聴解問題で使われる音声データを書き起こし、含まれる単語と表現のレベルを取得する。つづいて、各レベルの単語と表現の割合を算出する。レベル L_i の聴解問題の単語と表現のレベルを、式 (1) と式 (2) に示す 2 つのベクトルで表現する。

$$Word(L_i) = \{wr(L1), wr(L2), wr(L3), wr(L4)\}, \quad (1)$$

$$Exp(L_i) = \{er(L1), er(L2), er(L3)\}. \quad (2)$$

表 2: 実験被験者のデータ

回答レベル	実験群	統制群
L1	16	20
L2	25	21
L3	17	21

シーン S_j の単語と表現のレベルも同様に、式 (1) と式 (2) を用いて 2 つのベクトルで表現する。各シーンのベクトル $Word(S_j)$ と $Exp(S_j)$ が、どのレベルの分布のベクトル $Word(L_i)$ と $Exp(L_i)$ と似ているかを評価するために、式 (3) を用いて、類似度を評価する。類似度がもっとも高いレベルを一つ選択し、シーンのレベルとして付与する。

$$Sim(L_i, S_j) = \alpha * Cos(Word(L_i), Word(S_j)) + (1 - \alpha) * Cos(Exp(L_i), Exp(S_j)). \quad (3)$$

式 (3) の $Cos()$ は 2 つのベクトルのコサイン類似度である。 α は重みであり、単語と表現のどちらを重視するかを決めるものである。

3. 評価実験と結果

提案したシステムを評価する実験を行なった。実験では 3 つのアニメから合計 71 本のエピソードを用意し、提案システムへの入力として与えた。実験者は被験者を募集し、実験群と統制群に分けた。被験者は中国に住む日本語を学習している主には大学生であった。被験者には現状の日本語のレベルを質問し、一つ上のレベルを目指すという設定でグループ分けをした。例えば、現状 L2 の人は L1 を目指して学習するという意味である。結果として被験者のグループ分けは表 2 に示す通りとなった。

はじめに被験者たちの現状の日本語レベルを評価するためにテストを課した。テストは過去の JLPT の試験の聴解問題から 5 問ずつ用意し、各レベルの被験者に回答させた。このスコアを視聴前スコアとする。実験群の被験者はグループ分けされたレベルのシーンを 10 件から 12 件視聴した。統制群の被験者はグループ分けされた 3 つのレベルから 4 件ずつ、合計 12 件のシーンを視聴した。

視聴後に被験者は訓練効果を評価するためのテストを受験した。テストは過去の JLPT の試験の聴解問題から 5 問ずつ用意し、各レベルの被験者に回答させた。視聴前に行った問題とは異なる問題を用意した。このスコアを視聴後スコアとする。

3.1 実験結果：視聴前スコアと視聴後スコアの比較

表 3 に実験群の視聴前スコアと視聴後スコアを示す。表 4 に統制群の視聴前スコアと視聴後スコアを示す。

視聴前の正解率と視聴後の正解率を比較すると、実験群の L1 から L3 のグループと、統制群の L1 から L3 のグループのいずれにおいても、正解率が上昇した。アニメを視聴することにより、聴解テストのスコアは上昇することが示された。

上昇率を比較すると、実験群の上昇率の方が統制群の上昇率よりも高くなっていった。特に L2 では、実験群の上昇率は 26.4% であるのに対し、統制群の上昇率は 10.6% であり、もっとも差が大きかった。

表 3: 実験群の視聴前スコアと視聴後スコア

		Q1	Q2	Q3	Q4	Q5	正解率
L1	Pre	10	9	9	10	7	56.4%
	After	12	11	12	10	9	67.6%
L2	Pre	14	10	17	9	5	44%
	After	18	22	19	20	9	70.4%
L3	Pre	5	12	8	9	1	41.2%
	After	11	13	11	10	3	56.6%

表 4: 統制群の視聴前スコアと視聴後スコア

		Q1	Q2	Q3	Q4	Q5	正解率
L1	Pre	12	9	10	8	6	45%
	After	11	12	16	8	8	55%
L2	Pre	9	9	11	12	5	43.8%
	After	16	8	14	14	5	54.4%
L3	Pre	8	9	14	10	6	45%
	After	11	10	18	12	3	51.4%

4. おわりに

本論文では、アニメを用いた日本語の聴解スキル訓練システムを提案し、その実験結果を報告した。提案システムではアニメのシーンごとに、日本語能力検定試験の聴解問題のレベルとの類似度を評価することにより、シーンのレベルを 3 段階で評価する。類似度の評価にあたっては、シーンに含まれる単語と表現のレベルごとの分布と、聴解問題に含まれる単語と表現のレベルの分布を用いた。提案システムを評価する実験を行なった。提案システムを用いてレベルを評価されたシーンを視聴した被験者は、実験前のテストの点数に比べて、実験後のテストの点数が高くなった。このことから提案システムの有効性を確認した。

謝辞

本研究の一部は科研費若手 B(17K13254) と基盤 C(16K00307) の補助を受けた。記して謝意を申し上げる。

参考文献

- [国際交流基金 15] 国際交流基金: 2015 年度「海外日本語教育機関調査」結果, <https://www.jpf.go.jp/j/about/press/2016/dl/2016-057-1.pdf> (2016)
- [Lammers 04] Wayne P. Lammers: Japanese The Manga Way: An Illustrated Guide To Grammar And Structure, Stone Bridge Pr (2004).
- [Shan 17] Junjie Shan, Yoko Nishihara, Ryosuke Yamaniishi, and Junichi Fukumoto: Analysis of Dialogues Difficulty in Anime Comparing with JLPT Listening Tests, 21st International Conference on Knowledge-Based and Intelligent Information & Engineering Systems, Vol. 112, pp. 1345-1352 (2017).