

アクセント単位形の推測を用いた日本語複合語のアクセント句の合成

青柳詠美¹ 小島正樹²

¹東京薬科大学 生命科学研究科 ²東京薬科大学 生命科学部
 {s168002, mkojima}@toyaku.ac.jp

概要

現在のテキスト合成音声システム[1][2]では複合語句内において、話者による強調表現は再現できないという問題がある。そこで本研究ではNHKアクセント辞典[3]に掲載されているアクセント単位形に着目し、コンピュータで推測したアクセント単位形と話者の指定する強調から実際に発生しうるアクセント句を推定するアルゴリズムを作成した。まず話し言葉コーパス (CSJ) [4]に掲載されている長単位の複合語を対象に、アクセント核の位置からアクセント単位形を推定するモデルを作成する。そのモデルとCSJのBIラベルから予測される強調に基づいてアクセント句を合成し、コーパス内のアクセント句を再現できるか評価する。

1 はじめに

近年は様々なテキスト合成音声システム(text to speech)が開発され、AIアナウンサーや自動音声案内システムなど、人に情報を伝える重要な技術になりつつある。本研究では複合語に着目することで、伝え手が意図した韻律の音声を提供できるTTSの研究・開発に貢献することを目標とする。

1.1 強調を含めたアクセント合成とは

齟齬なく情報を伝達するためには、意味に対して正確なアクセント句を提供することが重要である。しかし上記に挙げた先行研究は共に実際に強調したか否かは入力値や特徴量にはなっていない。これは話者による強調を含んだ表現は再現できないため、システムでは正確なアクセントが付与された音声を提供できているとは言い難い状況である。

具体的には以下の例がある。

「長野新幹線車両センター」という場所が現実存在するが、この建物は北陸新幹線の建設の際に建設された歴史を考慮すると構文解析は[長野][新幹線車両センター]である。しかし「長野新幹線」という

固有名詞は存在する(かつて呼称として東京-長野間を「長野新幹線」と呼んでいた)ため、この知識がなければ[長野新幹線][車両センター]と捉えられても文法上の誤りではない。この単語であれ、他の似たような構文構造の単語であれ、システム利用者が後者のアクセント句を出力することを想定していないとは否定できない。

どちらの構文構造も文法上取ることが可能である。またアクセント辞典の記載では単語によっては、時代変化によって複合語の意味要素が薄れ、一つの単語として振る舞うようになる例も存在し、その変化は境目がなく、意味構造を持った発音と一つの単語として振る舞う発音の両方が見受けられる場合があるという指摘もある。

このことから一時的にかかり受け解析の予測等で確率として高い一つの回答に絞る事は可能であるが、どちらの選択肢も用意し、係り受けといった専門知識がなくてもユーザーが容易に選択できるように選択できるTTSの必要性を考えた。

そこでNHKアクセント辞典に掲載されているアクセント単位形に着目した。アクセント単位形は複合語のアクセント位置の考え方の一つである。強調時と強調以外のアクセント変化を説明する事が可能になるため、上記に挙げた構文変化に伴うアクセント変化を説明できると考えた。

本研究ではこのアクセント単位形をコンピュータで推測し、アクセント単位形と話者の指定する強調から実際に発生しうるアクセント句を推定するアルゴリズムを作成した。その上でコーパス内のアクセント句を再現できるか評価する。

2 実験方法

2.1 構築システムの概要

アクセント核の位置からアクセント単位形予測器を作成した。具体的にはアクセント核とアクセント単位形の関係からルールを作成して、自作でラベリ化、CRF(条件付き確率場:機械学習)で学習した。次

に強調を含めたアクセント予測実験を行なった。具体的には、上記アクセント単位形推定モデルと強調位置を入力値としてコーパス内の音韻を再現できるか実験した。

2.2 アクセント単位形の推測実験

まず、アクセント単位形を推測するために複合後のアクセント核の位置を入力値としたアクセント単位形を決定するアルゴリズムを以下の図のように作成した。

1. 該当アクセント核が平板型以外なら、次の単位とアクセント単位は独立(1)
2. 該当アクセント核が平板型で、次の単位も平板型ならアクセント単位は独立(1)
3. 該当アクセント核が平板型で、次の単位は平板型以外ならアクセント単位は継続(0)
4. 該当単位が最終単位ならアクセント単位は独立(1)

図 1 アクセント単位形決定アルゴリズム

次に、日本語話し言葉コーパス(CSJ)のコア文章 201 講演を対象に短単位が 2 つ以上で構成され、各品詞(自立語のみ)にエントリーされている長単位を抽出し、テストデータ:訓練+検証データを 2:8 に分け、CRF++[6]を用いてアクセント単位形が次の単位と結合する(0)・しない(1)を推定した。CRF(条件確率場:Conditional Random Fields)[5][6]とはある入力値に予め付与したラベルの確率分布を計算することで予測する機械学習のアルゴリズムの一種である。今回は特徴量に書く品詞情報や読み情報、表層形情報を入力し、アクセント単位形が次の単位と結合する(0)かしない(1)か単語末表現(EOS)かをラベルとして出力した。EOS は単語末に確実に出現するように設定したため、事実上アクセント単位形の結合可否(0, 1)の二値問題である。条件によって学習形態素数が異なるため別途表で示す。

評価には代名詞、副詞以外では 4-fold grid search を行い、正則化関数は Macro f 値の高かった L1 を、ハイパーパラメータ(-c オプション)は [0, 0.01, 0.05, 0.1, 0.5, 1, 5] から最も Macro f 値の高かった 1 を評価に用いた。代名詞、副詞ではデータ数が少なかったため、交差検定は行わず、訓練データと検証データで上記と同様にハイパーパラメータの調整を行った。特徴量は該当単位 2 つ前から 2 つ先までの trigram、unigram、該当単位 1 つ前から 1 つ先までの bigram を特徴量とした。

	表層形のみ	表層形と読み	表層形と品詞情報
名詞	17326	17371	18842
代名詞	48	48	49
動詞	3914	3913	4214
形容詞	251	252	281
形状詞	645	652	677
感動詞	56	69	70
接続詞	160	162	229
副詞	59	57	71
全体	22472	22493	24432

表 1 各条件における学習形態素数一覧

品詞情報を用いる場合は Unidic[7]のものを用いて、CSJに記載されている[表層形, 品詞(LUWPOS), 活用の種類(LUWConjugateType), 活用形(例:未然形 LUWConjugateForm), その他の情報 1(助詞の種類を示す LUWMiscPOSInfo1)]の unidic で相当する部分を抜き出して一単位として入力値とした。Unidic への変換には Mecab(0.996)[8]で変換を行い、辞書にはアクセントの位置が記載されている Unidic(unidic-cwj-2.3-2.0)を用いた。

2.3 アクセント句予測実験

アクセント単位形の推測結果と CSJ の BI ラベルを用いて、コーパスに記載された実際に発生したアクセント句を再現できるかを実験した。以下は本実験での BI ラベルの解釈である。

BI ラベルとは実際の音のピッチから判断された”次の単位との結合度”を示す[9]。

- 同じアクセント句である” 1”
 同じアクセント句でありながらポーズが入る” 1+p” ,
 違うアクセント句である” 2”
 違うイントネーション句である” 3”

今回実験では” 1”のみを” 強調なし(0)”、それ以外を” 強調あり(1)”と処理

図 2 本実験での BI ラベルの解釈

ベースラインの規則由来のアクセント合成には以下のものを使用した。作成には複合語のアクセント規則(句坂, 佐藤, 1983)や数詞のアクセント規則(宮崎, 1984)、OPEN J TALK[1]、TASET[2][10]を参考にした。

全ての条件で形態素解析に Mecab(ver0.96)と Unidic(ver2.3.0)辞書を用いた。

3 結果と考察

本結果は昨年情報処理学会で発表した名詞句のアクセント単位形推定実験の結果[11]も同時に載せている。

- ・sur_la:表層形のみが学習データ
- ・sur_pho_la:表層形と読みが学習データ
- ・sur_pos_la:表層形と読みと品詞情報が学習データである条件を示している。

3.1 品詞ごとにアクセント単位形推定モデルを作成した結果

品詞ごとにアクセント単位形を学習させた結果を付録 A-1 の図に載せる。どの品詞もアクセント単位形 1 ラベルの f 値より 0 ラベルの f 値が低いという結果になった。またデータが少ない副詞は 0 ラベルが学習できなかった。

動詞等活用形がある品詞も今回は実験の対象のため、品詞情報を付与した場合のアクセント単位形推定結果をまとめたが、どの品詞も結果が悪化したか大幅な改善は見られなかった。

3.2 品詞ごとの実験の考察

副詞の 0 ラベルの f 値や他の品詞の f 値の改善のため、データを他の組成が似た品詞と合わせる学習を試みた。理由として以下の表がある。これは本実験で用いたコーパス内(長単位)で各品詞のどの品詞 SUW(短単位)が多いか、種類ごとに分類したグラフとなっている。

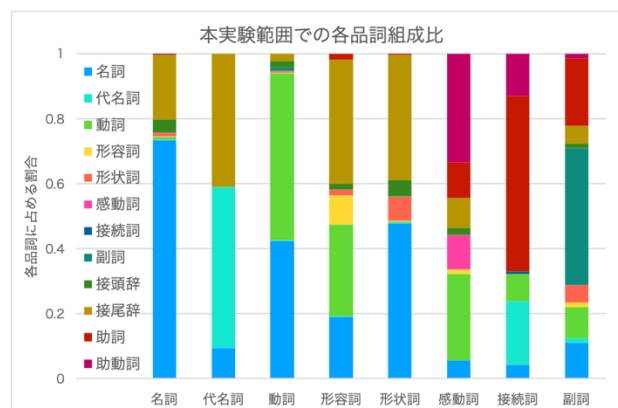


図 3 本実験範囲での各品詞組成

副詞は単独で似たような組成の品詞が見つからないが、副詞自身を除けば他の品詞が満遍なく存在すると考えると、品詞隔てなく学習することで副詞内

の他品詞の学習が進むのではと考えた。

3.3 全ての品詞を含めて推定モデルを作成した結果

全ての品詞を含めてアクセント単位形を学習させた結果を付録 A-2 の図に載せる。表層形のみ、読み追加の条件で副詞は 0 ラベルも含めて学習できたが品詞情報を入れると 0 ラベルの正解例がなく F 値が計算できなかった。また表層形のみで学習で接続詞の 0 ラベルが学習できなかった。これは組成表から考えると接続詞は助詞が多いため他品詞から助詞を含むような例が少なかったためと考えられる。なお、読みを入れた条件や品詞情報を入れた条件では 0 ラベルは学習可能であることがわかった。全体的に他品詞を含めて学習していると品詞情報を入れてもアクセント単位形の推測 F 値は 0 ラベル、1 ラベルともに大幅に上がらないという結果になった。

3.4 アクセント句合成実験結果・考察

次に上記で作成したアクセント単位形推定モデルを用いてアクセント句の推定実験を行った。実験結果について、品詞ごとにアクセント単位形を推定したモデルを用いた条件のものを付録 A-3 の図に、品詞を全て含めてアクセント単位形を推定したモデルを用いた条件のものを付録 A-4 の図に載せる。

品詞全体を俯瞰して、コーパス内に 1+p(強調するけれどアクセント句は切れない)に該当する単語が見られない品詞も多く、たまたまコーパスに無かったのか、本当に日本語としてそう言う事例がないのか疑問が生じた。事例がない場合はアクセント句の合成においてアクセント単位形を考えるまでもないということになる。これは今後検討が必要である。

また両方の実験でベースラインにした規則による強調指定よりもアクセント句の正解率が低い結果はなかった。このことはアクセント単位形の考え方が他品詞にも使える可能性が考えられる。

アクセント単位形の推測実験同様、品詞情報の付与によるアクセント句の正解率が大幅に改善したという例はなかった。

3.5 品詞ごとの考察

以下は各品詞の考察である。主だったものを載せる。

[名詞/代名詞]

品詞情報を付与してもアクセント単位形の推測、アクセント句の正解率ともに改善が見られなかった。これは名詞句というものは品詞情報が変化しないため恩恵を受けにくいと考えられる。また、数詞なしの条件において、規則でアクセント句を推定したものは1+p(強調あるもののアクセント句は結合:0ラベルによるもの)の推測で結果が良いことが分かった。これは提案手法のアクセント単位形の0ラベルの推測成績が悪いことが起因していると考えられる。提案手法の各条件でアクセント単位形推定タスクのf値が異なっていたが、この差はアクセント句合成実験では影響せず、ほぼ同じ正解率を示した。

[動詞/形容詞]

名詞、代名詞のときとは異なり、全体を通してアクセント単位形推測で全品詞同時学習時の結果が良い傾向が得られ、アクセント句推測実験のBIラベル2~のタスクで正解率の上昇に寄与していると考えられる。

[形状詞]

単独学習での品詞情報込みの学習結果のアクセント単位形の0ラベルの推測能力の低いとき、BIラベル2~の正解率の低下しており、品詞一括学習だと改善しているため、0ラベルの推測の正確性がアクセント句の正確な合成に重要であることが言える。

[副詞]

データ数の不足から副詞単体でのアクセント単位形0ラベルの学習は困難であった。副詞の品詞特性としては、副詞の連続(例:ぶるぶるぶるぶる)とそれ以外(例:一生懸命、必ずしも、比較的)に分かれ、副詞の連続以外には名詞や形状詞が含まれていた。このため、品詞を分けずにアクセント単位形を学習した場合には、副詞単独ではアクセント単位形の0ラベルが十分に学習できなかったものが、全品詞を学習させると0ラベルも判別できるようになったと考えられる。

4 まとめ

アクセント単位形の推測モデルを用いてアクセント句の推定を行った。全品詞を俯瞰して、アクセント単位形の推測タスクのF値、特にアクセント単位形の結合を示す0ラベルの推定結果が悪いと、アクセント句の正解率が悪くなる結果となった。またアクセント単位形の考え方は既存手法よりアクセント句の正解率が上がったため、アクセント句の推定において提案手法は有用である。

5 今後の課題

今回の研究の条件下では以下の3つの課題が存在する。

5.1 実験条件について

今回は複合語を入力したときの結果であり、文中でどのような品詞条件で、あるいは構文条件で出現したかについて全く考慮していない。例えば、ある文章中では1単位形として出現する単語が特定の構文条件で2単位形に分割する現象があったとしても、今回の実験で表現することは不可能である。このため、文単位で入力した際のアクセント単位形がどう変化するかを学習させる、あるいは単語の特徴量に品詞情報以外の文中での振る舞いを示す値を含めることが必要であると考えられる。

5.2 実験データについて

正確なアクセント核は人手でないと付与できない。特に平板型のアクセント定義はキャリア文との接合した際のアクセント変化で判断するため、とある文章を入力、発話したときに必ずしも判定できるとは限らない。こう言った事情から多量のアクセント単位形に関するラベルを用意する事は現実的ではない。既に短単位でアクセント核が付与されているコーパス等のデータは限られるため、今後ラベルの正解率やアクセント句の正確な推定を行うにあたり、機械学習の中でデータ不足を補える方法を検討する必要がある。

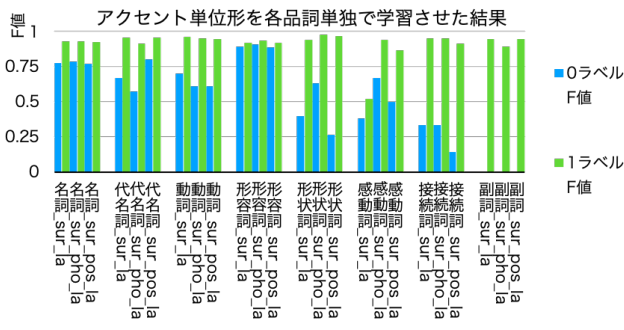
5.3 学習単語単位について

今回、学習する単語単位をUnidic(CSJのSUW)に固定した。本研究では人手で付与された正確なアクセント核の位置を利用したいという意図で採用したが、先行研究ではサブワード[12][13]という概念があり、文法に準拠した単語分割が計算機にとってアクセント単位形を推測しやすいかは検討材料である。分割した際のアクセントの位置について解決した上で、Unidic以外の単語分割での検討をしていきたい。

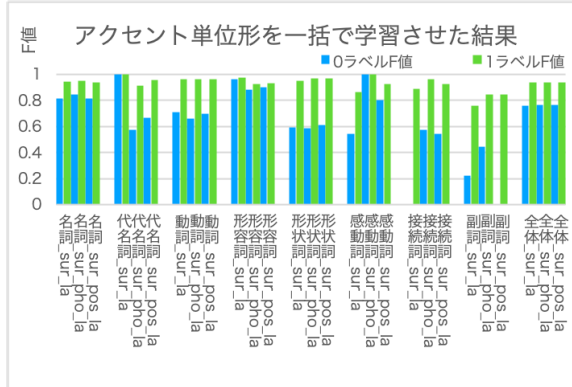
参考文献

1. Open JTalk
<http://open-jtalk.sourceforge.net/>
2. TASET Tokyo Accent Sandhi Estimation Toolkit,
<https://sites.google.com/site/suzukimasayuki/accent>
3. NHK 放送文化研究所・編『NHK 日本語発音アクセント新辞典』 2016 年
4. 国立国語研究所『日本語話し言葉コーパス (CSJ)』
https://pj.ninjal.ac.jp/corpus_center/csj/
5. J. Lafferty, A. McCallum, and F. Pereira, “Conditional random fields: Probabilistic models for segmenting and labeling sequence data,” Proc. 18th International Conference on Machine Learning (ICML), pp. 282-289, 2001.
6. CRF++, Yet Another CRF toolkit
<https://taku910.github.io/crfpp/>
7. Unidic, <https://unidic.ninjal.ac.jp>
8. MeCab, Yet Another Part-of-Speech and Morphological Analyzer
<https://taku910.github.io/mecab/>
9. 五十嵐 他 ” 報告書『日本語話し言葉コーパスの構築法』第 7 章 韻律情報”, 国立国語研究所報告 No. 124(2006)
10. 鈴木 他, ” 条件付き確率場を用いた日本語東京方言のアクセント結合自動推定”, 電子情報通信学会論文誌 D Vol. J96-D No. 3 pp. 644-654, 2013
11. 青柳 他 “アクセント単位形の推測を用いた日本語複合名詞のアクセント句の合成” 情報処理学会第 83 回全国大会学生セッション _7N-02. 2021
12. 工藤拓『実践・自然言語処理シリーズ 第 2 巻 形態素解析の理論と実装』(株式会社近代科学社者)2018 年
13. SentencePiece
<https://github.com/google/sentencepiece>

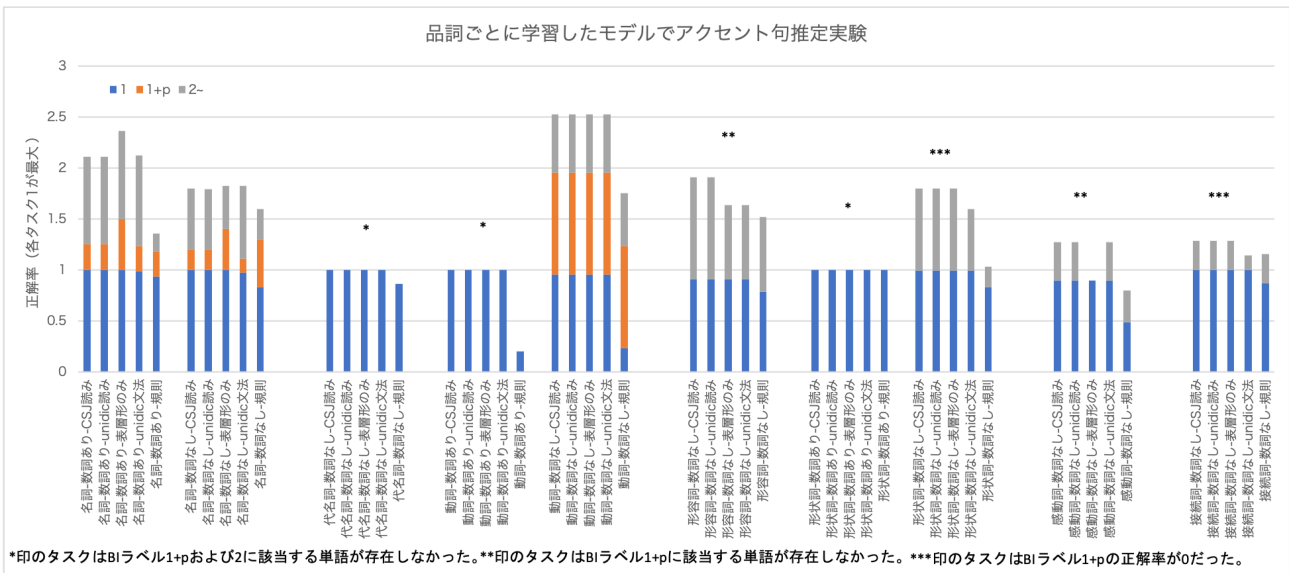
A 付録



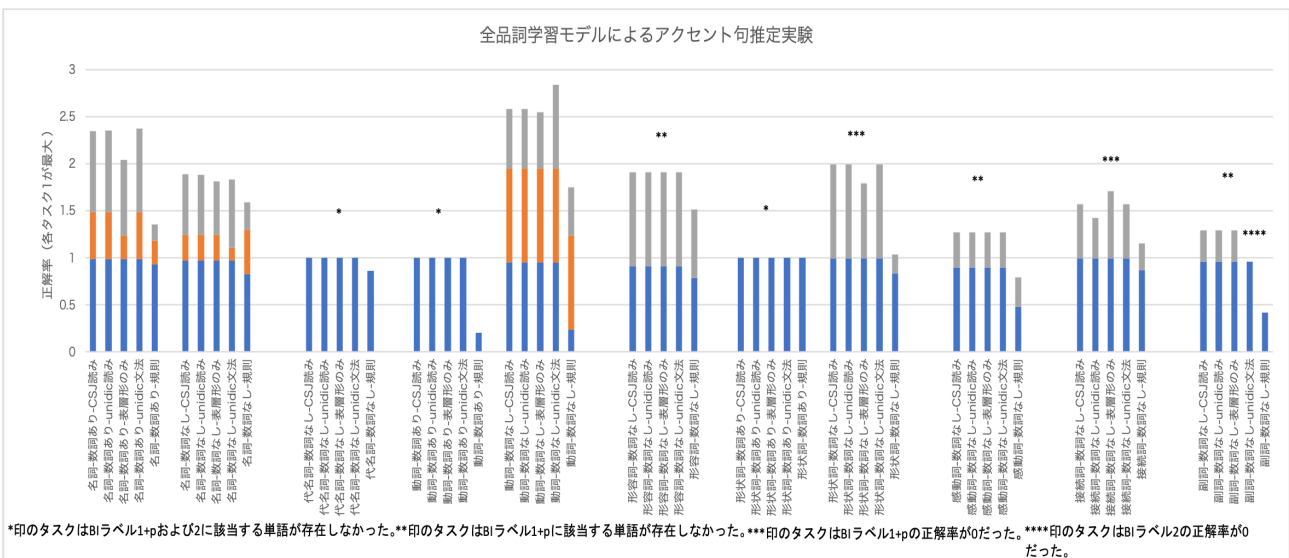
A-1 アクセント単位形学習結果(品詞ごと)



A-2 アクセント単位形学習結果(品詞一括)



A-3 アクセント句推定実験(品詞ごと学習モデル)



A-4 アクセント句推定実験(全品詞学習モデル)