

○前川喜久雄, 菊池英明 (国立国語研究所) △五十嵐陽介 (東京外語大大学院)

1 はじめに

筆者らが開発を進めている『日本語話し言葉コーパス』(以下CSJ)ではコーパスの一部, 50万語(約45時間)に対して韻律ラベルを提供する計画である[1]. 韻律ラベリング体系としては, 言語学・音声学の研究と相性が良い JToBI[2]を採用することにしたが, 多数のラベラーによるラベリング実験の結果, 自発音声に対して JToBI をそのまま適用すると, 種々の問題が生じることが判明した[3]. 本発表では JToBI に対して自発音声の韻律ラベリングに必要な拡張を施した新しい韻律ラベリング体系 X-JToBI (eXtended JToBI)を提案する[4].

2 分節音情報とトーンラベル

JToBI では分節音ラベルが供給されないために, 分節音の時間区分とトーン位置の関係に関する正確な情報を得ることができない. この問題を解決するために, X-JToBI では, 分節音ラベルを韻律ラベルの一部として供給することにした. これに伴って, トーンラベルに対する時間情報の付与方法にも変更を加えた. 例えば, 韻律句末に生じる HL% (上昇下降調), LH% (“反問の上昇”[5])等の複合境界音調について, JToBI では音調の全体に対してひとつのラベルを付与していたが, X-JToBI では複数のラベルを F0 変化の時点に付与することにした. これによって, 音調のカテゴリだけでなく, 上昇や下降のタイミング情報を得ることができる. 図1に「名古屋で」という発話における上昇下降調と反問上昇調のラベリングを示す. 図中の pH, pL はポイントと呼ばれるラベルで, 複合音調の第一要素(HL%のH)のタイミングを示している. 一方, 発話末には JToBI と同じ HL% が付与されており, 複合音調の存在を示している.

X-JToBI では, 境界音調に限らず, すべてのトーンラベルは対応する F0 変化の時点に付与される.

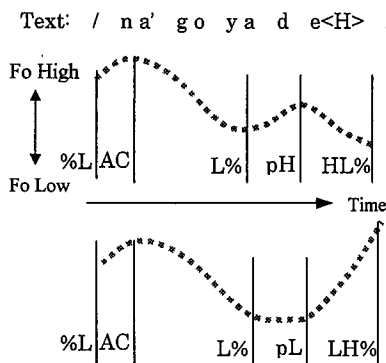


Figure 1. Schematic examples of the tone labeling of phrase-final complex boundary tones. Pointers (pH and pL) are used to denote the time-points of H and L of HL% and LH%. AC stands for lexical accent.

すなわち, F0 の最高値, 最低値, 屈折点などである. しかし, 母音の無声化や F0 の抽出誤り等の理由によって, 当該時刻の F0 値が測定できない場合がある. このような場合には, トーンラベルに補助記号(“x”, “?”)を付加することによって, F0 値の欠損(Ex. ‘L%x’)ないし信頼性の欠如(‘L%?’)を示すことにした.

3 エクステンダー

ある種のバラ言語情報の伝達では句頭や句末の境界音調が一定時間延長されることがあり[6], CSJ にも頻出する. この延長をエクステンダー(延長子)と呼ぶ補助ラベル“<”と“>”によって表示する.

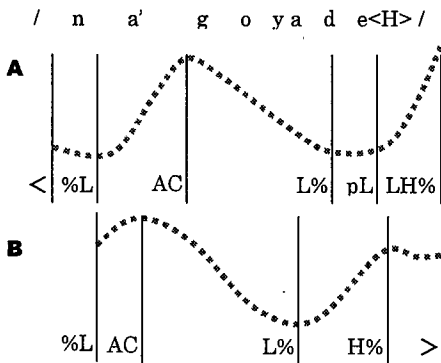


Figure 2. Examples of the usage of extenders. At the utterance initial (top) and utterance final positions (bottom). The rendition of A is a typical contour of “suspicion/disbelief”.

図 2A は, アクセント句頭の境界音調%Lが延長されている. F0 の最低値は延長区間の終端近くにあり, そこにトーンラベル(%L)が付与されている. エクステンダー("<")は低ピッチ区間の始端に付与されており, その右側に対応するトーンラベルがあることを不等号が開く方向によって示している. 図 2B は句末境界音調 H%の延長の例である. トーンラベルは F0 の最高値の時刻に付与され, エクステンダーは高ピッチが維持される区間の終端に付与されている.

4 ブレークインデックス

JToBI を自発音声に適用すると, ブレークインデックス(BI)値の選択に困難を覚えると訴えるラベラーが多い. この問題は, 自発音声においては韻律構造と統語構造との対応が朗読音声に比べてはるかにゆるやかであることに起因していると思われる. また, フィラー(言い淀み), 言い誤り, 語の断片化など, 朗読音声には通常現われることのない諸現象への対応も必要になる. 表1に X-JToBI におけるBIラベルの一

* An Intonation Labeling Scheme for Spontaneous Japanese: X-JToBI
K. Maekawa, H. Kikuchi (Nat. Inst. Japanese Lang.) and Y. Igarashi (Grad. School Tokyo Univ. Foreign Studies).

覧を示す。紙幅の関係ですべてを解説することはできないが、BI ラベルは整数値(0-3)または“整数値+補助符号”の形式で表示される。+記号は何かの理由で左側の整数値よりも強い韻律上の切れ目(break)が生じていることを示しており、補助符号はそのような判断がなされた根拠を示している。例えば“1+p”は、「ワタクシ<pauze>ガ」のように、通常ならばポーズが生じないアクセント句の内部にポーズが生じ、そのために典型的な語境界(BI は1)よりも強い(ただしアクセント句境界に該当する2よりは弱い)切れ目が「ワタクシ」と「ガ」の間に生じていることを示すラベルである。表1の備考欄に OP とある要素はオプションであり、これを利用しないラベリングも可能である。NR(No Recommend)は使用を禁止はしないが推奨しないラベルである。“1+”、“2+”のように根拠を明示せずに BI 値に中間値を用いることは推奨しない。文献[4]にはX-JToBIのBI体系についてのやや詳しい解説を載せた。

Table 1. Inventory of BI labels used in X-JToBI

ラベル	用法	備考
0	従来どおり	OP
1	従来どおり	
1+	1と2の間	NR
1+w	複合語内部の韻律境界	OP
1+p	アクセント句内のポーズ	
2	従来どおり	
2+	2と3の間	NR
2+p	句末に BPM 無 ポーズが後続 ピッチレンジリセット無	
2+b	アクセント句末に BPM ポーズ後続せず ピッチレンジリセット無	
2+bp	アクセント句末に BPM ポーズ後続 ピッチレンジリセット無	
3	従来どおり	
3+	J ToBI の“final”	OP
D	語断片 ピッチリセット無	
D+	語断片 ピッチリセット有	OP
P	語内部のポーズ ピッチリセット無	
P+	語内部のポーズ ピッチリセット有	OP
PB	境界音調の連鎖	
<F	韻律的フィラーの始端	
F	韻律的フィラーの終端	

5 フィラー

フィラー(「エー」「アー」「マソノ」の類)は、自発音声を彩る顕著な特徴である。ただし、何をフィラーと認定するかについて、一般的な合意は得られていない。X-JToBIによるCSJのラベリングでは、フィラー候補リストを準備した上で、リスト中の要素のうち、句頭のピッチ上昇やアクセント類似のピッチ下降などが一切認められないもの、つまり韻律的に平坦に実現された候補語だけを韻律的フィラーと認定し、通常のアクセント句とは異なる韻律ラベルを付与している。例えば「アノー」は、句頭の上昇が認められれば

通常のアクセント句として、全体が平板であれば韻律的フィラーとして処理される。韻律的フィラーはフィラー用のBI(表1参照)で区切られ、トーンラベルは全体に対して“FH”か“FL”のいずれかを与える。前者はフィラー全体のピッチ水準が高いことを示し、後者は低いことを示す。

6 卓立

自発音声には非アクセント句末位置にアクセントとは無関係にピッチの卓立(上昇・下降)が生じることがある。例えば「データが」「」はピッチの下降を示す。この場合アクセントによる「データ」のように発音される(「」はピッチの上昇)のがその例である。この例がそうであるように、この卓立は句末から数えて2モーラ目にピッチの頂点が位置することが普通である。X-JToBIでは、句末から2モーラにこの現象が生じた場合は、句末複合音調 HL%のピーク位置(通常は最終モーラの内部)が、前方に移動したものと解釈してラベリングすると同時に ToBI の備考層(miscellaneous tier)に“PNLP”(Penult, Non-Lexical Prominence)”というラベルを注記している。なお、卓立が感じられる発話であっても「見られて」「も」のように、偶々当該位置にアクセントが存在することがある。その場合は通常のアクセントとしての解釈を優先するが、備考層にはやはり注記を施す。

7 その他の現象

自発音声の韻律は想像を超えて多様であり、上で言及したもの以外にも多くの特異な現象が観察される。頭高型アクセントでないのに、発話冒頭が高く始まる発話、いわゆる半クエスチョンに続いてアクセント句末の H%が生じるために、句末境界音調がふたつ連続する発話などはその例である。X-JToBI はこれらの現象にも対処しているが[4]、今後一層新奇な現象に遭遇することも考えられるので、今後とも仕様の拡張ないし変更が必要となる可能性がある。

8 XML化

CSJの公開に際して、X-JToBI ラベルは書き起こしテキストや形態素情報と共にXML化して公開する予定である。詳細は他の機会に報告するが、韻律ラベルは、それ自体で利用することよりも、形態素情報などと共に利用することが多いと思われる。XML化によってデータ利用の便宜を図りたい。

謝辞: X-JToBI の仕様策定にあたっては 2000 年 11 月に国立国語研究所で開催された『自発音声韻律ラベリングワークショップ』における議論が有益でした。参加者の皆様に深く感謝いたします。

参考文献

- [1]前川 他 (2000).「日本語話し言葉コーパス」の設計. 音声研究, 4-2, 51-61.
- [2]Venditti (1995). Japanese ToBI Labelling Guidelines. (http://www.ling.ohio-state.edu/phonetics/J_ToBI/).
- [3]菊池 他 (2001). 自発音声に対する J ToBI ラベリングの問題点検討. 音響学会講演論文集, 1, 383-384.
- [4]前川・菊池・五十嵐(2001). X-JToBI: 自発音声の韻律ラベリングスキーム. 信学技報 SP2001-106, 25-30.
- [5]川上(1963). 文末などの上昇調について. 国語研究, 16, 25-46.
- [6]川上(1956). 文頭のイントネーション. 国語学, 25, 21-30.