

ニューラル音声を用いた日本語教育用 eラーニング教材の提案

中村しづか^{*1}, 和崎克己^{*1}

^{*1} 信州大学大学院 総合医理工学研究科

Proposal of e-learning materials for Japanese language education using neural voice

Shizuka NAKAMURA^{*1}, Katsumi WASAKI^{*1}

^{*1} Graduate School of Medicine, Science and Technology, Shinshu University

本稿は、従来の日本語教育現場では軽視しがちな発音能力、および、真正な文脈を提示した、正しい発話への架け橋となる教授法を検討した。非漢字圏学習者にとって学習の大きな壁となる日本語表記体系（ひらがな、カタカナ、漢字）による学習ストレスを軽減し、音声を用いた eラーニング用教材開発を行い、内容言語統合型学習（CLIL）、TTS（Text to Speech）および、STT（Speech to Text）による学習プラットフォームである Moodle 上にコースウェアとして公開した。日本語学習における教材開発の経過報告、および、非ニューラル音声、ニューラル音声を用いた教材開発の状況を報告する。

キーワード: TTS (Text to Speech), STT (Speech to Text), eラーニング音声教材, ニューラル音声, 内容言語統合型学習 (CLIL), 日本語教育

1. はじめに

日本語教育現場では、発音能力の土台を固めることが重要であると認識しているはずだが、文法や語彙への学習に偏りがちであり、発音の学習は軽視されがちである。しかしながら、発音のよい学習者は、文法やそのほかの点を間違ってしまったとしても、発音が悪い学習者よりも、対象言語をうまく話すことができるとみられる傾向がある。発音が不十分であることは、正当な教育を受けていない⁽¹⁾、知識が不足している、といった誤解を招くことがある⁽²⁾。さらには、文法的には正確な会話をしているにもかかわらず理解してもらえないこともある⁽³⁾。また、語彙を学習する際、直訳による導入によって、学習者の既存知識を用いた解釈の結果、発話の誤用が見られる。これらの要因から、業務や高度コミュニケーション時に必要な、前提としている概要や内容、および、状況に応じた正しい発話レベルの到達が困難である。

本稿では、日本語学習入門、初級レベルの非漢字圏

学習者に対し、効果的な学習法を提案するために、内容言語統合型学習（CLIL）、および、音声学習を合わせた LMS（Learning Management System）教材を Moodle 上に公開し、教材の状況について報告するものである。

2. 日本語音声教材の既存研究および現状

2.1 プロソディグラフ

日本語音声教材の既存研究においては、音声合成・音声認識技術を活用し、入力音声のリズムを視覚的に補正した「プロソディグラフ」を自動出力するシステムの開発がある⁽⁴⁾。プロソディグラフとは、音声分析機を用いて高さを表すピッチ曲線を抽出し、それを音節単位で区切ってわかりやすく示したグラフである。これにより、アクセント、イントネーション、ポーズ、プロミネンス、母音の無性化、音節感覚、リズム、発話速度、文レベルの声の高低などを総合的かつ簡易に学習させることができる。調査対象者である 14 名の

コメントによると、肯定的な意見があげられた。しかしながら、この既存研究の対象者の多くが中国話者であり、投稿者が担当するクラスの対象者、英語話者である非漢字圏学習者への適用がない。また、非漢字圏学習者にとって日本語表記体系（ひらがな、カタカナ、漢字）の学習は大きな壁となっている⁽⁶⁾ことから、日本語表記体系を用いた視覚情報に全てたよった教材提供の困難さを考慮した上で、学習コンテンツの提供方法は慎重に考える必要がある。

2.2 日本語学習者

日本語学習者の学習目的を見ると、アニメやファッションなどの興味から学習をする学習者が多いものの、将来の仕事や就職、今の仕事で必要などといった趣味以外の場面での目的も多く、これらの学習者については、正確に話す必要があることが分かる。また、海外における日本語学習者数を見ると、過去39年間で30.2倍に増加している。教師数も増加傾向にあるものの、学習者に対して教師数は2%にとどまっている⁽⁶⁾。これらの海外における日本語学習者にとっても、eラーニング教材により、内容言語統合型学習（CLIL）の4つの概念である4C（3節にて後述。図1：内容 Content, 言語知識・言語使用 Communication, 思考 Cognition, 協学・異文化理解 Culture）を用いた、ある一定レベル以上の学習目的として、文化や高度な内容・概念を通じた学習が不可欠である。また、この4つの概念を用いた音声教材は現在存在しないため、新たなeラーニング教材として検討する必要がある。



図 1 CLIL の概念である 4C

2.3 教育現場の変化

2019年3月以降、筆者が担当する日本語講座の全てがオンラインへと移行した。2020年11月現在、今

もなお、オンラインでの授業が続いている。同期型授業では、録画することができ、また、非同期型授業では、自分のスケジュールで授業を受けることができ、欠席者や社会人学習者にとって、授業に遅れることなく自己学習できる点においてメリットが大きい。また、発話については、マスクを着用するよりもはっきり音声聞き取り、発話することができるため、オンラインでのメリットも大きい。With コロナ時代をきっかけに、教育DXが急速に一般化したため、教育提供はeラーニングを基軸とするようになった。そのため、日本語教育現場においてもオンライン授業が急激に増え、eラーニングベースの新たな教材開発が求められている。人間系の代わりに、機械系（Text to Speech, Speech to Text）を利用することによる受講者のスクレーパビリティの確保がより大切な時代の流れとなっている。

3. 教授法

3.1 内容言語統合型学習（CLIL）の適用

内容言語統合型学習（CLIL）とは、特定の内容（強化やテーマ、トピック）を、目標言語を通して学ぶことにより、内容と言語の両方を身につける教育法である⁽⁷⁾。学習者が学ぶべき内容について、動機を高め、情報収集、分析、発表、検討を行うなどの言語活動を行うことにより、自然な言語運用を通してより高度な知識力、思考力へと働きかけることができる。ヨーロッパを中心に英語教育法として急速に広まり、日本では、「New Horizon Elementary」（東京書籍）を使用した小学生の英語の教科書にも採択されている。食物連鎖、食糧産地、栄養など、社会や理科の教科と統合し、英語を学ぶ試みが教育現場で行われている。指導にCLILを用いるメリットは、動機付けが高まることや、意味のあるインプットが与えられること、アウトプットの必要性が生まれること、深い思考に伴い、言語知識が記憶に定着しやすいことなどを挙げている。また、CLILの概念である4Cを意識することで、内容と言語を統合した授業が行いやすくなり、授業の質も向上する⁽⁸⁾。日本語学習においては、CLILを用いた日本語教育の試みが報告されているが、いずれも対面授業、および、同期型授業における報告であり、eラーニング教材としてのCLILの適用はまだない。今後のeラー

ニング教材としての、日本語習得と実践力を兼ね備えた力を身に着ける教材開発に期待したい。

3.2 内容言語統合型学習 (CLIL) の実践

筆者が担当した日本語講座のうち、「Summer Session 2021 Intermediate Course 1」では、CLILを意識し、全8回の短期集中講座を開催した。CLILの概念である4Cを用いて、日本文化を知り(内容)、自国文化と比較し(思考)、高度な日本語を学び(言語)、発表(協学)を通じた授業を展開した(図2, 図3)。受講生6名は事前にMoodleのコースウェアにアクセスし、自己学習を進めたのち、同期型のオンラインクラスに参加した。事前に自己学習を進め、自由に調べ、準備をしておくことで、同期型授業での緊張を緩和し、自信をもって発表することができたとの意見があった。また、日本語の学習を通じて、日本文化に興味を持ち、深く理解することができたとの意見も多かった。語彙導入には音声教材を利用したが、繰り返し練習することで、同期型授業内でも容易に聞き取りが可能となり、従来授業内で行っていた語彙の導入を省くことも可能であった。また、受講生は、今後のコンテンツに期待したいという意見が多数あった。

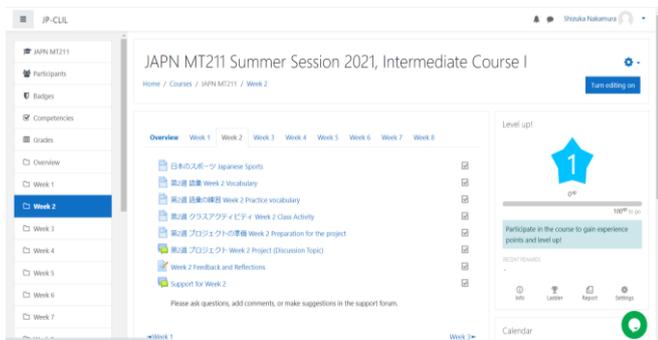


図 2 Summer Session 2021 コースウェア

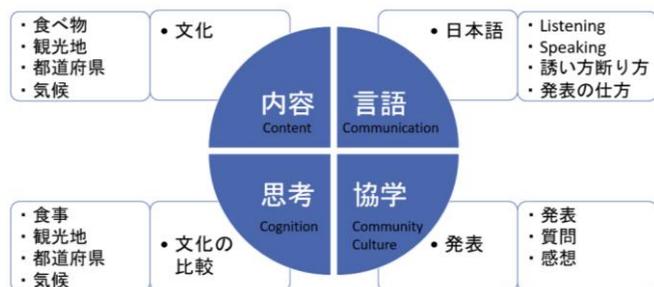


図 3 CLIL の概念である 4C の中級クラス適用例

4. e ラーニング教材開発

4.1 教材開発

教材には、GNU GPL (General Public License) の下で配布されているオープンソース LMS (Learning Management System) である Moodle 上にコースウェアとして展開した⁹⁾。Moodle には、ユーザ管理の機能をはじめ、多くの機能が備わっているため、学習者の学習状況をはじめとする管理を一元化し、試験の自動化など指導者側の作業においても自動化することができる。自動化することで、教師の負担を軽減し、効率化を図ることができる。

また、教材には語彙の導入、練習問題をはじめ、音声教材を導入している(表1)。非漢字圏学習者4名が受講する Basic II Course の宿題として、「Basic II Homework」での練習問題では、日本文化について英語で説明を聞いたのち、音声を使った練習を行う(図4)。Basic II では、漢字の導入は行っていないが、補足情報として日本語表記体系での説明も追加した。これにより、音声を聞き、自分の回答を録音し、音声にて回答を提出する。受講者は、インタラクティブなデザインにより、宿題が楽しくなった、日本文化に興味を持った、自分の音声を入力し、正解するのが嬉しかったなど、肯定的な意見が多かった。しかしながら、問題によってはなかなか正しく発音しているつもりでも次の問題に進むことができないといった不具合も見つかった。



図 4 Basic II Homework 練習問題

4.2 Text to Speech, Speech to Text の導入

Text to Speech (TTS) とは、テキスト(文章)を入力し、望む言語内容の音声を生成する機能である。また Speech to Text (STT) は、TTS とは逆に、音声か

らテキストを生成する機能である。本教材では、合成音声を使うことで、作業の効率化を図り、また多くの音声データ作成を可能にしている。英語の音声合成データは多く存在するが、日本語の音声合成は、現時点において数に限りがある。開発段階においては、非ニューラル音声、および、ニューラル音声の検証を行い、本教材に導入をしている。「Basic II Homework」では、TTS, STT を用いた練習問題を実施している (図 5)。



図 5 Basic II Homework 語彙練習問題

4.3 教材開発の計画

日本語学習 e ラーニング教材の開発ロードマップを図 6 に示す。開発中の e ラーニング教材には、音声教材を利用し、学習者にとって負担の大きい日本語表記体系（ひらがな、カタカナ、漢字）による視覚情報を主な教材として使わず、音声にフォーカスを当てている。開発の Step1 では、インターネット上で利用可能な TTS を用いた。音声ファイルを作成し、Moodle 上に、「Elementary Japanese Course I」として公開した (表 1)。次に、Step2 では、Moodle のプラグインである Poodll を導入し、「Elementary Japanese Course I」の教材をアップデートした。さらに、「Summer Session 2021 Intermediate Course I」、「Basic II Homework」他に適用した。Poodll は、オーディオとビデオの記録、教室用のメディアプレイヤー、タブやストップウォッチなどのウィジェットなど、Moodle の機能のツールボックスであり、音声学習に活用できる機能を多く兼ね備えている。主に英語学習に用いられている。音声合成には、Amazon Polly の非ニューラル音声を利用している。Polly の日本語音声には、Takumi, Mizuki の 2 種類の音声があり、いずれも非ニューラル音声のため、音声教材として利用するには、人間の音声とかけ離れた発音やトーンもあり、検討が必要である。

現在開発中の Step3 の教材では、音声教材としてニューラル音声を用いた。TTS を利用し、音声ファイル生成後、Moodle に音声ファイルとして埋め埋め込み、

評価を行った。また、TTS の合成音声には、Microsoft Azure を用いた。Azure での日本語音声では、2 種類のニューラル音声を提供されており、Nanami (Nerul), および、Keita (Neural) を用いている。しかしながら、汎用性の面から不十分であり、API (Application Programming Interface) を利用した教材開発が必要である。さらに、教師が指導の内容に合わせて、ニューラル音声のトーンやピッチ、スピードを変更できるような GUI (Graphical User Interface) に変更し、簡単に、かつ、自由に変更できるような仕様へと開発中である。Moodle 上で、真正な文脈を提示し、学習者の高速な学習を促進する。

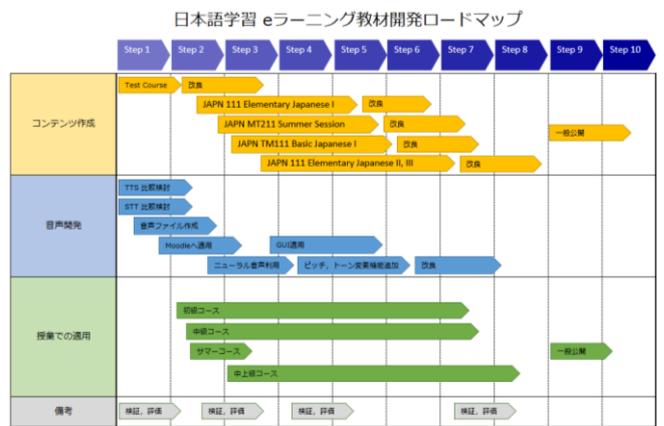


図 6 日本語学習 e ラーニング教材開発ロードマップ

表 1 教材開発過程と音声教材

開発過程	音声教材を用いたコース
Step1	簡易音声教材 - JAPN 111 Elementary Japanese I
Step2	Poodll による音声教材 - JAPN 111 Elementary Japanese I - JAPN MT211 Summer Session 2021 - JAPN TM111 Basic Japanese I - JAPN TM 112 Basic Japanese II - JAPN 111 Trial Lesson
Step3	ニューラル音声による教材 - Sample Voices
Step4	API を利用した開発
Step5	GUI を用いた音声選択

5. 試作中の日本語音声教材の評価

5.1 音声教材 (Step2) の評価 (非ニューラル音声)

開発段階 Step2 では、非漢字圏学習者 20 名に評価を行った。代表的な評価結果 (自由記述によるフィードバック) を表 2 に示す。音声教材を用いた e ラーニング教材を利用することで、何度も音声を聞きながら練習を繰り返すことができ、日本語表記体系 (ひらがな、カタカナ、漢字) の学習が完全に進んでいない状態でも、楽しく音声のみで練習し、理解することができた、CLIL の要素を用いていることで、日本語の文化を理解し、夢中で学習した、といった肯定的な意見が 80% あった。しかし、一方で音声面白い声だった、教師からのフィードバックをその場でほしいといった、否定的な意見が 20% あった。否定的な意見における音声への指摘は、非ニューラル音声の声のトーンやイントネーションが不自然である点があげられる。そのため、次の開発段階である Step3 ではニューラル音声を導入し、評価を行った。

表 2 開発教材 (Step2) における学習者評価

● 音声を聞いて練習を繰り返すことで、ひらがなを見なくても理解することができた
● 日本文化も学ぶことができ、夢中で学習した
● 単語を覚えるのが簡単だった。また、場面想定できた
● 自分の音声認識されないと、何度も練習し、認識されたときは嬉しく効果的な学習だった
● 早く日本に行って日本人と日本語で話し、自分の実力を試したくなった
● 自己学習できるのは効率がいい
● 先生からのフィードバックが毎回楽しみだ
● 音声面白い声だった
● 先生に質問したいときに質問できる対面型授業が望ましい
● なかなか音声認識してくれない

5.2 ニューラル音声を用いた TTS 品質の評価 (Step3 準備)

教材には、非ニューラル音声 6 パターン、および、ニューラル音声 6 パターンを用いたサンプル音声を合計 12 パターン用意し、非漢字圏学習者 6 名、日本語

教師 5 名による評価を行い、有意性を検証した⁽¹⁰⁾。非漢字圏学習者の日本語レベルは、CEFR (Common European Framework of Reference for Languages: Learning, teaching, assessment) の A1 レベルの学習者である。CEFR とは、2001 年にヨーロッパの言語教育・学習・評価の場で共有される枠組みとして欧州評議会より発表され、以後、世界でも広く外国語教育や評価の場で利用されている評価基準である⁽¹¹⁾。A1 レベルは、自分や他人を紹介することができ、どこに住んでいるか、誰と知り合いか、持ち物などの個人情報について質問したり、答えたりできるレベルである。また、相手がゆっくり、はっきりと話して、助け舟を出してくれるなら、簡単なやり取りをすることができるレベルである。サンプル音声に用いた内容は、学習者が十分に内容を理解できないニュースを用い、音声のトーンや速度により、喜怒哀楽の場面認識が可能かどうかも加え評価を行った。また 12 パターンの音声には、調整していないもの、微調整したものに加え、印象の違いを検証した。学習者は、音声の種類により、楽しい声、悲しい声として場面をとらえることができた。非ニューラル音声では、ロボットのように聞こえるとの回答が多かったが、ニューラル音声については、スムーズに聞こえ、アナウンスのように聞きやすく、音声学習を進めるのに問題ないとの意見が多かった。また、教師の意見では、授業の中ではどうしても文法説明に偏りがちのため、e ラーニングで聴解、および、発話の練習を繰り返すことができると効率がいいとの意見が多かった。また、ニューラル音声については学習者同様、聞き取り易く、授業で使いたいとの意見を得ることができた。開発を続けるにあたり、教材としての有意性があることを確認することができた。

表 3 非ニューラル音声、ニューラル音声を用いた評価用サンプル音声

対象者	非漢字圏学習者、CEFR A1 レベル 日本語講師
音声	非ニューラル音声: Mizuki, Takumi ニューラル音声: Nanami, Keita
記事	235 文字 2021 年 9 月 1 日 NHK 政治マガジン 「眞子さま 小室圭さんと年内にも結婚の方向で調整」 ⁽¹²⁾ より一部引用

表 3 非ニューラル音声, ニューラル音声を用いた
評価用サンプル音声

1	非ニューラル (男性, 調整なし)
2	非ニューラル (女性, 調整なし)
3	非ニューラル (男性, 速度 1.1, 高低マイナス 20)
4	非ニューラル (女性, 速度 1.1, 高低マイナス 20)
5	非ニューラル (男性, 速度 1.1, 高低プラス 20)
6	非ニューラル (女性, 速度 1.1, 高低プラス 20)
7	ニューラル (男性, 調整なし)
8	ニューラル (女性, 調整なし)
9	ニューラル (男性, ピッチ 0.8)
10	ニューラル (女性, ピッチ 0.8)
11	ニューラル (男性, ピッチ 1.5)
12	ニューラル (女性, ピッチ 1.3)

6. まとめと今後の課題

本稿では, 従来の日本語教育現場では軽視しがちな音声能力, および, 日本語学習者にスムーズな発話を促し, 真正な文脈を提示した, 正しい発話への架け橋となる教授法を検討した. 真正な文脈を提示するため, 音声教材を用いた e ラーニング教材を提案し, 学習者, および, 教師における評価を行った. 音声教材には, 非ニューラル音声, ニューラル音声を用い, 速度やピッチを変更することで, より学習者に真正な文脈の学習に適応可能かどうか検証を行った. 非ニューラル音声では, 長い文章になるとイントネーションが不自然になることが多く, 学習者の学習段階が上がるにつれ, 不十分であることが評価より確認できた. しかしながら, 満足のいく音声もあり, 容易に利用可能な Moodle のプラグインである Poodll の継続利用も一部可能であることが分かった. 細かな音声の変更については, ニューラル音声に頼る必要がある. ニューラル音声では, より人間に近い合成を期待でき, 学習者, 教師共に教材適用において, 高評価を得ることができた. し

かしながら, 既存のニューラル音声は, スムーズで聞き取り安い, 感情が加味されておらず, 学習内容においては, トーンやピッチ, スピードを細かく変更し, 場面に応じた音声教材を使う必要がある. ニューラル音声のカスタマイズし, Moodle に公開中の教材へ API を用いた音声を導入する必要がある. 人間に劣らない高度な合成音声により, より真正な学習教材を期待したい. また, 今後の e ラーニング教材における日本語教授法の一つとして有効であるかどうか, 検証を続けていきたい.

参 考 文 献

- (1) Sato, C.J.: “Sociolinguistic variation and language attitudes in Hawaii, English around the world, Sociolinguistic perspectives, pp647-663 (1991)
- (2) Gilakjani, A.P.: “The Significance of Pronunciation”, English Language Teaching, pp97-107 (2012)
- (3) Shibuya, T., Hurtig, R.: “Prosody acquisition by Japanese learners”, In understanding Second Language Process, pp176-203 (2007)
- (4) 松崎寛: “音声認識技術を活用したプロソディグラフ自動出力システムの開発”, 日本語教育方法研究会誌, Vol 19, No. 1, pp.72-73 (2012)
- (5) 中村かおり: “非漢字圏学習者の負担を軽減する漢字指導の試み”, 拓殖大学日本語教育研究 (4), pp.34-51 (2019)
- (6) 国際交流基金: “2018 年度 海外日本語教育機関調査, pp25-26 (2020)
- (7) 和泉伸一, 池田真, 渡部良典: “CLIL 内容言語統合型学習 上智大学外国語教育の新たな挑戦 第 2 巻 実践と応用”, 上智大学出版, (2012)
- (8) 小林明子, 佐藤礼子, 元田静, 渡部倫子: “日本語教師のための CLIL (内容言語統合型学習) 入門”, 株式会社凡人社, 東京 (2018)
- (9) Moodle Organization (2021 年 11 月 1 日確認), <https://moodle.org/>
- (10) Moodle サイト TTS 音声の評価用サンプル (公開) <https://ns.wakasato.jp/moodle/course/view.php?id=21>
- (11) JF 日本語教育スタンダード, (2021 年 11 月 1 日確認) <https://jfstandard.jp/cefr/ja/render.do>
- (12) NHK 政治マガジン, (2021 年 9 月 1 日確認) <https://www.nhk.or.jp/politics/articles/lastweek/66845.html>