

レビューに含まれる感情を用いた児童向け絵本分類の検討

白石 諒† 大竹 恒平‡ 生田目 崇‡

中央大学大学院 理工学研究科 経営システム工学専攻†

中央大学 理工学部 経営システム工学科‡

1. はじめに

児童に対し絵本を読み聞かせることは、娯乐的側面があるとともに、言葉とイラストを関連付け言葉の意味を学習させる教育的な側面がある。そのため、年齢や季節、テーマなどによって分類された絵本リストが数多く存在し、絵本選択のための情報源となっている。一方で、児童へ読み聞かせを行う絵本の選択の際には、児童の状態や家庭の教育方針などを考慮する必要があり、これらの状況を加味した絵本の推薦が求められている。

絵本に関する研究の多くは発達心理学における児童教育に関する研究である。例えば、上原らは絵本レビュー中から子供の反応に関する記述から、絵本が子供に与える発達の効果に対して分析を行っている[1]。しかし、絵本のレビュー中に用いられる表現やその表現が有する感情については触れられていない。

本研究では絵本に関するレビュー中の児童の反応ではなく、絵本が持つ感情的特徴を発見することを試みる。具体的には、保護者や保育者が投稿したレビューを収集し、レビュー内に出現する語彙についての分析を行い、その絵本を読み聞かせた際の感情の抽出および、絵本についての特徴付けを行う。これらの結果を用いて、絵本の読み聞かせによりもたらされる感情を加味した絵本の分類を試みる。

2. データ概要

本研究では絵本情報サイト「絵本ナビ」[2]から読者が書き込んだレビューを収集した。収集した分析対象の作品数は417作品であり、1作品あたりの平均レビュー数は104件であった。

3. 分析手順

図1に本研究の分析方法を示す。分析方法としては、はじめに感情を表す単語表（以下、感情単語表）を作成する。次にレビューを対象に形態素解析を行い、文を単語に分割する。分割

した単語に対し、感情単語表を用いて感情のラベル付けを行う。ラベル付けの結果を用いて、絵本ごとに感情ベクトルを算出し、絵本のクラスタリングを行う。

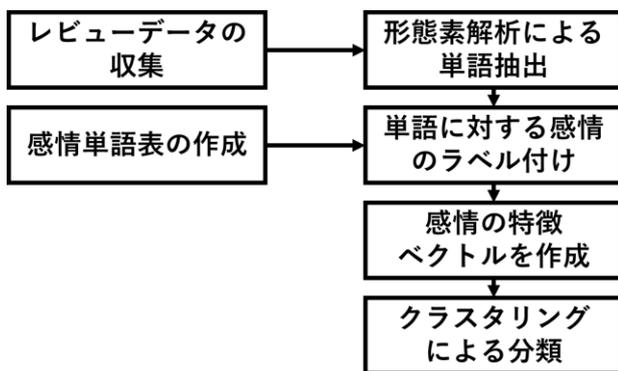


図1：分析手順

3.1. 感情単語表の作成

レビューの感情評価のための感情単語表を作成した。感情単語表の作成には WordNet Affect [3]を用いた。WordNet Affectは WordNet [4]という、シソーラスで単語間の関係を保持している概念辞書に感情が付与されたものである。本研究では WordNet Affect の中に含まれる感情のうち joy (喜び), sadness (悲しみ), negative-fear (恐れ), anger (怒り), dislike (嫌悪), surprise (驚き) の6つを使用した。なお、これらの感情はエクマンが提唱した基本感情(怒り・嫌悪・恐れ・幸福感・悲しみ・驚き)に対応するものである[5]。表1に使用した感情単語表の各感情の単語数と一例を示す。

表1: 感情単語表

感情	単語例	単語数
喜び	明るい, 面白い	398
悲しみ	悲しい, みじめに	277
恐れ	怖い, うろたえ	274
怒り	怒り, むかつと	118
嫌悪	憎む, 嫌い	179
驚き	驚かせる, 唐突	112

3.2. 形態素解析による単語抽出

本研究ではレビューの単語抽出のために形態

Classification of the Children's Illustrated Book Using Emotions in the Review

†Ryo Shiraiishi, Graduate School, Chuo University

‡Kohei Otake and Takashi Namatame, Chuo University

素解析エンジンの MeCab を用いて形態素解析を行う。今回は単独でも意味の通じる名詞、動詞、形容詞、形容動詞の 4 つの品詞に含まれる単語を使用した。

3.3. 単語に対する感情のラベル付け

形態素解析により分割された単語と感情単語表にある単語とのマッチングを行い、それぞれの単語に対して感情ラベルを付与する。また、作品毎に全レビュー中の単語と感情単語表の単語が一致した回数を感情毎にカウントする。

3.4. 特徴ベクトルの作成

作品毎の感情のカウント結果を用いて、絵本の感情を表す特徴ベクトルを作成する。ただし、今回使用した感情単語表の単語数が感情毎に差があり特に「喜び」が多くカウントされてしまう可能性がある。そこで、感情毎にカウントした結果を出現した各感情の単語の種類数で割り、作品毎に感情を割合に直した上で特徴ベクトルを作成した。

3.5. クラスタリングによる分類

絵本の分類を行うため、求めた特徴ベクトルを用いて k-means 法によるクラスタリングを行った。なお、クラスタ数は感情の数と同じ 6 とした。

4. 分析結果および考察

表 2 にクラスタリングを行った結果を、表 3 に各クラスタに含まれる感情の割合を示す。

表 2: クラスタリング結果

クラスタ	1	2	3	4	5	6
作品数	106	72	100	45	46	48

表 3: 各クラスタの感情割合の平均

クラスタ	喜び	悲しみ	恐れ	怒り	嫌悪	驚き
1	0.58	0.10	0.09	0.04	0.17	0.00
2	0.30	0.14	0.10	0.21	0.05	0.15
3	0.36	0.19	0.29	0.01	0.04	0.11
4	0.96	0.01	0.00	0.00	0.19	0.00
5	0.32	0.46	0.11	0.02	0.04	0.05
6	0.34	0.11	0.12	0.04	0.31	0.08

クラスタ全体の特徴として「喜び」の感情がどのクラスタにも含まれている。特徴ベクトル作成の際に、出現した単語の種類数で感情にカウントされた値を割ったが、それでも「喜び」が現れた。その理由として、絵本は元々娯楽的側

面を持っているため、「面白い」といった単語が多く出現しこのような結果が出てきたと考えられる。

次に、絵本がどのクラスタに分類されているかをいくつかの代表的な作品について確認した。「100 万回生きたねこ」（講談社、1977）はクラスタ 5 に属する。この作品は 100 万回生きたねこが最後に愛を知り亡くなってしまうという内容である。そのため、「喜び」よりも「悲しみ」の感情が強く表れているクラスタ 5 に分類されたと考えられる。「ちいちゃんのかげおくり」（あかね書房、1982）はクラスタ 3 に属している。この作品は戦争に関連する話のため、「悲しみ」や「恐れ」の感情が強く表れているクラスタ 3 に所属したと思われる。これらの結果から、レビューに出現する単語から感情の一端を抽出できたと考える。しかし、レビュー中に作品の内容が書かれていると感情の一部が強くなってしまふ作品があった。一例として、作品内容が「お風呂が嫌い」という内容であれば「嫌い」といった単語がレビューに使用され「嫌悪」の感情が強くなる結果となっていた。

5. まとめおよび今後の課題

本研究では、まず作品のレビューに含まれる単語に感情ラベルを付与した。付与した結果を用いて特徴ベクトルを作成し、絵本の分類を行った。分析の結果、絵本が有する感情の一端を捉えることができたと考える。今後、実際に絵本の感情がどの程度捉えられているのか絵本の読み聞かせを行う保護者や保育者に対し、評価実験を行う必要がある。

参考文献

- [1] 上原宏, 馬場瑞穂, 宇津呂武仁, “発達心理学の観点から見た絵本レビュー中の子供の反応の分析”, 言語処理学会第 21 回年次大会論文集, 2015.
- [2] 絵本ナビ, <http://www.ehonnabi.net/>
- [3] Y. Torii, D. Das, S. Bandyopadhyay and M. Okumura, “Developing Japanese WordNet Affect for Analyzing Emotions”, ACL-HLT 2011, pp. 80-86, 2011.
- [4] F. Bond, T. Baldwin, R. Fothergill and K. Uchimoto, “Japanese SemCor: A Sense-tagged Corpus of Japanese”, The 6th International Conference of the Global WordNet Association, 8 pages, 2012.
- [5] P. エクマン, W. V. フリーゼン, 「表情分析入門—表情に隠された意味をさぐる—」, 誠信書房, 1987.