

◎菊池英明 籠宮隆之 前川喜久雄 (国立国語研究所) † 竹内京子 (東京大学大学院) ‡

1 はじめに

近年、大規模な音声コーパスの構築が各方面で進められており、音声学的な分析や工学的利用が可能になりつつある。こうしたコーパスにおいて韻律情報は有用な情報であり、これを利用しやすくするコーディング手法が期待されている。

これまでに、日本語音声の韻律情報のコーディング手法として ToBI (Tones and Break Indices) モデルに基づく J-ToBI と呼ばれるスキームが提案されており [1]、朗読音声のラベリングスキームとして利用されている。しかしながら、自発性の高い音声にこれを適用する場合、フィラーや言い淀み、句末の多様な韻律構造に対しては問題が生じる [2]。

本研究では J-ToBI を自発音声に適用する際の問題の所在を、多数のラベラーが同一サンプルをラベリングした結果に基づいて分析し、明らかにする。

2 分析方法

現在我々は「日本語話し言葉コーパス」(Corpus of Spontaneous Japanese: CSJ) の構築を進めている [2]。CSJ には、学会講演などから収集した約 700 万語の自発的な音声の一部について、転記情報に加えて形態論情報、分節音情報、韻律情報などの付加情報が収められる予定であり、韻律情報の付与には J-ToBI を採用する。

ここでは、CSJ から抽出したサンプルデータに対して多数のラベラーが J-ToBI ラベリングを行った結果から、J-ToBI を自発音声に適用する際の問題の所在を明らかにする。サンプルデータとして、音声的に特徴のある平均約 30 秒程度の講演音声を CSJ から 9 つ抽出して用意した。ラベリングに先立って、語境界および語のアクセント情報を記す word 層のラベルを形態論情報に基づいて作成し、単語間の韻律的区切りの度合を記す break index 層 (以下、BI 層) の初期値として、word 層ラベルと同じ位置に BI1 を与えた。音韻的トーンを記す tone 層の初期値は特に与えなかった。なお、ラベラーは音声学、言語学、音声工学などの分野を背景に持つ 23 人の研究者であり、分析に際しては後述するように正解率に基づいた選別を行った。

3 tone 層の問題

3.1 tone ラベルの構文正解率

J-ToBI ラベリングのガイドライン [1] に示されたスキームから、tone 層のラベリング構文規則は以下の様に設定できる ("?" は 0 または 1 回の繰り返し、"+" は 1 回以上の繰り返し、"|" はラベル間の境界、"|" は選択可能であることを示す)。

$$((\% (w) ? L,) ? (H - | H * + L | H - , H * + L | * ? | H - , * ?), (w) ? L \% (H \% | H L \%) ? , ?) +$$

この構文規則を用いて構文正解率を求めたところ全体の正解率は 82.2% であり、ラベラー毎の正解率が 80% を下回る 8 人を以降の分析対象から除外した。

表 1 にデータ毎の構文正解率を示す。この表から、データによって難しさの度合が異なるものの、平均して 10% 強の構文誤りが見られることがわかる。続いて表 2 に構文誤りの内訳を示す。

表 1: tone ラベルの構文正解率 (BPM は Boundary Pitch Movement の略)

データ	講演種別	特徴	正解率 [%]
1	学会講演	BPM が特徴的	89.4
2	学会講演	BPM が特徴的	87.1
3	学会講演	BPM が特徴的	89.0
4	学会講演	BPM が特徴的	88.3
5	学会講演	言い淀みが多い	90.6
6	模擬講演	文節内のポーズが多い	90.4
7	模擬講演	半疑問が多い	86.7
8	学会講演	フィラーが多い	85.9
9	学会講演	アナウンサの自発音声	93.1
全体			88.9

表 2: tone ラベルの構文誤り内訳

誤りの種類		割合 [%]
S1	アクセント句内に tone ラベルがない	26.8
S2	同じ boundary tone ラベルが二つ続く	26.8
S3	アクセント句内に同じ tone ラベルが複数	7.0
S4	boundary tone ラベルの方向が逆	28.2
S5	その他	12.2

表 2 の S3, S4, S5 の多くは単純な作業ミスでありラベリング時の構文チェックなどの支援により解消できる。一方、表の S1, S2 はフィラーの位置に多く現れており、現在の J-ToBI ではフィラーについて明確なラベリング基準を与えていないことによるものである。

* Problems on J-ToBI Labeling for Spontaneous Speech

† by Kikuchi Hideaki, Kagomiya Takayuki, and Maekawa Kikuo (Nat'l. Lang. Res. Inst.)

‡ by Takeuchi Kyoko (Tokyo University)

フィルターの定義をはじめ、その出現位置および音調によって tone ラベルをどのように与えるべきかを明確にする必要がある。以下ではフィルターの前後についてのラベルを分析対象から除外する。

3.2 tone ラベルの誤り解析

次に、我々が想定した tone ラベルの正解に対する誤りの要因を解析した結果を表 3 に示す。

表 3: tone ラベルの誤り内訳

誤りの種類		割合 [%]
A1	アクセントの知覚誤り	46.8
B1	BPM の知覚誤り	42.4
B2	boundary tone の強弱の誤り	10.5
その他		0.3

この表において、A1 は、有核語に対してアクセントラベルを付与していないなどの誤りであり、アクセント辞書の整備、初期値作成によるラベリング支援によってある程度解消可能と思われる。

B1 は、韻律句末のピッチの局所的な上昇に対して下降調の boundary tone ラベルを付与しているなどの誤りである。この種の誤りには、多様な BPM に対するラベリング基準を明確にするとともに、ラベラーの訓練によって対処する必要がある。

B2 は、後続の語に影響を受けてピッチが上昇しきらない場合に弱い boundary tone ラベル (%wL, wL%) を付与していないあるいはその逆のケースである。ここで、現在の J.ToBI ではこれらのラベルの区別において後続の語が重音節で始まるかどうかを基準としているが、これについては辞書情報からの自動的な識別が可能であり、韻律情報として付与する必要性はあまりないため今後スキーム自体を見直す必要がある。

3.3 tone ラベル位置の揺れ

最後にアクセント句内に付与するラベル (H*+L, H-) の位置について誤差をはかったところ、H*+L については 60msec、H- については 80msec の平均誤差が見られた。現在の J.ToBI では、F0 曲線が明確な場合には付与する位置が定義されているが、F0 曲線がはっきりしない場合の位置については基準が明確とはいえない。CSJ については分節音情報が利用可能であるため、より厳密な位置に付与することが可能であるが、そうでないケースに対してより厳密な位置に付与するための基準を明確にする必要がある。

4 BI 層の問題

BI 層について、我々が想定した正解ラベルに対する正解率をはかったところ、全体で 78.0% であった。ラベラー毎の正解率が平均を下回る 4 名を以降の分析対象から除外した。

4.1 BI ラベルの一致率

表 4 に BI ラベルの正解との一致率を示す。この表において、あらかじめ初期値として与えた BI1 の精度が最もよく、ラベルの揺れとしては BI2 を BI1 とするケースと BI3 を BI2 とするケースが多い。J.ToBI における BI ラベルの判断基準は主観によるところが大きく、これらの揺れが全て誤りとはいえないが、今後安定性を増すために何らかの工夫が必要である。

なお、ラベラー間の一致率 (安定性) を、偶然の一致率を考慮した Cohen の κ [3] により求めたところ 0.59 であり、我々が想定した正解に対しての一致率 (正確性) については 0.76 であった。このことからまだ安定性を向上させる余地があるといえる。

表 4: BI ラベルの正解との一致率

		付与されたラベル			
		1	2	3	その他
正解	1	94.4	3.0	0.4	2.2
	2	25.4	57.0	5.0	12.6
	3	0.5	16.2	68.4	14.9

4.2 mismatch, disfluency の一致率

BI ラベルのうち、tone と BI の間に矛盾が生じる場合に付与する "mismatch", 言い淀みや言い直してポーズがある場合に付与する "p" については、著しく一致率が低かった。それぞれこれらの特殊な記号を付与しない誤りが最も多いが、同種の現象に対して一人のラベラーが異なるラベルを与えているケースが数多く見られた。

今後、こうした特殊記号を付与すべき対象を数多く提示し、さらにラベラーの訓練を行う必要がある。

5 おわりに

本稿では、自発音声データに対する J.ToBI ラベリングの結果に基づき現行基準の問題の所在を明らかにした。今後、基準を改訂し、大規模コーパスへのラベリング環境の整備を進めていく予定である。

謝辞

ラベリング結果を提供していただいた自発音声韻律ラベリングワークショップ (2000 年 11 月 14, 15 日開催) 参加者の皆様に感謝します。

文 献

- [1] Venditti, J., "Japanese ToBI Labelling Guidelines," Manuscript, Ohio State University, USA., 1995.
- [2] Maekawa, K. et al., "Spontaneous speech corpus of Japanese," *Proc. LREC-2000*, pp.947-952, Athens, 2000.
- [3] Cohen, J., "A Coefficient of agreement for nominal scales," *Educational and Psychological Measurement*, Vol.20, No.1, pp.37-46, 1960.