

Bi-LSTM による音象徴モデルと文書埋め込みモデルに基づく義訓生成

Gikun Generation with a Bi-LSTM Sound Symbolism Model and a Sentence Embedding Model

増田陽星[†] 目加田慶人[†]
Yosei Masuda Yoshito Mekada

概要

義訓とは、漢字に対して文脈に応じたその場限りの読みを当てる技法で、創作物を中心に多様な用法がある。本発表では、日本語の入力文に対して自動的に義訓を付与する手法を提案し、評価を行う。提案手法では、親文字とルビの関係から義訓の候補語を取得したのち、選定する。選定には、文脈的妥当性を評価するために文書埋め込みモデルである SBERT、あるいは候補語の持つ音象徴を予測するために構築したモデルを用いる。この音象徴モデルは、語から想起される強弱に注目し、Bi-LSTM を用いて学習を行う。被験者実験を実施し、提案手法によって生成された義訓と実際の用例とを比較することで、その表現的妥当性と有効性を検証した。

1. 背景・目的

「暖^{はる}[1]」のように、漢字に対して文脈に応じたその場限りの読みを当てるルビの用法を義訓といい、その義訓が付与されている文字を親文字という。

義訓には様々な用例があり、そのルビと親文字の組み合わせにより独自の効果を持つ[2]。また、ルビの持つその多様な役割により、現代の文字言語によるコミュニケーションにおいて、音声言語では実現困難な表現をも可能にしているといえる[3]。これらは、インターネット、または書籍の発展により顕在化している。しかしながら、義訓の分析に関する報告はあるものの、親文字とその義訓に関するデータセットが少ないことから、機械学習を使用した義訓の生成手法はない。本研究では自然言語処理によって日本語の入力文に対して自動的に義訓を付与することを目標とする。

2. 関連研究

2.1 音象徴の機械学習による再現

Ramachandran らは、実験により音と概念について一致する連想関係を人が持つことを示した[4]。これを心理学の分野では共感覚 (Synaesthesia) というが、言語学の分野では音象徴 (Sound symbolism) と呼ばれる。しかし、人間による実験から得られた結果には、意味バイアスが含まれており、純粋な音象徴を観察するためにはこれを分離する必要がある。

この問題に対応するため、三浦らは「強さ」という抽象的な概念が音象徴に存在するかを調査し、SVM を使用した機械学習により純粋な音象徴を観察可能にした[6]。

2.2 ルビの体系化

吉田の研究では、多様なジャンルから収集した義訓の用例を分析し、多様化する特性を捉えた義訓の類型化を試み

ている[3]。結果として、自由な文章表現がなされていないが義訓と親文字との間には音声や意味の点で何らかの関係性が保たれていることが示唆された。このように、親文字と義訓において意味的な関係は重要であるといえる。また、渡辺らは歴史的観点から振り仮名の機能の変遷をたどり、現代の出版物にみられる多様な機能性を収集、分類および分析を試みた[2]。

本研究では親文字とルビにおける意味的な関係に注目し、その分類の一部を WordNet による類義語取得と、BERT による Mask 箇所予測の結果から得られるとし、義訓の候補となる語を取得する。

3. 提案手法

義訓を付与する対象となる日本語文を分かち書きしたのち、義訓の候補語を取得し選定する。

3.1 候補語取得

ここでは、親文字に付与する義訓の候補となる語を日本語 WordNet または BERT を用いることにより取得する。

3.1.1 日本語 WordNet を用いた類義語取得

義訓を付与する対象語の標準形を入力とし、日本語の意味辞書である日本語 WordNet を用いて類語を取得する。ここでは、入力と同じ概念 (Synset) に属する全ての単語を対象語の類義語とし、義訓の候補語とする。

以下に、対象語の標準形を入力として WordNet から得られた類義語を付与した例を示す。

たびだち 出発 いま 今 チャリンコ 自転車

これらの親文字とルビの関係は、渡辺らによって挙げられた用例にあるような同義語、つまり「漢字などで表された単語に対してそれと辞書的な意味で相互関係にあるもの[2]」のような関係となる。したがって、親文字に対する義訓の候補となり得ると考える。

3.1.2 BERT による Mask 箇所予測

義訓を付与する日本語文において、親文字を[MASK]トークンに置換し、学習済みの日本語 BERT モデルとトークナイザーにより、Mask 箇所の予測を行い、予測結果上位 5 語を義訓の候補語とする。

以下に、日本語 BERT に対象語を[MASK]トークンに置換した文を入力として得られた予測結果を付与した例を示す。

ひそと 秘密 うきと 浮気 しと 湿原

これらの親文字とルビの関係は、渡辺らによって挙げられた用例にあるようなコンテキストに依存するもの、つまり「ある単語が用いられている環境の中で動的意味において

[†] 中京大学 Chukyo University

関係のあるもの[2]]のような関係となる。したがって、予測された語は対象語に対する義訓の候補となり得ると考える。

3.2 候補語選定

ここでは、取得した義訓の候補語を SBERT、または作成した音象徴予測モデルにより選定する。

3.2.1 SBERT による選定

文脈的妥当性を評価するため、文書埋め込みモデルである SBERT による選定を行う。SBERT を用いて、義訓を付与する対象語が含まれる文のベクトル表現と、その文の対象語を候補語に置換した文のベクトル表現を得る。これらのコサイン類似度が最大の候補語を選定する。

3.2.2 音象徴予測モデルによる選定

人があえて義訓を付けるということは、分かりやすさよりも重視したいものがある場合が存在する。それらを考慮するために、音象徴をモデル化し、その特徴を考慮した選定を行う。モデルの作成には Bi-LSTM[6]を用いる。Bi-LSTM は RNN の一種であり、2つの LSTM レイヤーを持つことで時系列データを双方向から学習でき、様々な NLP タスクにおいて優れた性能をもつ。

4. 実験

作成した音象徴予測モデルと、候補語選定手法の評価のため、2つの被験者実験を実施した。

4.1 実験に用いたモデル

分かち書きには Unidic を辞書とした MeCab を、BERT モデルとトークナイザーとしては東北大学幹・鈴木研究室が公開している bert-basejapanese-whole-word-masking を、SBERT モデルとしては Hugging Face が提供するライブラリである SentenceTransformers の事前学習モデルの stsb-xlm-r-multilingual[7]をそれぞれ用いた。

音象徴予測モデルは、語から想起される強弱に注目し、学習データとして、Web 上のサイトにおいて強そうな言葉として挙げられているもの、また弱そうな言葉として挙げられているものを収集した。収集された語はそれぞれ 254 語、97 語となった。

丹野の実験では、<強い>という形容を得られた音の上位に「ぜ、ダ、グ、ゴ、ゾ」を挙げている[8]。鈴木は、オノマトペ以外の一般語彙においても何らかの意味で音声の清濁音が語の意味に関連している例を挙げている[9]。このように日本語の音象徴を表す場合には、五十音表記を使用するとき、音素表記を使用するときがある。したがって表 1 の例のように、本システムではどちらの特性も学習できるよう、元の表記に加えてカタカナに変換したものと、Julius で用いられる音素表記に変更したものを並列して学習モデルへの入力とする。音素表記への変換は、日本語文字変換ライブラリである jacnv を利用した。

表 1：学習データの表記変形例

元の表記	カタカナ表記	音素表記
金城鉄壁	キンジョウテツペキ	kiNjo:teqpeki
わたあめ	ワタアメ	wata:me

4.2 音象徴予測モデルの評価

語に対する音象徴予測モデルの予測結果が、人間が語から想起する印象とどの程度一致しているか確認するためにアンケートを実施し、6人から回答を得た。

選択肢を 10 語用意し、その中から「最も強そうと感じた語」と「最も弱そうと感じた語」を 1 語ずつ選択させた。これを、意味を持たない語のセットと意味を持つ語のセットに対して 3 問ずつ実施した。また、意味を持たない語としてはランダムなカタカナを 5 文字連結したものを使用し、意味を持つ語としては日本語 WordNet からランダムに取得した語を使用した。

4.3 候補語選定手法の評価

提案手法によって生成された義訓の表現的妥当性と、選定手法の違いによる影響を見るため、実際の用例と比較を行うアンケートを実施し、7人から回答を得た。

実際の用例としては、「当て字・当て読み 漢字表現辞典 [10]」に掲載されているものを使用した（以下、用例と呼ぶ）。また、その用例を基に以下の 5 つの方法で義訓を選定し、選択肢とした。

- 選択法 1 辞典の元の表記
- 選択法 2 候補語からランダムに選択
- 選択法 3 SBERT の類似度が最も高いもの
- 選択法 4 音象徴モデルが最も強いと予測したもの
- 選択法 5 音象徴モデルが最も弱いと予測したもの

被験者には、「最も良い」または「次に良い」と感じた義訓を選択させた。これを 6 問実施した。

ここで、選択肢に候補語からランダムに選択したものを含んだ理由としては、「多様化するルビの働き[3]」が求められる義訓においては有用である可能性があると考えたためである。提示した選択肢の一例を表 2 エラー! 参照元が見つけられません。に示す。

表 2：提示した選択肢例

選択法 1	<small>あらし</small> 戦争だけじゃ得られない。誰も何も。
選択法 2	<small>コンフリクト</small> 戦争だけじゃ得られない。誰も何も。
選択法 3	<small>こうそう</small> 戦争だけじゃ得られない。誰も何も。
選択法 4	<small>しゅら</small> 戦争だけじゃ得られない。誰も何も。
選択法 5	<small>もみあい</small> 戦争だけじゃ得られない。誰も何も。

5. 結果・考察

5.1 音象徴予測モデルと人間の比較

音象徴予測モデルは、語の持つ強弱に関する音象徴を予測として出力し、語から受ける音象徴が<強い>と予測した際には正の値を、<弱い>と予測した際には負の値を出力する。

モデルによる予測結果と、語から受ける強弱を尋ねた設問の回答結果をまとめたものを表 3 に示す。表の数値は、

語の持つ強弱に関する音象徴の予測値を、その設問における選ばれた人数だけ掛けたもの合計である。

表 3：音象徴予測モデルと人間の比較

	意味を持たない語		意味を持つ語	
	強そう	弱そう	強そう	弱そう
1 問目の合計	1.846	-2.848	1.135	-0.314
2 問目の合計	0.655	-0.330	3.904	1.132
3 問目の合計	-1.630	-2.456	-0.191	4.842
平均	0.290	-1.878	1.616	1.887

音象徴予測モデルの結果として、意味を持たない語を弱い、意味を持つ語を強いと予測する傾向にあったため、全体としては意味を持たない語に関する数値は小さく、意味を持つ語に関する数値は大きい傾向にあった。これは、モデルが熟語または漢字が含まれている語が持つ音象徴を強いと予測する傾向にあったためであると考えられる。

また 3 回の平均では、意味を持たない語の数値の差に対して、意味を持つ語の数値の差が小さかった。このことから、意味を持たない語の音象徴に関しては有用性がある程度見られたが、意味を持つ語に対しての有用性はあまり確認できなかったといえる。加えて、意味を持たない語に関しては、人間とモデルのどちらも濁音が含まれる語を「強い」、半濁音が含まれる語を「弱い」と判断する傾向が見られた。これは、丹野の実験でも、<強い>という形容を得られた音の上位に「ゼ・ダ・グ・ゴ・ゾ」を挙げており [8]、本実験と一致する部分がある。したがって、対象となる語の意味による影響が小さいとき、このモデルは人間の感じる音象徴をある程度再現できていると考えられる。

5.2 用例と候補語選定手法の比較

実際の用例と選定手法の違いによる影響を見るために用意した 5 つの選択肢に対する回答結果を表 4 に示す。ここでの表の数値は、選択肢ごとに選択された回数を 6 問分合計したものである。

全ての質問において、辞典の元の表記を「最も良い」と回答した者が大多数であった。また、「次に良い」とされたものは SBERT による選定が比較的多かった。これは、人間である書き手が考えたものが受け手にとっても自然であるという前提に基づくならば、被験者がより違和感の少ないものをよいものとして選択した可能性が考えられる。つまり、文脈を考慮できるコサイン類似度が違和感の少なさに繋がり、SBERT による選定が二番目に選ばれたといえる。したがって、一般に広く受け入れられるための表現にすることを重視するのであれば、文脈を考慮できる SBERT を用いたコサイン類似度に基づいて義訓を選定することが適していると考えられる。

しかし、投票結果は実際の用例に偏っており、提案手法による生成の結果が有用であるとは言い切れない。したがって、取得した候補語と用例を比較することでより相応しい候補語の取得手法を吟味する必要があると考える。同様に、音象徴が選定に有用な場面を調査するなどして、より妥当な義訓を付与可能にする必要があると考える。

6. まとめ

本研究では、ユーザーの入力文に対して義訓の候補となる語を取得したのち、SBERT、音象徴予測モデルのいずれかに基づいて義訓を自動生成する手法を提案した。SBERT を用いたコサイン類似度に基づいて選定することで、文脈に沿ったより自然な義訓を選択でき、音象徴予測モデルにおいては、Bi-LSTM を用いて語の強さを学習することで、単語の意味による影響が小さいときに音象徴をある程度予測することができる。これらが評価実験により明らかになった。今後はさらに多様な用例に対して分析を行い実験の規模を拡大することにより、提案手法の有用範囲を検証したい。

表 4：選定手法と実際の用例の比較

	最も良い (人)	次に良い (人)
選択法 1	32	5
選択法 2	2	7
選択法 3	6	17
選択法 4	7	10
選択法 5	0	4

参考文献

- [1] 日吉盛幸, “万葉集に於ける漢字の用字法的研究(2)-歌句漢字の「氣」-”, 大東文化大学紀要, No. 38 (2000).
- [2] 渡辺鉄太, 佐々木真, “送り仮名の現代的類型とその機能”, 日本語科学研究会プロシーディングス, pp. 68-84 (1987), https://www.agu.ac.jp/~makoto/makoto/papers/Watanabe&Sasaki_1987.pdf (accessed 2024-12-28)
- [3] 吉田敬, “ルビの役割: その多様化と文章表現における働き”, 立教大学日本語研究会, Vol. 26, pp. 169-177 (2020).
- [4] V.S. Ramachandran, E.M. Hubbard, “Synaesthesia--A window into perception, thought and language”, *Journal of Consciousness Studies*, Vol. 8, pp. 3-34 (2001).
- [5] 三浦智, 村田真樹, 保田祥, 宮部真衣, 荒牧英治, 東京大学, “音象徴の機械学習による再現: 最強のポケモンの生成”, 言語処理学会第 18 回年次大会発表論文集, pp. 65-68 (2012).
- [6] Alex Graves, Jürgen Schmidhuber, “Framewise phoneme classification with bidirectional lstm and other neural network architectures”, *Neural networks*, Vol. 18, No. 5-6, pp. 602-610 (2005).
- [7] Nils Reimers, Iryna Gurevych, “Sentence-bert: Sentence embeddings using siamese bert-networks”, In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics (2019), <http://arxiv.org/abs/1908.10084> (accessed 2024-12-28).
- [8] 丹野眞智俊, “日本語音韻における音象徴の存在”, *児童教育学研究*, Vol. 22, pp. 1-10 (2003).
- [9] 鈴木孝夫, “音韻交替と意義分化の関係について”, *言語研究*, pp. 23-30 (1962).
- [10] 笹原宏之, “当て字・当て読み 漢字表現辞典”, 株式会社三省堂, 第 3 版 (2012).