

仙台市方言を合成する方言音声合成の構築と 小学校2年生向け方言学習授業における活用報告

高道 慎之介^{1,2,*} 丹治 尚子^{1,*} 庄司 潤子³ 佐藤 照一⁴ 田村 文子⁵

¹ 東京大学 ² 慶應義塾大学 ³ 仙台文学館 ⁴ 昔話採訪家 ⁵ 読み語り手 * 同等の貢献
shinnosuke_takamichi@keio.jp

概要

次世代への方言継承の手段の一つとして、学習教材としての方言音声資料を作成する方法がある。これまで、この目的に対しては、方言話者による語りが主に採られてきた。他方で、音声合成技術の発展により、標準日本語だけでなく、日本語方言についても比較的自然的な音声を合成できるようになった。本技術の利用は、音声資料の新規生成を可能にするだけでなく、いわゆる「しゃべるAI」を用いた新たな教育体験をもたらす可能性を有する。本論文では、仙台市方言を対象として方言音声合成を構築し、小学校2年生向けの方言学習授業に活用した結果を報告する。授業計画および音声合成の活用方法を述べたのち、児童から得られた反応を分析する。

1 はじめに

日本においては標準語モノリンガル化が進行しており、日本語方言およびそれに伴う文化が消滅の危機にある。文部科学省は、小学校学習指導要領において「(前略) 共通語と方言の違いを理解すること」¹⁾を掲げ、言葉の経時的变化に対する理解を促している。このような教育に向けた教材として、これまでに方言かるた [1] や方言俳句 [2] といった文字教材に加え、方言の語り [3] のような音声教材が用いられてきた。音声教材は、前述の理解を聴覚を通して促す点に加え、特に昔話の語りを基盤とする教育においては、昔話自体が口承文芸であることから、その形式を継承するという意義ももつ [4]。

一方で、テキストから音声を合成する音声合成技術は近年目覚ましい発展を遂げている [5]。発音および韻律の規則をデータ駆動的に学習可能となったことで、合成対象となる言語は、言語資源の豊富な



図1 動画のプロローグ。昔話の語りをイメージさせる。「どんな昔話を聞かせようか」に続き語りが始まる。方言音声に標準語字幕を添えている。

言語から希少言語へと拡大している [6, 7, 8, 9]。これらの音声合成技術を前述の教材作成に転用することで、新たな方言音声教材を作成できる可能性がある。

以上の背景を踏まえ、本論文では、方言音声合成技術を用いた方言教育教材の作成と、その教材を用いて実施した小学校での授業内容について報告する。具体的には、仙台市方言を対象とした方言音声合成を構築し、仙台市内の小学校2年生向けの授業を計画・実施した。以降では、音声合成の構築および授業計画の内容について述べたのち、児童から得られた反応を分析し、所望の教育効果が得られたかを検証する。

2 仙台方言音声合成の構築と映像付き音声資料の作成

小学校2年生向けの授業を想定し、方言で昔話を語る映像付き音声資料を作成した。図1はそのプロローグ、図2は語り部分のスクリーンショットである。以降では、本資料の作成手順について述べる。

2.1 音声合成の構築

仙台市方言の音声を合成可能な方言音声合成を構築した。学習データには、東北地方民話コーパ

1) https://www.mext.go.jp/component/a_menu/education/micro_detail/___icsFiles/afieldfile/2019/03/18/1387017_002.pdf



図2 動画の語り。昔話「蛇がによろによろ」において蛇がとても長いことをイメージさせる。最後は蛇は地球を超えて宇宙まで伸びるが、それでも尻尾は見えてこない。

ス [10] に含まれる話者 F001 の約 3 時間の音声を用いた。当該コーパスの音声データはオープンリールに保存されていたため、音響的な歪みが認められる。そこで、学習済み音声強調モデル Sidon [11] を用いて音声復元を行った。音声合成モデルの学習には、style-bert-vits2²⁾を使用した。

2.2 資料の作成

昔話の選定。 合成音声を作成する昔話を選定した。選定基準として、(1) 子供の興味を引きやすい、動物が登場する話であること、(2) 合成音声の長時間聴取は負担が大きいため、比較的短い話であること、(3) 現代において不適切とされる表現を含まないこと、の三点を設定した。その結果、形式譚に分類されている「長い話 (蛇がによろによろ)」を選定した。本昔話は、山に住む蛇が非常に長く、いつまでたってもしっぽが見えないという内容である。東北地方民話コーパスのアノテーションを担当した佐藤氏 (第四著者) が、コーパスに含まれている昔話を元に文を作成した、この作品に加え、プロローグおよびエピローグをそれぞれ同方言で作成した。

音声の合成。 昔話の読み語りをしている田村氏 (第五著者) の指導のもとで合成音声を作成した。合成後の音声を後加工せずに同氏に提供したところ、(1) コーパスの元音声の話速が反映されており、小学生に聞かせるには話速が速い、(2) 昔話全体にリズムの緩急を付けたほうがよい、(3) 一部の音素に発話の不明瞭さが見られる、という指摘を受けた。(1) については、audiotsm³⁾ の WSOLA を用いて、話速を約 0.8 倍に調整した。(2) および (3) については、音声の再合成による修正が困難であったため、音声

波形を手作業で直接編集することで対応した。

映像資料の作成。 映像については、コーパス音声収録当時の昔話の語りを想起させることを意図し、おばあさんが囲炉裏の前で子供に語りかける場面を設定した。また、「蛇がによろによろ」では、蛇の大きさを視覚的にイメージできる場面を設けた。これらの場面設定をもとに、本郷制作局⁴⁾ に映像制作を依頼し、動画を作成した。方言音声の聴取において児童がつかまづかないよう、プロローグ部分の動画には標準語の字幕を付与した。

3 授業計画

前述の資料を一部に用いた、方言学習の授業を計画した。

3.1 授業目的

授業目的を「昔話の語りや AI 音声による語りの視聴体験、言葉のクイズを通して、昔話や言葉のおもしろさに気付く。」と定めた。文部科学省の定める要領を踏まえつつ、AI の技術そのものへの興味を持てるように指導することを目的とした。

3.2 授業の流れ

授業を以下のように構成した。

導入。 昔話を聞いた経験を問うことにより、昔話に対する興味・関心を高める。

展開 1。 昔話の語りを通して、方言のおもしろさと親近感に気付かせる。言葉のクイズを通して友達と協働して答えを追求することで、互いの考えを伝え合う力を高める。

展開 2。 前述の資料を視聴することにより、AI の技術に興味をもつことができるようにする。加えて、(1) AI とは何か、(2) 朗読を通して児童が漢字の読み方や物語の面白さに気付くように、AI も朗読 (すなわちモデル学習) を通して喋るようになっていくこと、(3) 将来的には「しゃべる AI」が児童の身の回りにあふれることを説明する。

終末。 感想カードを活用して、目的が達成できたかどうか確認する。

2) <https://github.com/litagin02/Style-Bert-VITS2>

3) <https://pypi.org/project/audiotsm/>

4) <https://hongoproduction.com/>

4 実施報告

4.1 実施日時・クラスなど

2025年12月に仙台市内の小学校第2学年に対して授業を実施した。『「せかい一の話」～お話を聞いて、おもしろいところを伝え合おう～』の単元内の単発授業として実施した。2クラスに対し実施し、各クラスの児童数はそれぞれ30名前後であった。

以降では、3.1節に述べた目的を達成しているかを調査するため、児童から得られた反応を分析する。

4.2 「AIは何か知っていますか」に対する児童の回答

展開2においてAIとは何かを説明するために表題を児童に問うた。児童の回答を以下に列挙する。

- 人工知能のことである。
- 0と1からできている。
- ロボットに使われている。
- 家にあるタブレットでよく使っている。

問いに対しいずれのクラスにおいても積極的に手が挙がったことから、多くの児童がAIという言葉を知っていることが伺える。また、いずれの回答も的を射ており、AIの存在を前提にした授業計画が妥当であることが示される。本件について、クラス担当からは「児童がYouTubeの動画を日頃より視聴しているため、AIという言葉を知っているのでは」のフィードバックを得た。

4.3 児童からの授業後コメントの分析

カードに書かれた感想を、授業内容と照合しながら分析した。分類には、GPT-5.1⁵⁾を使用した。学校名や個人名など、個人名を特定する情報は入力しなかった。また、当該入力をOpenAIの学習に使用されない設定にて使用した。以下は、その分類と件数である。

- 昔話に関する感想：10件
 - 知らない知識が増えたことへの喜び：14件
 - ことばの学習、クイズについて：11件
 - 授業全体の感想：5件
 - AIおよびAI音声による昔話に関する感想 13件
- 以上より、本授業の目的であった(1)昔話に対する

関心を高めること、(2)AIの技術に興味をもつことのいずれについても、十分に達成できたといえる。

以降ではさらに、AI関連の感想を以下のように小分類した。

- **AIが何かを知れたこと(4件)**：「ぼくはAIについてあまりわからなかったんです。だけど始めてAIのことについてAIも勉強してるんだあと思いました。」「AIがどんなのか知れて楽しかったです。」など、ポジティブな感情語を含む文がすべてであった。
- **AIの音声は自然であること(3件)**：「AIが読む、むかし話が人げんみたいによんでいてとてもおもしろかったです。」のように、人間の音声と比較して音声の自然性を記述した感想がすべてであった。AIによる音声と人間による音声が異なることを認識しつつも、児童にとって前者が後者と同程度に自然と感じられることは、合成音声を用いた今後の教材作成の参考になると思われる。
- **AIも学習すること(2件)**：「AIがいっぱい音読したり、れんしゅうしたのをはじめてしました。」という、AIの学習過程に触れる感想があった。人間と同様にAIにも学習過程が存在することを理解した児童がおり、AIの学習過程を小学2年生にも伝えられることが明らかになった。
- **「AIのコーナー」自体への感想(4件)**：「AIのお話がおもしろかった。むかしばなしのことがいろいろわかった。」「エーアイの長いへびが、いつまでたってもしっぽが岩からでてこないからいつになったらへびのしっぽがでてくるのかなっておもった。」など、話に関する感想が散見され、合成音声を用いた教材でも、語りの面白さを提供できることが分かった。

総括すると、AIに対する感想はポジティブであり、その構築課程に対する理解や、合成音声が昔話への関心を高める効果をもたらす例があることが明らかになった。

5 おわりに

本論文では、仙台市方言の音声合成を開発し、小学校第2学年向けの方言学習授業に活用した事例を報告した。合成音声付きの動画を用いた授業では、合成音声を用いた方言教育へのポジティブな反応を

5) <https://openai.com/ja-JP/index/gpt-5-1/>

観測できた。

謝辞

本研究は、国立国語研究所 異分野融合型共同研究「歴史的音源アーカイブに向けたオープンコーパスの整備と AI 音声復元技術の開発」、JSPS 科研費 25K00460 の支援を受けて実施した。東北地方民話コーパスを音声合成学習に利用することについて許諾して下さった、民話を収集された佐々木徳夫氏のご遺族に感謝申し上げます。

参考文献

- [1] 佐藤高司. 小学校教員養成課程における方言教材の作成. 共愛学園前橋国際大学論集, Vol. 18, , 2018.
- [2] 平瀬正賢, 前田桂子, 森下幸子, 橋元良太, 中村慧亮. 「我が国の言語文化」としての長崎方言を学び, 詩の鑑賞・創作を通して語彙力と表現力を高め, 郷土愛を育てる小学校国語授業の研究. 長崎大学教育学部教育実践研究紀要, Vol. 19, pp. 9–18, 2020.
- [3] 細恵子, ホソケイコ. 方言を取り入れた昔話の語りの授業プログラム構想—保育者・小学校教員の養成に向けて—. 広島女学院大学児童教育学科研究紀要, No. 9, pp. 13–23, 2023.
- [4] 高木史人. 「昔話」から考える幼小連携——キクチからの涵養を中心に——. 月刊国語教育研究, Vol. 54, pp. 56–57, 2018.
- [5] Xu Tan, Tao Qin, Frank Soong, and Tie-Yan Liu. A survey on neural speech synthesis, 2021.
- [6] Perez Ogayo, Graham Neubig, and Alan W Black. Building tts systems for low resource languages under resource constraints. In **1st Workshop on Speech for Social Good (S4SG)**, 2022.
- [7] Alex Peiró-Lilja, José Giraldo, Martí Llopart-Font, Carme Armentano-Oller, Baybars Külebi, and Mireia Farrús. Multi-speaker and multi-dialectal Catalan TTS models for video gaming. In **Interspeech 2024**, pp. 999–1000, 2024.
- [8] Agathe Wallet, Ilaine Wang, Emmett Strickland, and Pierre Magistry. Evaluating Speech Synthesis in a Nonstandardized, Multidialectal Context: A Teochew Case Study. In **13th edition of the Speech Synthesis Workshop**, pp. 137–142, 2025.
- [9] Kazuki Yamauchi, Yuki Saito, and Hiroshi Saruwatari. Cross-dialect text-to-speech in pitch-accent language incorporating multi-dialect phoneme-level bert. In **2024 IEEE Spoken Language Technology Workshop (SLT)**, pp. 750–757, 2024.
- [10] 高道, 慎之介, 丹治ほか. 東北方言昔話に関する歴史的音声コーパスと機械学習ベース自動音声復元の試み. じんもんこん 2022 論文集, Vol. 2022, pp. 27–32, 2022.
- [11] Wataru Nakata, Yuki Saito, Yota Ueda, and Hiroshi Saruwatari. Sidon: Fast and robust open-source multilingual speech restoration for large-scale dataset cleansing, 2025.