

CRFによる日本語東京方言アクセント結合推定の改善*

☆橋本浩弥, 峯松信明, 広瀬啓吉 (東大)

1 はじめに

日本語のアクセントは、モーラを単位として主に音の高さの高低によって表現されるが、その様子は孤立単語として発声した時と、文として発声した時では大きく異なる。文中の語のアクセントが文脈によって孤立発声と比べて変化する現象を accent sandhi と呼び、特に複数の語が1つのアクセントで発話される現象をアクセント結合と呼ぶ。アクセント結合を記述する規則としては、句法による規則が音声合成を目的として開発されているが [1]、アクセント結合を完全に記述するものではない。これに対して、人手でラベリングしたコーパスを利用し、機械学習の1つである CRF(Conditional Random Field) によってアクセント結合を推定する手法が提案されている [2]。

本研究では、この CRF による手法を元に、次に述べる3つの改良を行った。まず、コーパスの利用効率を高め、その効果を検証した。次に、特徴量に検討を加えることで、さらに推定精度を向上させた。そして、推定が誤る原因を調べ、文節境界に基づくアクセント核位置推定を試みた。

2 アクセントラベルデータベースの利用効率の改善

CRF の学習、評価に用いるコーパスについて述べる。このコーパスは、日本音響学会新聞記事読み上げ音声コーパス (JNAS) から抽出された 6334 文に対して [3]、音感に優れた東京出身東京方言話者 1 名がアクセント句境界とアクセント核位置を一般的な発話速度を想定してラベリングしたものである。

鈴木らの手法では [2]、テキストに対して、Unidic を用いて形態素解析をし [4]、読みがコーパスと一致したものに対してだけ用いていた。しかし、読みが一致しないケースの多くは、濁音、半濁音、促音、撥音等の間違いであったため、Unidic の辞書データから直接正しい読みを探し出し、修正する作業を行った。これにより、6334 文のうち、従来は使用可能なデータが 4775 文 (75%) まで減少していたのに対し、6109 文 (96%) まで利用可能になった。

ここで、コーパスの増加に伴う効果を検証した。従来の利用可能なデータ 4775 文を、学習データ 3776 文と評価用データ 999 文にわけ (Set A)、今回の利用可能なデータ 6109 文を、5110 文と 999 文にわけた (Set B)。そして、各セットの学習データで CRF を学習し、各セットの評価用データを推定した。CRF の学習、推定に用いた特徴量は鈴木らの手法と同様の条件である [2]。結果を Table 1 に示す。アクセント

Table 1 コーパス量とアクセント結合推定精度 (行: 学習コーパス、列: 評価コーパス)

	(a) アクセント句境界推定 (F 値 [%])			
	Set A		Set B	
	全体	複合名詞	全体	複合名詞
Set A	93.7	73.8	90.4	71.0
Set B	94.1	74.9	94.1	78.6

	(b) アクセント核位置推定 (アクセント句中の主核正解率 [%])			
	Set A		Set B	
	全体	複合名詞	全体	複合名詞
Set A	96.4	95.4	96.2	94.1
Set B	96.7	96.1	96.7	96.0

ト句境界の評価指標 F 値は適合率と再現率から算出される。ここで評価に用いられている複合名詞とは、一般に複合名詞になる連続名詞または、名詞の直前の品詞が接頭辞、形状詞、記号-文字、形容詞-語幹を対象とする。Set A は Set B に比べて読みが変化する語が少ないため、Set A で学習したものは、読みの変化が起きる場合のアクセント変化の学習が十分でなく、評価用 Set B に対する性能が低下していると考えられる。

3 特徴量の改良によるアクセント結合推定の改良

鈴木らの手法では [2]、CRF の学習、推定に用いる特徴量の中に、後部要素の語頭変化を記述した語頭変化型、語頭変化形、語頭変化結合型があるが、これに、前部要素の語尾変化を記述した語尾変化型、語尾変化形、語尾変化結合型を追加した。さらに、アクセント核位置推定には、1つ前のアクセント句が連体詞句であるかどうかの2値ラベルを加えた。これは、「人」、「国」などは0型の単語であるが「その」などの連体詞や連体形で修飾する用言が前に付くと、2型に変化するような名詞があるからである。結果を Table 2 に示す。アクセント句境界推定については、名詞連続の推定精度が向上しているものの、全体としての正解率はさほど変わらなかった。アクセント核位置推定は、若干性能向上が見られた。

4 エラー分析とアクセント句境界の考察

アクセント句境界推定結果について、エラー分析したところ以下の問題が判明した。複合語のアクセ

*Improvement in accent sandhi prediction of Japanese Tokyo-dialect utterances using CRF by Hiroya HASHIMOTO, Nobuaki MINEMATSU, and Keikichi HIROSE (The University of Tokyo)

Table 2 素性の違いによるアクセント結合推定の精度

	アクセント句境界		アクセント核	
	(F 値 [%])		(主核正解率 [%])	
	全体	複合名詞	全体	複合名詞
従来手法	94.1	78.6	96.7	96.0
提案手法	94.0	79.8	97.0	96.6

ントについて、複数のアクセント句にそれぞれ1つずつアクセント核があると解釈する場合と、1つのアクセント句の中に複数のアクセント核があると解釈する場合の両方が混在していることである。例えば、「関与せぬ」という複合動詞は、「か' んよせ' ぬ」のように1つのアクセント句に副次アクセントがあると解釈する方法と「か' んよ/せ' ぬ」のように2つのアクセント句があると解釈する方法の2通りが考えられる。(“'”がアクセント核、“/”がアクセント句境界)

これはそもそもアクセント句をどのように定義するかによって変わってくる問題である。つまり、アクセント句を

- 実発話に表れる典型的な韻律パターンのひとつかたまり
- 日本語アクセントを理論的・体系的に解釈(しようと)するための単位

のどちらを目的にするかによって、アクセント句の長さが変わってくる。前者は、実際の発話に近く、直感的に近いものになっているのでわかりやすいという利点があるが、発話速度等によって変化するため、曖昧性が避けられないという問題がある。コーパスはこの前者の考えに基づいてラベリングされており、比較的短いモーラの単語や付属語的な役割を持つ単語は、通常アクセントが強くは表れないため、副次アクセントで表現し、そうでない場合は複数のアクセント句に分けるといったことがされているが、同じような組み合わせの単語列に対して異なるラベリングがされている場合が存在した。

5 文節境界に基づくアクセント句とアクセント核位置推定

音声合成を目的にアクセント結合推定を応用する場合は、本来、話者に依存せず、言語的な制約によって生じる“アクセント”の方が望ましい。なぜなら、テキスト読み上げ音声合成の場合、合成時に使える情報はテキストのみだからである。つまり、前章におけるアクセント句の定義として、後者の立場から、テキストのみによって決定されるアクセントの表現が求められる。

ここで、日本語には単語より長い単位として文節という単位が知られている。一般に文節は、1つの自立語と0個以上の付属語からなり、複合語は1つの自

Table 3 アクセント句境界の違いによるアクセント核位置の推定精度 [%]

アクセント句境界推定	全体	複合名詞
正解	97.0	96.6
文節	96.7	95.9
CRF	96.5	95.5

立語と扱うことが多いが、ここでは、複合語については、いくつかのアクセント句に分かれると考える¹。このように考えると、複合語以外は文節境界をアクセント句境界と考えることで統一的にアクセント句境界を決定できる。そこで、今回は複合名詞のみアクセント句境界を既知とし、それ以外は前述の定義に基づく文節境界をアクセント句境界として、アクセント核位置推定を行った。また、CRFによって推定されたアクセント句境界を用いた場合とも比較した。ただし、条件を揃えるため、CRFによって推定されたアクセント句境界のうち、複合名詞に関しては文節境界を用いるときと同様に正解を与えた。結果をTable 3に示す。文節境界を用いた方が、CRFによって推定されたアクセント句境界を用いるときより、性能が高いことが分かる。これは文節を単位とした方が、句法規則との相性が良く、規則の推定性能が向上しているためと考えられる。

6 まとめ

本稿では、CRFを用いたコーパスベースによる日本語東京方言アクセント結合推定について、コーパスの利用効率を高め、CRFの学習、推定に用いる特徴量の改良をすすめることにより、推定性能を向上させた。また、複合名詞以外については、文節境界を用いることによって、推定されたアクセント句境界を用いるより推定性能が向上した。従来のアクセント結合推定は、現在、日本語オンライン辞書(OJAD)において[5]、韻律読み上げチュータズズキクンにて実運用されている。そこで報告されたズズキクンが正しく解析できないいくつかの例において、提案手法は正しいアクセントを出力している。今後の課題として、複合名詞のアクセント句境界、アクセント核位置推定の精度を改善していく必要がある。

参考文献

- [1] 句坂, 佐藤, 電子情報通信学会論文誌, vol. J66-D, no. 7, pp. 849-856, 1983.
- [2] 鈴木, 他, 電子情報通信学会論文誌, vol. J96-D, no. 3, pp. 644-654, 2013.
- [3] JNAS, <http://research.nii.ac.jp/src/JNAS.html>
- [4] 伝, 他, 日本語科学, 22号, pp. 101-122, 2007.
- [5] 平野, 他, 電子情報通信学会論文誌, vol. J96-D, no. 10, pp. 2496-2508, 2013.

¹CRFの特徴量として用いている日本語係り受け解析器cabochaから推定された文節とは異なる