

絵本の読み聞かせを目的とした 感情を込めた合成音声の生成

向 直人 今井理絵 福永智子

概要

絵本の読み聞かせは、幼児保育・教育現場で広く取り入れられており、保育所保育指針や幼稚園教育要領では重要な活動として位置づけられている。一方で、共働き世帯の増加により、親が子どもと関わる時間が減り、絵本の読み聞かせの機会も減少していると予想される。そこで、本研究では、子どもが気軽に絵本の読み聞かせを楽しむことができるように、感情を込めた読み聞かせのための合成音声の生成を試みる。絵本の文章に含まれる単語に対してネガポジ判定を適用し、ネガポジのスコアに応じて「喜び」と「悲しみ」の感情を付与した合成音声を生成する。大学生17人を対象とした実験の結果、「喜び」の感情を込めた合成音声聞き取りやすく、読み聞かせに適していることが明らかとなった。

1 はじめに

絵本の読み聞かせは、幼児保育・教育現場で広く取り入れられており、保育所保育指針や幼稚園教育要領では重要な活動として位置づけられている [1]。その目的は、言葉の獲得だけではなく、保育士や親との心を通わせることで、愛情に基づいた願いや気遣いを、子どもが絵本と一緒に受け止めるという経験でもあるとされる。楊らは、熟練保育者への聞き取りから、読み聞かせの本当の意味は、絵本の世界に入り込み共に「生きる」感覚であると述べており、話し手となる保育士や親と、聞き手である子どもとの関係形成に大きく影響することが示唆されている [2]。

男女共同参画局がまとめた令和2年の男女共同参画白書¹⁾によると、女性の家事・育児・介護時間は年齢により傾向が大きく異なり、男性は年齢による差が小さいことが示されている。特に25歳～29歳の女性の家事・育児・介護時間は昭和

51年から大きく減少している。一方で、仕事等時間は1.7倍に増加しており、仕事に偏重する傾向が確認できる。男性の家事・育児・介護時間は微増しているものの、女性に比べれば低水準であり、仕事等時間に大きな変化はみられない。また、共働き世帯は、昭和55年には約600万世帯であったが、令和元年には約1200万世帯を超え、男性雇用者と無業の妻から成る世帯をはるかに上回っている。この状況から、親が余裕を持って育児や保育に取り組むことは困難になっていることが予想され、絵本の読み聞かせの機会も減少していると考えられる。

近年、読み聞かせの機会を増やすための手段として、絵本アプリの活用が注目されている。株式会社エバーセンスの「電子絵本読み放題サービスPIBO (ピーボ)」²⁾ もその一つであり、360冊以上の絵本を無料で提供するスマートフォン向けのアプリである。また、Amazonが提供する「Audible (オーディブル)」³⁾ はプロのナレーターが朗読した本をアプリで聴けるサービスであり、ビジネス書や小説だけでなく、絵本・児童書も提供されて

いる。

上記のアプリを利用することで、気軽に自宅での絵本の読み聞かせが可能になる。しかし、読み聞かせの音声の作成は簡単ではなく、誰にでもできるわけではない。また、アプリの多くは有料のサービスであり、その費用も必要になる。そこで、本研究では、絵本の読み聞かせに適した合成音声を生成することを目的とする。感情表現のない無機質な合成音声では、話し手と聞き手の関係性の構築を目的とするような読み聞かせは難しいと考えられる。そこで、喜び、怒り、悲しみなどの感情付きの合成音声を生成することで、心を通わせるための読み聞かせを再現し、学習効果の向上を狙う。実験では、感情を付けた合成音声と、感情を付けない無機質な合成音声を比較し、その有用性を検証する。

本稿の構成は以下である。第2章では、絵本の読み聞かせに関連する研究について紹介する。第3章では、ネガポジ判定を利用した感情付きの合成音声の生成方法について解説する。第4章では、生成した合成音声による読み聞かせの実験とその評価について述べる。第5章で本稿をまとめる。

2 関連研究

本章では、本研究で対象とする絵本の読み聞かせに関連する研究について述べる。関連研究は多数あるが、ここでは、読み聞かせに用いる音声に注目した研究について着目する。

祝儀園らは、「拡張現実を利用した絵本読み聞かせシステムの構築 [3]」において、子どもが集中して絵本を読むために、拡張現実を用いた仮想立体絵本のシステムを構築した。紙媒体の良さであるページをめくる動作は残しつつ、絵本のキャラクターの動作や音声などが再生される仕組みとなっている。キャラクターの音声や動きを取り入れることで、子どもの絵本に対する集中力が向上

することが観察された。一方で、読み聞かせの音声には合成音声を採用しており、抑揚のなさが、集中力を途切れさせる要因になったと述べている。

松村らは、「絵本の読み聞かせ時の演じ分けが子どもの物語理解と物語の印象に与える影響[4]」において、キャラクターの声の演じ分けが、子どもに与える影響に関して評価した。事前に録音した「演じ分けをした読み聞かせ」と「演じ分けをしない読み聞かせ」を比較し、物語の理解度と印象に関して調査した。実験の結果、物語の理解度に関しては、演じ分けによる有意な差がないことが確認された。一方で、演じ分けした読み聞かせは、物語の印象を高める効果があることが示唆された。

本研究では、子どもにとって聞きやすく、親子の読み聞かせを再現した感情表現が豊かな音声を生成することが目的である。上記の関連研究の結果は、肉声に比べ合成音声は集中力の低下を招く可能性を示しているが、演じ分けなどキャラクターの感情を表現することで、物語の印象を高め、効果的な読み聞かせが実現できる可能性を示している。

3 ネガポジ判定を利用した感情付き合成音声の生成

本章では、本研究で使用した「Oseti」と「AITalk WebAPI」を使った合成音声の生成方法について述べる。Osetiは日本語の感情分析ライブラリであり、文章に含まれる単語から評価極性(ポジティブ、ネガティブ)のスコアを算出する。また、AITalk WebAPIは合成音声を生成するクラウドサービスである。

3.1 テキストの生成

最初に、絵本に書かれている文章をテキスト化する。ここでは、誰もが知っていて、ストーリー展開が分かりやすい「それいけ! アンパンマン

[5]」を例に説明する。「それいけ! アンパンマン」の冒頭部分を示す、

あんぱんのかおをした
アンパンマンがかぜにこげちゃいろの
マントを
なびかせて、たっています、
「おや、だれかおなかを
すかして しにそうに
なっているこえがきこえる、
すぐたすけにいかなくちゃ、
それいけ!
アンパンマン!

出典:「それいけ! アンパンマン」(抜粋)

AITalkでは、テキストを1行ずつ読み込んで音声ファイルを生成する仕様となっている。このため、オリジナルのテキストを直接用いるのではなく、文章のまとまりを表すために改行を挿入したテキストに変換する必要がある。また、「それいけ! アンパンマン」は子ども向けの絵本であることから、平仮名が多用されているが、漢字に変換することで自然な合成音声が生じやすくなる。加えて、文章中の句点「。」や鉤括弧は、合成音声で再現されないため、テキストには含める必要がない。上記を踏まえ、「それいけ! アンパンマン」の冒頭部分は次のようにテキスト化する。

あんぱんの顔をしたアンパンマンが
風に、焦げ茶色のマントをなびかせて、立っ
ています
おや、誰かお腹をすかして、死にそうになっ
ている声が聞こえる
すぐ助けに行かなくちゃ
それいけ、アンパンマン

同様の基準で、「それいけ! アンパンマン」の全体の文章をテキスト化する。

3.2 文章のネガポジ判定

合成音声に客観的に感情を付与するため、日本語評価極性辞書 [6, 7] を利用した「Oseti」を利用する⁴⁾。Osetiは文章の評価極性(ポジティブ、ネガティブ)のスコアを算出するためのライブラリであり、日本語感情分析やネガポジ判定と呼ばれる。文章に用いられる単語や用言がポジティブの場合のスコアは1.0、ネガティブの場合のスコアは-1.0になる。上述のテキストの一行ごとにスコアの平均値を算出し、この値に基づいて合成音声の感情を設定する。表1がスコアの例である。「おや、誰かお腹をすかして。死にそうになっている声が聞こえる」には「死ぬ」というネガティブな単語が含まれるため、スコアは-1.0である。また、「すぐ助けに行かなくちゃ」には「助け」というポジティブな単語が含まれるため、スコアは1.0である。一行のテキストにポジティブとネガティブの両方の単語が含まれる場合がある。「ひもじい人の味方。アンパンマンの顔は。とびきり美味しいのだ」には、ポジティブな「味方」と「美味しい」と、ネガティブな「ひもじい」が存在する。この場合は、スコアの平均値 $\frac{1.0+1.0+(-1.0)}{3} = 0.33$ が用いられる。

表1 ネガポジのスコア

文章	スコア	ポジティブ	ネガティブ
おや、誰かお腹をすかして。死にそうになっている声が聞こえる	-1.0		死ぬ
すぐ助けに行かなくちゃ	1.0	助け	
足を滑らせて。崖から落ちた子ザルが泣いています	-1.0	落ちる	
ひもじい人の味方。アンパンマンの顔は。とびきり美味しいのだ	0.33	味方、美味しい	ひもじい

ここで、読み聞かせの候補となる4冊の絵本のポジティブ・ネガティブの単語数を比較する。比較対象は、上述の「それいけ! アンパンマン」に加えて、「ねずみくん [8]」「おこる [9]」「おじいちゃんのパン [10]」である。表2に、各絵本の行数、ポジティブの単語数、ネガティブの単語数をまとめる。この4冊の中では、「それいけ! アンパンマン」は、ストーリーが長過ぎず、ポジティブとネガティブのバランスもよい。そこで、後述の実験では「それいけ! アンパンマン」を採用することにした。

表2 絵本に含まれるポジティブ・ネガティブの単語数

絵本	行数	ポジティブ	ネガティブ
アンパンマン	76	16	11
ねずみくん	146	37	23
おこる	39	5	22
おじいちゃんのパン	35	12	2

3.3 合成音声の生成

合成音声の生成には株式会社エーアイが提供する「AITalkWebAPI⁵⁾」を利用する。このサービスでは、感情表現が可能な話者を選択し、「喜び」「怒り」「悲しみ」の3種類の感情を込めた音声を生成することができる。ここでは、爽やかで若々しい声の特徴の女性「のぞみ」を採用する。本研究では、上述のようにポジティブ、ネガティブの2種類の判定を基に感情を設定することから、「感情なし」「感情あり (喜び)」「感情あり (喜び+悲しみ)」の3パターンの合成音声を生成する。「感情なし」は、いずれの感情も込めず淡々とした話し方である。「感情あり (喜び)」は、スコアが正のテキストに「喜び」の感情を込めて話す。「感情あり (喜び+悲しみ)」は、スコアが正のテキストに「喜び」、スコアが負のテキストに「悲しみ」の感情を込めて話す。各合成音声の生成パターン

を表3にまとめる。ポジティブ・ネガティブの評価極性では判断が難しい「怒り」は合成音声に用いないことにした。

表3 合成音声のパターン

スコア	感情なし	感情あり (喜び)	感情あり (喜び+悲しみ)
0	無	無	無
正	無	喜び	喜び
負	無	無	悲しみ

4 実験・評価

本章では、生成した合成音声による読み聞かせの実験とアンケート結果に関して述べる。

4.1 実験方法

相山女学園大学の大学生17人を被験者として実験を実施した。このうち10人は司書課程を履修している。実験では、上述の3種類の合成音声を被験者に比較してもらう。その際、合成音声の再生位置に合わせて、「それいけ! アンパンマン」のページを前方のスクリーンに表示する。合成音声の再生順番は、「感情なし」「感情あり (喜び)」「感情あり (喜び+悲しみ)」であるが、被験者にはどの順番で再生するかは伝えない。実験の終了後に、表4の設問で構成されるアンケートに回答してもらった。設問5, 6, 7は、「とてもそう思う」「そう思う」「どちらでもない」「あまりそう思わない」「まったくそう思わない」の5段階評価で回答してもらった。

表4 アンケートの設問

設問番号	設問内容
1	子どもに読み聞かせをした経験はありますか?
2	読み聞かせのときに重視したことを教えてください
3	最初の合成音声と2番目の合成音声を比較して。感じたことを簡潔に述べてください
4	最初の合成音声と3番目の合成音声を比較して。感じたことを簡潔に述べてください
5	感情表現が豊かである (5段階評価)
6	音声が聞き取りやすい (5段階評価)
7	子どもの読み聞かせに適している (5段階評価)

4.2 実験結果

4.2.1 読み聞かせの経験

設問1「子どもに読み聞かせをした経験はありますか?」に対して、7人の被験者が読み聞かせの経験があると回答した。設問2「読み聞かせのときに重視したことを教えてください」に対して次のコメントがあった。

- 聞きやすいようにゆっくりと読むこと
- 抑揚をつけて重要なところを強調させて読むこと
- 子どもたちの質問に答えながら読むこと
- 大袈裟に読むこと

子どもが聞きやすいような声の大きさや速度に加え、絵本の内容に合わせて抑揚を付けて読むことを重視していることがわかる。

4.2.2 合成音声の比較

設問3「最初の合成音声と2番目の合成音声を比較して。感じたことを簡潔に述べてください」では、次のコメントがあった。

- 1回目の方は平坦な音声だったけど、2回目の方は声の違いがあって聴きやすかった。
- 楽しい時や喜んでいる時の感情が分かりやすかった。
- 最初の合成音声は感情を表現しないで。淡々と話している感じがした。2番目の合成音声は、最初の合成音声よりも情熱的な話し方で、場面がより想像しやすかったです。ポジティブな明るい声で表現されていて、子供もうれしく感じると思いました。

最初の合成音声は「感情なし」、2番目の合成音声は「感情あり(喜び)」である。被験者には、どちらの音声かを事前に伝えていないが、合成音声の違いを的確に感じ取っていることがわかる。「感情なし」の合成音声は「平坦な淡々とした音声」という印象を与える。一方、「感情あり(喜び)」の合成音声は「ポジティブな明るい音声」という印象を与え、子どもにとっても聴きやすいという意見があった。

設問4「最初の合成音声と3番目の合成音声を比較して。感じたことを簡潔に述べてください」では、次のようなコメントがあった。

- 最初の合成音声よりも3番目の合成音声のほうが抑揚があったり、声色の変化があった。
- 3番目は高い声と低い声両方が混ざった音声に感じました。最初の音声に比べると深刻な場面が明確に分かり、声だけでも感じとることができました。
- 最初の音声より3番目の音声の方が感情が入っていて良いと思いましたが、暗い読み方であり、同時に怖いと感じました。

最初の合成音声は「感情なし」、3番目の合成音声は「感情あり(喜び+悲しみ)」である。「感情あり(喜び+悲しみ)」は、抑揚があることで、

状況を反映した表現ができる一方で、「暗くて怖い」という印象を与えることがわかった。話し手が人間であれば、子どもの様子を見ながら、悲しみの感情の込め方をコントロールできるが、合成音声にはそれができない。合成音声による過剰な悲しみの表現は避けるべきと思われる。

4.2.3 合成音声の評価

設問5「感情表現が豊かである」、設問6「音声聞き取りやすい」、設問7「子供の読み聞かせに適している」の5段階評価に合わせて加重平均を算出した結果が表5である。

感情表現が最も豊かであると感じた合成音声は「感情あり（喜び+悲しみ）」であった。喜びと悲しみの感情の落差が生じるため、他の合成音声よりも感情表現の変化を強く感じ取れたと考えられる。聞き取りやすいと感じた合成音声は「感情あり（喜び）」であった。感情豊かさと評価された「感情あり（喜び+悲しみ）」は、喜びと悲しみの切り替えに違和感があることや、極端な音の高低差があるため、聞き取りにくいと感じたと思われる。子どもの読み聞かせに適していると感じた合成音声は「感情あり（喜び）」であった。「感情無し」と「感情あり（喜び+悲しみ）」の読み聞かせに対する評価は低かった。淡々とした無機質な音声より、感情がある音声の方が読み聞かせに適していると判断された。一方で、「喜び」は子どもにポジティブな印象を与えるが、「悲しみ」は恐怖を与える可能性が指摘されており、評価を大きく下げた。

上記の結果から、「感情あり（喜び）」が、聞き取りやすさと読み聞かせの適正が高く、本研究で生成した合成音声の中では最も評価が高かったと言える。

表5 合成音声の評価（加重平均）

評価項目	感情なし	感情あり (喜び)	感情あり (喜び+悲しみ)
感情表現が豊かである	1.35	3.71	4.18
音声が聞き取りやすい	2.94	4.18	3.24
子どもの読み聞かせに適している	2.18	4.06	2.65

5 まとめ

本研究では、子どもが気軽に読み聞かせを楽しめるように、読み聞かせに適した合成音声の生成することを目的とした。話し手と聞き手の関係構築を再現するため、「喜び」と「悲しみ」の感情を込めた合成音声の生成した。学生17名を被験者としたアンケートでは、感情が無い無機質な読み聞かせより、感情付きの読み聞かせの方が聞き取りやすく、読み聞かせに適していることが分かった。一方で、悲しみの感情表現が過剰な場合、子どもに恐怖を与えてしまう可能性があることも分かった。

絵本には複数のキャラクターが登場するため、文章にはキャラクターのセリフやナレーションなど性質が異なる場合がある。本研究では、「のぞみ」という話者のみを採用したが、今後は、キャラクターごとに異なる話者を用いることを検討したい。さらに、「それいけ！アンパンマン」といった決め台詞に対しても感情を設定するなど、ネガポジ以外の基準にも着目して、読み聞かせに適した合成音声を開発することを目指したい。

注

- 1) 男女共同参画白書：https://www.gender.go.jp/about_danjo/whitepaper/index.html

- 2) PIBO : <https://pibo.jp/>
- 3) Audible : <https://www.audible.co.jp/>
- 4) Oseti : <https://github.com/ikegami-yukino/oseti>
- 5) AITalk WebAPI : https://www.ai-j.jp/products/cloud_webapi/

参考文献

- [1] 杉山喜美恵. 幼稚園教育要領・保育所保育指針における「絵本」の位置づけ。東海学院大学短期大学部紀要, No. 48, pp. 55-66, 2022.
- [2] 楊奕, 多治見里美. 熟練保育者の「語り」から考える絵本の読み聞かせの意味—ことばとからだの関係に着目して—。現代教育学部紀要, Vol. 13, pp. 15-27, 03 2021.
- [3] 祝儀園茜, 水村明日香, 築地立家. 拡張現実を利用した絵本読み聞かせシステムの構築。第80回全国大会講演論文集, 第2018巻, pp. 607-608, mar. 18
- [4] 松村敦, 森円花, 宇陀則彦. 絵本の読み聞かせ時の演じ分けが子どもの物語理解と物語の印象に与える影響。日本教育工学会論文誌, Vol. 39, No. Suppl, pp. 125-128, 2016.
- [5] やなせたかし. それいけ! アンパンマン。フレーベル館, 1981.
- [6] 小林のぞみ, 乾健太郎, 松本裕治, 立石健二, 福島俊一. 意見抽出のための評価表現の収集。自然言語処理, Vol. 12, No. 3, pp. 203-222, 2005.
- [7] 東山昌彦. 述語の選択選好性に着目した名詞評価極性の獲得。言語処理学会第14回年次大会発表論文集, 2008, 2008.
- [8] なかえよしを. ねずみくんのきもち。ポプラ社, 2007.
- [9] 中川ひろたか, 長谷川義史. おこる。金の星社, 2008.
- [10] ちびこママ. おじいちゃんのパン。あかえほ, 2019.

むかい・なおと / 文化情報学部准教授
E-mail : nmukai@sugiyama-u.ac.jp