

# 『日本語話し言葉コーパス』のイントネーションラベリング

Version 1.1

前川喜久雄<sup>†</sup>・五十嵐陽介<sup>‡</sup>・菊池英明<sup>\*</sup>・米山聖子<sup>\*\*</sup>・小磯花絵<sup>†</sup>

<sup>†</sup>国立国語研究所 <sup>‡</sup>広島大学 <sup>\*</sup>早稲田大学 <sup>\*\*</sup>大東文化大学

## 目次

1. 本文書の内容 .....	2
2. X-JToBI の構造 .....	2
2. 1 単語層 .....	3
2. 2 分節音層 .....	3
2. 3 トーン層 .....	3
2. 4 BI 層 .....	3
2. 5 プロミネンス層 .....	3
2. 6 注釈層 .....	4
3. 単語層のラベリング .....	5
4. トーン層のラベリング .....	5
4. 1 アクセント .....	5
4. 2 句頭音調 .....	6
4. 3 句頭境界音調 .....	7
4. 4 句末境界音調 .....	7
4. 5 ポインター .....	8
4. 6 エクステンダー .....	9
4. 7 フィラーのトーンラベル .....	9
5. BI 層のラベリング .....	9
5. 1 中間値 .....	9
5. 2 非流暢性 .....	10
6. プロミネンス層のラベリング .....	12
7. 注釈層のラベリング .....	15
8. 非流暢性の BI ラベリングについての補遺 .....	17
8. 1 フィラー .....	17
8. 2 語断片 .....	18
9. ラベルファイルの構造と形式 .....	20
謝辞 .....	21
文献* .....	21

## 1．本文書の内容

本文書は、『日本語話し言葉コーパス』(以下 CSJ)における韻律ラベリングの仕様を解説したものである。ラベリング体系としては、Venditti (1997)に記述された J\_ToBI (Japanese Tones and Break Indices) を国立国語研究所で独自に拡張した X-JToBI(eXtended J\_ToBI)を用いている。X-JToBI の仕様については、前川・菊池・五十嵐(2001), Maekawa, Kikuchi, Igarashi & Venditti (2002)等において報告してきたが、これらの発表後に新しく改定した部分があるので、CSJ のイントネーションラベルを利用する方は、本文書の情報に依拠していただきたい。また CSJ 第三刷作成(2011年)に際して、ラベリング方式を一部変更したので、本文書の Version を 1.0 から 1.1 に変更した。第三刷における変更については 8.2 参照。

X-JToBI は、日本語(東京方言)の音韻的構造分析に基づくラベリング体系である点においては、J\_ToBI と同一である。しかし以下の説明から明らかになるように、

- 1) 分節音層の導入、
- 2) Tone の時間情報の正確な表現、
- 3) Tone 要素の拡張、
- 4) BI 要素の拡張、
- 5) 種々の非流暢性(disfluency)への対処、

など、さまざまな拡張が施されている。これらの拡張は、ひとつには CSJ に格納された自発音声の韻律が朗読音声に較べてはるかに大きな多様性を示すため、それに対処する必要から実施したものである。しかし、自発音声の問題とは別に、J\_ToBI 固有の問題点を解消するために実施した拡張もあり、上記 1)と 2)がこれに該当する。

このうち 1)については、別途マニュアルを提供する(segment.pdf 参照)。また、CSJ に含まれる転記情報、形態論情報についてもそれぞれマニュアルを提供しているので、参照していただきたい(transcription.pdf, pos.pdf)。

X-JToBI によるイントネーションラベリングの対象は CSJ のコア部分である。コアに格納された音声の話者は、東京方言の韻律特徴を有すると判断された話者であり、1名を除いて東京ないし首都圏で出生している(overview.pdf 参照)。

## 2．X-JToBI の構造

X-JToBI のラベル情報は、J\_ToBI 同様、複数の層(tier)から構成されている。すなわち、以下の 6 層であるが、このうち分節音層とプロミネンス層は X-JToBI において新規に導入したものである。

- 1) 単語層(word tier)
- 2) 分節音層(segment tier)
- 3) トーン層(tone tier)
- 4) BI 層(break index tier)
- 5) プロミネンス層(prominence tier)
- 6) 注釈層(miscellaneous tier)

本節では各層の性格について簡単な説明をくわえる。各層におけるラベリングについては 3 節以下で詳しく触れる。

## 2.1 単語層

単語層は、発話を構成する形態論的単位（≒単語）の境界を表示する。CSJ では長短 2 種類の形態論的単位を採用しているが、X-JToBI では、このうち「短単位」とよばれる単位の境界を表示している（pos.pdf 参照）。CSJ の短単位には、韻律情報（アクセント）は含まれていないが、X-JToBI の単語層ではアクセント核の位置がアポストロフによって表示されている。これはアクセント辞書等に記載されている規範的なアクセント位置ではなく、実際の発音を反映した位置である。

## 2.2 分節音層

分節音層は、発話を構成する分節音のラベルを表示する。CSJ で用いた分節音ラベルは、基本的には音素ラベルであるが、一部は音声レベルの情報も表現している。分節音ラベルの詳細は segment.pdf 参照。分節音層をもつことによって、X-JToBI では各種トーンと分節音との時間関係の詳細な検討が可能となっている。

## 2.3 トーン層

トーン層は、イントネーションを構成するトーンの種類と時間軸上の位置を表示する。J\_ToBI においては、トーン層のラベルは、アクセント、句頭の H、3 種類の境界音調(boundary tones)にはほぼ限られており、これらは純粋に音韻論的な要素であったが、X-JToBI では、イントネーションの表現力を向上させるために、句末境界音調の要素を拡張し（4.3 参照）、さらに、ポインター（4.5 参照）、エクステンダー（4.6 参照）等の新要素を導入している。その結果、X-JToBI は音韻的情報だけでなく、音声的な情報も豊富に提供することになったが、同時にラベリングも複雑化している。また、X-JToBI におけるトーン層ラベルは、可能な限り、ラベルに対応する Fo イベントの時刻に付与されている（J\_ToBI は必ずしもそうっていない）ので、トーンにかかわる時間情報も J\_ToBI より豊富に提供している。

## 2.4 BI 層

BI(break index)とは、韻律特徴からみた発話の階層構造を表現する情報である。基本的には 1、2、3 の整数によって表現され、この順により深い階層の境界を表現する。おおまかには、1 は語境界、2 はアクセント句境界、3 はイントネーション句境界に該当すると考えてよい。単語層における語境界はすべて最低でも 1 の BI を与えられている。

ここまでは J\_ToBI における BI と共通であるが、X-JToBI では、自発音声の韻律的な多様性に対応するために、BI を大幅に拡張している。詳しくは 5 節で解説する。

## 2.5 プロミネンス層

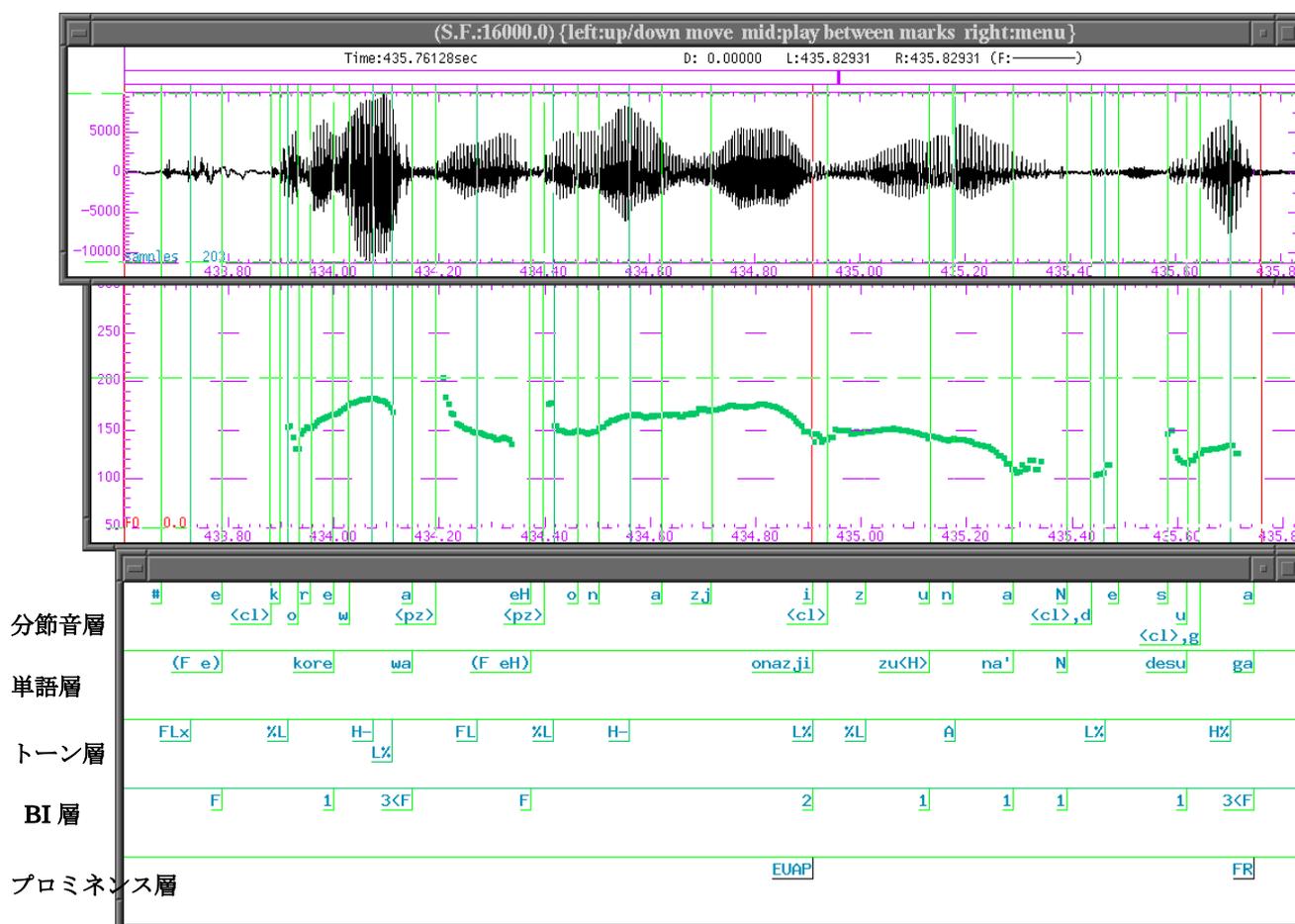
プロミネンス層は X-JToBI で新設された層である。日本語のイントネーションを解釈するうえで必須と考えられる情報のうち、トーンや BI には分類できない情報がこの層に記入されている。その多くはトーンの実現位置の変異に関する情報である。

例えば、東京方言には「デー「タ」が」「ディズニーワールド」で」のように、アクセントとは無関係にアクセント句の句末から2番目のモーラを選択的に卓立する話し方がある。X-JToBIではこれを通常の句末境界音調のひとつであるL%HL%の時間軸上の変異形とみなしており、トーン層にはただL%HL%としての指定だけがなされている。しかし、この場合プロミネンス層に“PNLP”というラベルが付与されるので、それによって通常のL%HL%（つまりアクセント句の最終音声ないしモーラ内部で上昇下降する句末イントネーション）と区別できるようになる。

## 2.6 注釈層

X-JToBIの注釈層は、J-ToBIのmiscellaneous tierと同様、種々雑多な情報を記入するために設けられている。自発音声の韻律的特徴はきわめて多様である。そのために、現在のX-JToBIの仕様でも完全には対応できないことがある。そのような現象が生じている場合に、利用者の注意を喚起するために、注釈層にラベルを記入している。

X-JToBIによるラベリングのサンプルを下に示す。「エ これは エー 同じ図なんですが」という発話の音声波形とFo曲線が上部に示されており、その下に、上から順に分節音層、単語層、トーン層、BI層、プロミネンス層のラベルが表示されている（紙幅の関係で注釈層を省略している）。プロミネンス層の情報(6節参照)からは、発話の中ほどに強調を受けた無核アクセント句があること(“EUAP”)、また、発話末尾に「浮き上がり調」と呼ばれる特殊な句末上昇イントネーションが生じていること(“FR”)が分かる。



### 3．単語層のラベリング

単語層には、発話を構成する単語がアクセント核記号とともに表示している。単語の単位としては短単位(pos.pdf 参照)が用いられており、転記テキストの発音形(transcription.pdf 参照)を分節音マニュアル(segment.pdf)の付録1に拠って分節音ラベル列に変換したものが表示される。単語層ラベルの付与位置は、各短単位の終端である。この位置は同時にBIラベルの付与位置ともなっている(5節参照)。なお単語ラベルの付与位置は、各短単位を構成する最後の分節音の終端である(詳細はsegment.pdf、特に同文書7節を参照)。

### 4．トーン層のラベリング

トーン層に生じるラベルの一覧を表1に示す。またトーンラベルと組み合わされて利用される補助記号の一覧を表2に示す。

表1：X-JToBIのトーンラベル

ラベル	分類	出現位置
A	語彙アクセント	
H-	句頭音調	アクセント句頭からアクセントまでの位置
%L	句頭境界音調	アクセント句頭
(H%)	仮想的句末境界音調	例外的に高いピッチで始まるアクセント句の冒頭
L%	句末境界音調	アクセント句末
H%	句末境界音調	アクセント句末、L%の後
HL%	句末境界音調	アクセント句末、L%の後
LH%	句末境界音調	アクセント句末、L%の後
HLH%	句末境界音調	アクセント句末、L%の後
pH	ポインター	HL%ないしHLH%の前
pL	ポインター	LH%ないしHLH%の前
>	エクステンダー	
FH	フィラー音調	フィラー区間
FL	フィラー音調	フィラー区間

表2：X-JToBIのトーン層補助記号

記号	使用例	意味
x	Ax, L%x 等	トーン時刻のFo情報が存在しない、もしくは信頼できない
?	A?, HL%?	要素の存在(あるかないか)ないし選択(どれか)に自信がない
!	A!	位置が不明確*

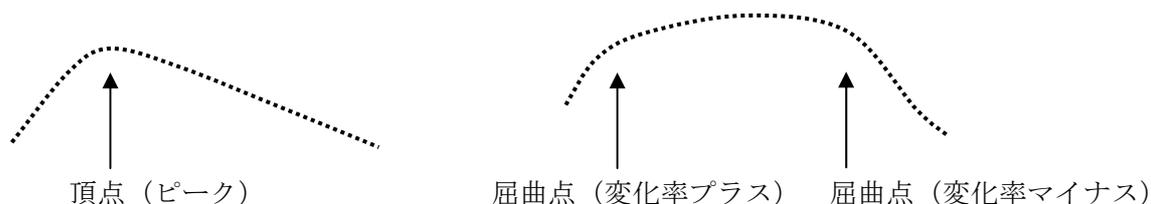
\* E\_ToBI (English ToBI) における"! "の用法とは異なっているので注意

#### 4.1 アクセント

東京方言の語彙的に定まったアクセント核をラベル"A"で表示する。アクセント辞書に有核語(アクセントを持つ語)として記載されている語であっても、無核語(アクセントをもたない語)として発音されていればラベル"A"は付与していない。同じくアクセント辞書に無核語として記載されていても、有核語として発音されていればラベル"A"を付与している。"A"が付与された語に対応する単語

層のラベルでは、アクセント核の位置がアポストロフで示されている。

ラベル”A”の付与位置は、アクセントによる **Fo** の下降が始まる時刻である。**Fo** 曲線に明瞭な頂点（ピーク）が存在していれば、その位置がラベル位置になる。**Fo** が下降しつづけている場合は、下降の加速度が急激に変化する時刻（屈曲点と呼ぶ）にラベルを付与している。ここで頂点とは、**Fo** 曲線が上昇から下降に転ずる（ピークをなす）時刻をさし、屈曲点とは **Fo** 曲線の変化率が符号を変えずに顕著に減少ないし増大する時刻をさす（下図参照）。



**Fo** にはピークも屈曲点も見あたらず、しかも、聴覚的には明瞭にアクセントが知覚されることがある。その場合は、分節音情報を参照して、アクセントを担う母音領域の近傍にラベル”A!”を付与している。このラベルが生じやすいのは主に以下のふたつの環境である。

- 1) H%で終わるアクセント句に頭高型（第1モーラにアクセントをもつ）アクセント句がポーズ無しに後続する場合、
- 2) 先行語のピッチレンジが拡大されており対象とする語のピッチレンジが狭められている場合（先行語が強調されている場合）。

ラベル”Ax”は、**Fo** 抽出アルゴリズムの問題等によって、アクセント核が存在すると想定される母音近傍に **Fo** 値が存在しない、もしくは信頼できない場合に使用する。ラベルの位置はアクセントを担うと想定される母音の終端である。ただし、アクセント核を有すると知覚される母音に無声化が生じている場合の処理は特別である。この場合、無声化母音に後続する有声区間の **Fo** 頂点（始端が最高値であればそこ）にラベル”A”を付与している。この処理は、アクセント核を有する母音が無声化した場合、後続する有声区間の **Fo** が代償的に上昇するという音声学的な観察を根拠としている。

最後に、ラベリング作業者がアクセントの有無の判断に迷い、どうしても自信を持って判断できない場合は、ラベル”A?”を付与している。アクセントの連続などによって、ピッチレンジが極めて狭くなった発話の句末ではこのラベルを用いることが多い。

## 4.2 句頭音調

東京方言のアクセント句頭では通常、第1モーラから第2モーラにかけてピッチの上昇が知覚される。この上昇に該当する **Fo** の頂点ないしプラスの屈曲点にラベル”H-”を付与している（なお、X-JToBI ではアクセント句境界には必ず2以上の強さの **BI** が付与される）。

頂点ないし屈曲点の位置は、分節音からみると、必ずしも第2モーラの内部とは限らない。第1モーラにも第3モーラにも生じることがある。X-JToBI では、句頭の上昇が知覚されるかぎりにおいて、実際に頂点ないし屈曲点が観察された位置にラベル”H-”を付与している。第2モーラが促音であったり、母音に無声化が生じたりした場合は、第3モーラにラベルが付与されることが多い（注：第3モ

ーラも無声化している場合は、第2モーラの終端にラベル”H-x”を付与している)。また稀に、上昇が知覚されるにもかかわらず Fo 曲線がほとんど平坦であったり下降しつづけていたりすることがある。その場合にはラベル”H-!”を第2モーラを中心部に付与している。

ただし、すべてのアクセント句に”H-”が付与されているわけではない。句頭におけるピッチの上昇が知覚されず、Fo も上昇していないアクセント句にはラベルは付与されていない。第1ないし第2モーラにアクセント(核)を有するアクセント句において”H-”が省略されやすい。

#### 4.3 句頭境界音調

アクセント句の始端にはラベル”%L”が付与されている(唯一の例外はラベル”(H%)%L”に関するものである。7節の HBP 参照)。このラベルは、前節で言及した句頭の上昇の起点を示すものであり、句頭の上昇は、”%L”とそれに後続する”H-”ないしアクセント(”A”)の組合せによって表現されている。”%L”は Fo の局所的最低値の時点に付与するのが原則である。最低値が確定できない場合の処理は以下のとおりである。

- 1) アクセント句頭の Fo が全体的に信頼できない場合は、ラベル”%Lx”をアクセント句頭モーラの母音始端に付与している(母音が無声化している場合も同様)。
- 2) Fo の局所的最低値が見つからない場合は、ラベル”%L!”を付与している。
- 3) アクセント句頭の Fo に先行子音の影響などで乱れが生じている場合は、アクセント句頭モーラの母音区間のうち Fo が安定して抽出されている箇所に付与している。
- 4) 頭高型のアクセント句では、Fo 曲線の開始時刻に Fo が最高値をとることがある。この場合、”%L”と”A”は同じ箇所(最高値)に付与している。
- 5) 先行するアクセント句末に句末境界音調”H%”、”LH%”、”HLH%”(4.4 参照)が生じており、アクセント句間にポーズが介在していない場合、当該アクセント句の冒頭が高い Fo 値ではじまり、そのまま単調に下降し続けることがある。このような場合、下降区間中の屈曲点を探してその位置に”%L”を付与している。Fo が途切れなく下降していて屈曲点を確定できない場合は先行句の”H%”、”LH%”、”HL%”、”HLH%”と同時刻に”%L”を付与している。
- 6) アクセント句頭に音声の延長が観察され、Fo も低い値のまま延長されている場合は、”%L”とエクステンダー記号”>”を組み合わせることによって、Fo の延長区間を表現している(4.6 参照)。
- 7) 自発音声には、稀にはあるが、アクセント句の冒頭が高いトーンで始まると解釈すべき発話も生じている。その場合、ラベル”(H%)”を用いたラベリングをおこなう。この現象については7節の HBP 参照。

#### 4.4 句末境界音調

日本語のアクセント句末には多彩なイントネーションが生じる。これを句末境界音調と呼ぶ。表3に X-JToBI における句末境界音調ラベルの一覧を示す。句末境界音調は2種類に大別できる。アクセント句末が単純に下降して終わる場合とそれ以外の場合、すなわち句末に上昇をともなう複雑なピッチ変動が観察される場合である。前者を単純境界音調、後者を複合境界音調(boundary pitch movement; BPM)と呼ぶ。複合境界音調の第1要素は必ず”L%”である。

表 3：句末境界音調のラベリングに使用されるラベル

ラベル	使用するラベル列	特徴
L%	L%	下降調（単純境界音調）
H%	L%, H%	上昇調 1（BPM）
LH%	L%, pL, LH%	上昇調 2（BPM）
HL%	L%, pH, HL%	上昇下降調（BPM）
HLH%	L%, pH, pL, HLH%	上昇下降上昇調（BPM）
pH	HL%,HLH%参照	BPMにおける上昇の頂点
pL	LH%,HLH%参照	BPMにおける上昇の開始点

### 単純境界音調

**L%（下降調）**：句末に複合境界音調が生じていなければ（すなわち下降調であれば）、アクセント句末にラベル”L%”を付与している。なお、アクセント句末には **creaky phonation** が生じたりするためにはしばしば **Fo** 抽出が不安定になり、母音が無声化することも多い。このような場合にはラベル”L%x”（対応する **Fo** が存在しない）を句末モーラの終端に付与している。

### 複合境界音調（BPM）

**L%H%（上昇調 1）**：句末境界音調のうち上昇調 1 は、単純な上昇調である。上昇の始端に”L%”、頂点に”H%”を付与している。

**L%LH%（上昇調 2）**：これは上昇が始まる前に一定期間低いピッチが持続されるタイプの上昇調である。これに対しては、低ピッチ区間の始端に対応する屈曲点に”L%”を、低ピッチ区間の終端に対応する **Fo** 上昇の始端に”pL”を、そして上昇の終端に”LH%”を付与している。上昇調 1 と 2 は、異なるパラ言語的意味を伝達するが、ときに両者の区別が困難なことがある。その場合には **L%H%** に分類している。

**L%HL%（上昇下降調）**：これはピッチが上昇したのち下降するタイプの **BPM** である。上昇の始端に”L%”、上昇の頂点に”pH”、**BPM** 全体の終端に”HL%”を付与している。

**L%HLH%（上昇下降上昇調）**：これは、自発性の高い対話音声で稀に観察される **BPM** である。**CSJ** の音声にはほとんど出現しないが、念のためにラベリング方法だけは規定してある。最初の上昇の始端に”L%”、その上昇の頂点に”pH”、下降の谷に”pL”を付与し、**BPM** 全体の終端に”HLH%”を付与している。

## 4.5 ポインター

以上のラベリングで利用されたラベル”pH”と”pL”をポインターと呼ぶ。ポインターは、**BPM** における **Fo** 変化の時刻を明示するために付与されるラベルである。**BPM** のタイプそのものは、アクセント句末に与えられたラベルが”L%”、”H%”、“LH%”、“HL%”、“HLH%”のいずれであるかによって識別できる。

朗読音声では、句末境界音調は原則として発話の最終音節（ないしモーラ）内部に生じる。しかし

CSJ に記録されているような自発音声では、上昇の始端や頂点が最終音節以前に位置することがある。その場合は、プロミネンス層に特別なラベルを付与して、通常の BPM と区別する（6 節参照）。

#### 4.6 エクステンダー

自発音声では、様々な要因によって、音声の延長（引き伸ばし）が生じることがある。すなわち、発音すべきテキストの一部が言語的要因（長母音など）とは無関係に、知覚可能な程度に延長される現象である。音声の延長はアクセント句の句頭ないし句末に生じることが多い。音声の延長が生じ、なおかつ延長部分のピッチもほぼ一定値を保ったまま延長されている場合、トーン記号と記号">"との組合せによって、延長の区間を表示する。この記号をエクステンダーと呼ぶ。

例えばアクセント句頭で分節音の延長が生じており、かつ当該箇所がほぼ一定の低いピッチで発音されていれば、延長区間（低ピッチ区間）の始端にラベル"%L"を付与し、終端にエクステンダーを付与している。

同様の現象（つまり分節音の延長と一定のピッチの持続）は、アクセント句末でも生じる。その場合は、延長区間の始端に句末境界音調ラベルに含まれる最後のラベルを付与し、終端にエクステンダーを付与している。なお、CSJ のラベリングでは、句末においてエクステンダーを適用しうるトーンラベルを、"L%"と"H%"に限定して作業をおこなった。これら以外の句末境界音調に延長が生じるとは極めて稀であると判断されたからである。

#### 4.7 フィラーのトーンラベル

フィラー（filled pause）にはフィラー専用のトーンラベル"FH"と"FL"を付与している。フィラーのラベリングについては、別途 8.1 で解説する。

### 5. BI 層のラベリング

BI ラベルは、J\_ToBI と同様、すべての語境界（短単位境界）に付与される。J\_ToBI の BI ラベルは、韻律境界の深さに応じて基本的には 1 以上 3 以下の整数値をとる。しかし、自発音声においては種々の原因によって韻律境界が曖昧化するために、従来の整数値の中間値を認め、さらに種々の非流暢性(disfluency)に対処するために、数値を用いない特殊な BI を考案した。以下、X-JToBI における拡張部分について説明する。

#### 5.1 中間値

中間値を示すラベルは、整数 1、2、3 を N とし、アルファベットを X とすると、N+X の形で与えられる。ここで、X は 2 文字となることもある。N+は、その韻律境界の深さが N と N+1 の中間値であることを示しており、X はその境界を中間値と判断する音声学の根拠を示している。X をヌルとすることは許しているが、できるかぎり避けるべきである。

表 4 に BI 層のラベリングに用いられるラベルのうち従来の BI ラベルと中間値ラベルの一覧を示す。表中の「イントネーション句」はアクセント句よりも深い韻律上の境界をさす。イントネーション句の特徴は、そこでピッチレンジがリセットされることである。いわゆるダウンステップ（アクセント核によって引き起こされるピッチレンジの狭め）はイントネーション句境界でリセットされる。

表4：X-JToBIのBIラベル（その1：中間値）

ラベル	用法	備考
1	語境界	CSJでは短単位境界
1+p	アクセント句内のポーズ	ポーズ位置は語と語の境界
2	アクセント句境界	
2+p	アクセント句境界にポーズが後続	
2+b	アクセント句境界にBPMが存在するが、ピッチレンジのリセットがない	ピッチレンジがリセットされていれば3
2+bp	アクセント句境界にBPMが存在し、ポーズが後続するがピッチレンジのリセットがない	ピッチレンジがリセットされていれば3
3	イントネーション句境界	

**1+p**：アクセント句境界を”|”、ポーズを”<pz>”で示すと「フクオカ<pz>ニ | イキマシタ」のように、アクセント区内部の語境界にポーズが生じた場合に付与している。なお、ポーズが単語の内部に生じた場合は表5の”P”を用いることになるので注意。

**2+p**：アクセント句境界にポーズが後続することによって、アクセント境界における韻律上の不連続性が、ポーズのないものよりも強く感じられる場合に付与している。

**2+b**：アクセント句境界に句末複合境界音調(BPM)が生じると、BIは3となるのが普通であるが、ときとしてピッチレンジがリセットされないことがある。その場合に付与している。

**2+bp**：アクセント句境界にBPMが生じており、ポーズも後続しているが、ピッチレンジはリセットされていない場合に付与している。

## 5.2 非流暢性

表5もX-JToBIにおいて拡張されたBIラベルである。これらのラベルは表4のラベルと違ってその強さを1次元上に配列することが困難であるので、整数値は与えていない。

表5：X-JToBIのBIラベル（その2：非流暢性など）

ラベル	用法	備考
P	語中のポーズ	
<D	語断片の直前	8.2も参照
D	語断片の終端	8.2も参照
<F	フィラー区間の開始	“3<F”,”2+b<F”のように先行アクセント句のBIとともに示す
F	フィラー区間の終了	“<F”と共に利用
F2	対話の応答詞のうち、言語情報を保有していないと判断される部分の終端	対話音声にのみ使用
W	短単位内部のアクセント境界	下記参照
PB	句末境界音調の継起	下記参照

**P**：「フク<pz>オカデ」（福岡で）のように、短単位の内部にポーズが生じた場合に付与している。

**<D および D** : 短単位が断片化した発話で、かつ、そこに韻律上の非流暢性が知覚される場合、その終端に”D”を付与している。また、韻律的不連続性が当該要素の前側境界に知覚される場合には、語断片の直前に”<D”を付与する。例えば「先日、ト、名古屋で」という発話には、「名古屋」を「東京」と言い誤りかけたところで発話を中断したために生じた語断片「ト」が含まれている。この発話の BI を( )に入れて示すと「センジツ(2<D)ト(D)ナゴヤ(1)デ(2)」ようになる。本ラベルの用法については、8.2 で別途説明しているの必ず参照すること。転記テキストにおける語断片の扱いについては transcription.pdf の(D)タグに関する記述を参照。

**<F および F** : ”<F”はフィラー区間の始端に付与している。フィラーに先行するアクセント句の BI とともに用いて、後続区間がフィラー区間であることを示す。一方”F”はフィラー区間の終端を示す(”F”は単独で使用し、後続するアクセント句との関係は示さない)。例えば「私がエー学生だったとき」という発話の BI を( )に囲って示すと、「ワタクシガ(2<F)エー(F)ガクセーダッタトキ」ようになる。フィラーについては 8.1 でも別途説明しているの必ず参照すること。

**F2** : このラベルは対話音声においてのみ使用される。対話音声には「ウン」「フーン」「シー」等の応答表現やあいづちが生じる。これらの表現は転記テキストでは(F)タグで囲われている。しかし(F)で囲われたテキストのなかには、発音が明瞭でなく、また、言語情報を保有しているかどうか不明らかなものがある。例えば相手の発話を聞きながら「アー」「シー」と母音や鼻音を発音しつづけているような場合である。この種の音声にラベリングを施すことは困難であるので、ラベリングの対象から除外した。BI 層にラベル”F2”が記入されている短単位区間には、トーン情報は何も記入されていない。ただし、転記テキストにおいて(F)で囲われている表現であっても、以下の条件を満たす発話は通常のアクセント句として認定している。1)は有核アクセント句、2)は無核アクセント句である。

- 1) 発話中にアクセント類似のピッチ下降が知覚される場合
- 2) アクセント類似のピッチ下降は知覚されないが、テキストが明確な言語情報を保有している場合 (ほぼテキストが「ハイ」の発話に限られる)

**W** : ひとつの短単位内部において、複数のアクセント核が生じるかアクセント核の後ろに句頭の上昇類似の上昇が生じた場合、その境界に付与している。下の例では // の位置に”W”が付与される (↑はアクセント核、「は」は句頭の上昇、そして | は短単位境界)。

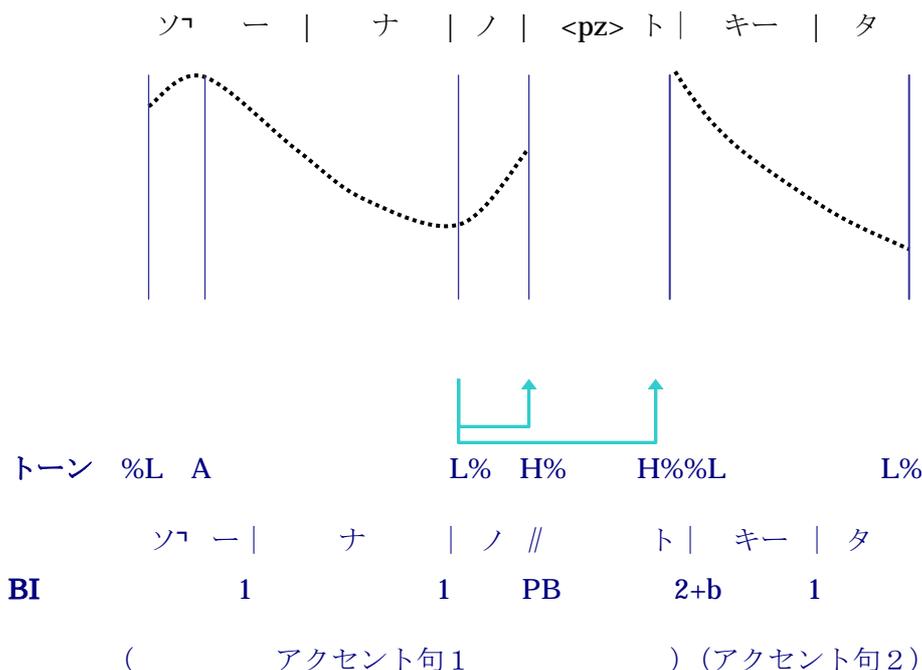
ファ<sup>↑</sup>ッショ // ナ<sup>↑</sup>ブル | ナ | カタ<sup>↑</sup> | ガ | オ<sup>↑</sup>ーク | テ (ファッションナブルな方が多くて)  
W            1   2            1   2            1

エ<sup>↑</sup>キ // マ<sup>↑</sup>エ | ワ | カワラ | ナ<sup>↑</sup>イ | シ (駅前是不変ないし)  
W            1   1            1            1

**PB** : PB は”Parasitic Boundary”の略である。句末境界音調は、原則として表 3 上部に示された 5 種類のどれかひとつが選択される。しかし自発音声ではふたつの境界音調が隣接して継起したと解釈できる発話が生じることがある。その場合、最初の境界音調末尾に BI ラベル”PB”を付与し、2 番

目の境界音調には通常の BI を付与している。

PB が生じる典型的環境は何らかの語句が引用された場合である。例えば『『そうなの』と訊いた』という発話は、下図のようなイントネーションで発音されることがある (<pz>はポーズ)。この場合、被引用句「そうなの」の末尾に L%H%が生じているが、後続する引用の助詞「と」も高いピッチで発音されているので、境界音調 H%を保有しているとみなせる。「そうなの」末尾の L%は、図下部の矢印のように、直後の”H%”と上昇句末境界音調を形成すると同時に、2番目の H%とも上昇句末境界音調を形成していると解釈できる(なお、ここで助詞「と」の Fo 最高値の時点に”H%”と”%L”が付与されているのは、上記 4.3 のケース 5) に該当するからである)。



**PB についての注意：**”PB”は注釈層における”AYOR”、”QQ”ラベルと共起することが多いので、本文書 7 節の記述も参照すること。

## 6 . プロミネンス層のラベリング

プロミネンス層は X-JToBI において新たに設けられた層であり、表 6 に示す 4 種類のラベルが記入される。これらのラベルは、句末境界音調の下位分類に関わる情報と無核アクセント句のプロミネンスに関わる情報を提供することによって、トーン層ないし BI 層のラベルの解釈を補助することを目的としている。

表 6 : X-JToBI のプロミネンスラベル

ラベル	用法	備考
PNLP	アクセント句次末モーラに生じた非語彙的な卓立	L%HL%と共起
FR	川上肇の「浮き上がり調」	L%H%と共起
HR	川上肇の「つり上げ調」	L%H%と共起
EUAP	強調された無核アクセント句	

**PNLP** : PNLN は”Penult Non-Lexical Prominence”の略である。アクセント句の最後から 2 番目のモーラ（以下では次末モーラと呼ぶ）に非語彙的な卓立（プロミネンス）がおかれる現象をさす。例えば「ディズニーランド」<sup>3</sup>という発話において「ラ」から「ン」にかけてのピッチ下降はアクセント核に起因しているが、「ド」における上昇・下降は、アクセント核とは無関係であるので”PNLP”として記述される。**X-JToBI** ではこの現象を句末境界音調”L%HL%”と同一のトーン列の時間的配置が変更されたものと解釈している。すなわち、通常は最終モーラ（「デ」）内部に生じる上昇下降調の句末境界音調が、発話末の 2 モーラ（「ドデ」）に広がって実現されたという解釈である。通常の L%HL%と PNLN 化した L%HL%とはトーン層のラベルを見るだけでは区別できない。そこで両者を区別するために PNLN 化した L%HL%については、プロミネンス層にラベル”PNLP”を付与している。ラベルの付与位置はアクセント句末の **BI** 位置である。

**PNLP に関する注意** : CSJ では、音声学的には PNLN と同じイントネーション形状を持つ発話であっても、次末モーラにアクセント核が存在する可能性がある場合は、PNLN としては解釈していない。例えば「名古屋まで」が「ナゴヤ」<sup>3</sup>と発音されていても PNLN とはしていない。助詞「マデ」が単独の発話ではアクセント核をもちうるからである。上記の発話は助詞がアクセント句として独立していると解釈して以下のようにラベリングする。

	ナゴヤ		「マデ」	
<b>BI</b>			<b>3</b>	<b>3</b>
トーン	%L A		L% %L A	L%

**FR** : FR は”Floating Rise”の略であり L%H%(上昇調 1)の変種を示す。通常の L%H%では Fo の上昇は専ら発話の最終音節（ないしモーラ）内部で生じるが、FR に分類される変種では、上昇の始端が、次末ないし次々末モーラ（を含む音節）に置かれる。これは、川上泰氏が 1963 に「浮き上がり調」と呼んだ句末境界音調現象である。上昇調 1 と浮き上がり調では、トーン層のラベルの種類はともに L%H%であり、ラベルの位置（時刻）だけが相違する。浮き上がり調の句末境界音調をもつアクセント句にはプロミネンス層にラベル **FR** を付与している。ラベルの位置はアクセント句末の **BI** ラベル位置である。なお、浮き上がり調は稀にはあるが上昇下降調(”L%HL%”)に生じることもあるので、川上の「浮き上がり調」とは完全には一致しない。

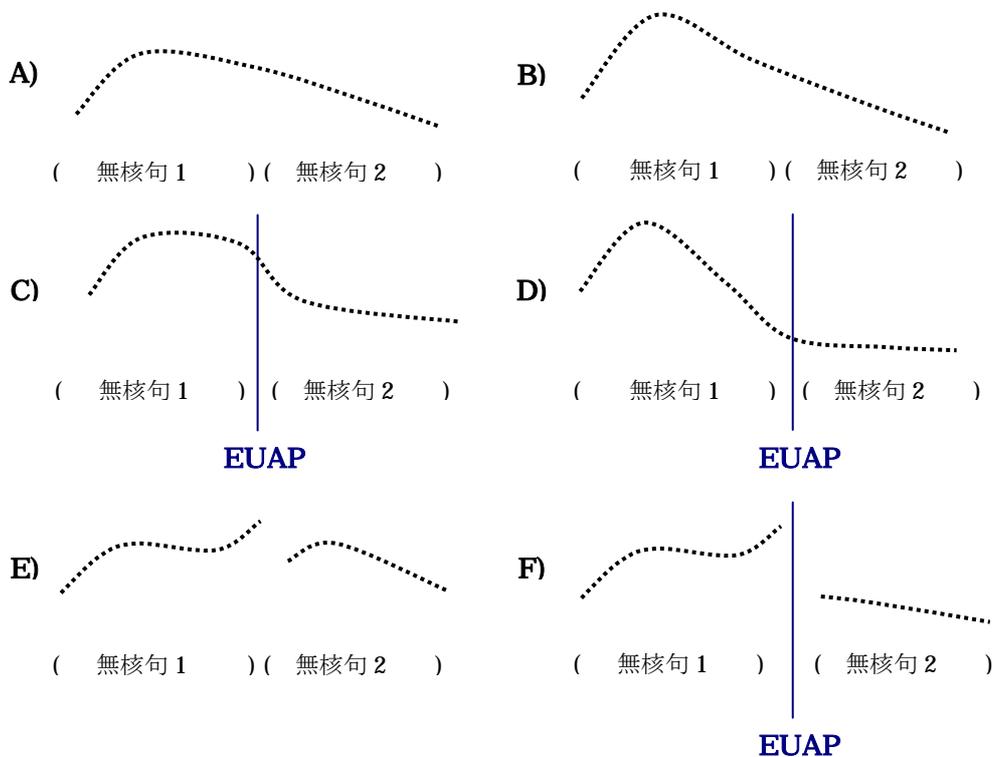
**HR** : ”Hooked Rise”の略。このラベルは、やはり川上泰氏が「つり上げ調」と呼んだ句末境界音調のために用意されたものである。川上は、この音調の特徴を、アクセント句末が「～です」「～ます」で終わる発話に生じ、末尾の 2 モーラが **des, mas** と 1 音節に発音され、その始端から急にピッチが上昇するタイプの句末境界音調と説明している。CSJ でも、ほぼこの環境に限ってつり上げ調が観察されている。つり上げ調音調のトーン層ラベルは”L%H%”のままであるが、プロミネンス層にラベル”HR”を記入している。ラベルの位置は句末 **BI** の終端である。

**EUAP** : EUAP は”Emphasized Unaccented Accentual Phrase”の略である。このラベルは強調を受けた無核アクセント句と後続アクセント句との境界に顕著なピッチレンジの差が観察される場合に付与している。ラベルの位置は強調を受けた無核アクセント句終端の **BI** にあわせる。**EUAP** が必要とされる理由は、東京方言ではアクセント句が有核であるか無核であるかによって強調の音声学的実現形が異なってくるからである。

東京方言の強調は、基本的にはアクセント句のピッチレンジの拡大によって実現される。有核句の場合アクセントに起因する **Fo** のピークが顕著に高まることによってピッチレンジの拡大が実現されるが、無核アクセント句は、**Fo** 曲線に顕著なピークが存在しないために強調を実現しにくいことが知られている。無核句における強調を実現するために、以下のような手段がもちいられることがある。

- 1) 強調の置かれた無核句のピッチレンジを拡大すると同時に後続するアクセント句のピッチレンジを狭める。
- 2) 無核句のピッチレンジはほとんど変化させないが、句末に **BPM** を置くと同時に後続するアクセント句のピッチレンジを狭める。

いずれの場合も、強調を受けた無核句から後続句にかけて、あたかも有核句におけるダウンステップが生じたかのような **Fo** 形状が観察されることになる。その場合強調を受けたとみなされる無核句終端の **BI** ラベルの位置に”**EUAP**”を付与している。



上の模式図中 A)-D)は、上記 1)の典型例として、ふたつの無核アクセント句が継起した場合の **Fo** 形状を示している。A)はどこにも強調がない場合（ふたつの句が融合している）、B)は無核句 1 に強調がおかれているが句境界に **EUAP** の対象となるダウンステップ類似の **Fo** 形状が観察されない

場合、C)と D)は EUAP が付与される場合である（先に 2.6 で示した X-JToBI ラベリングのサンプル図には上記 C) に該当する EUAP が生じている）。

E)と F)は上記 2)の典型例として無核句 1 の末尾に”L%H%”が生じた場合を示している。E)では無核句 2 のピッチレンジが狭められていないので EUAP の対象とはならない。

**EUAP に関する注意：**

- 1) 有核句に強調が置かれていることを示すラベルは存在しない。
- 2) EUAP の対象とした無核句の BI は 2、2+b、2+p、2+bp のいずれかであり 3 とはならない。

**7. 注釈層のラベリング**

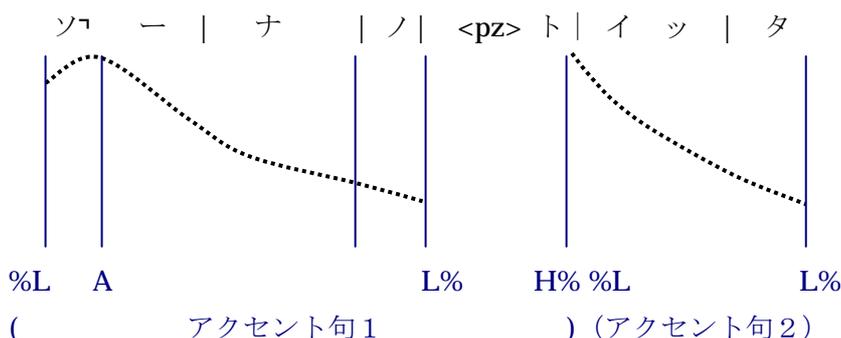
注釈層に用いられるラベルは表 7 に示した 3 種類である。これらはいずれも、X-JToBI による自発音声ラベリングにおいて十分には解決されていない問題の所在を示している。

表 7：X-JToBI の注釈層ラベル

ラベル	用法	備考
HBP	H で始まるアクセント句	High Beginning Phrase
QQ	いわゆる「半疑問」イントネーション	Quasi-question
AYOR	韻律の解釈に成案を得ていない箇所	At your own risk

**HBP：**”High Beginning Phrase”の略。アクセント句の冒頭が例外的に高いトーンで始まっている場合に当該アクセント句冒頭の短単位終端に付与している。「例外的に」とは以下の意味である。音韻論では東京方言のアクセント句は必ず低いトーン(%L)から始まると考えられている。実際にはアクセント句頭で Fo が最高値をとる発話が少なからず存在しているが、その多くは、下記のいずれかに該当する。

- A) アクセントが句頭にあるケース
- B) 直前のアクセント句が H%ないし LH%境界音調で終了しており、その高さが当該アクセント句頭まで持ち越されていると分析できるケース
- C) 直前の語が低いトーンで終わっているときに、高いトーンで始まる後続アクセント句冒頭の語が機能語であり、かつ、統語的には先行アクセント句末の語とともにアクセント句を形成すると解釈できるケース（下の模式図参照）

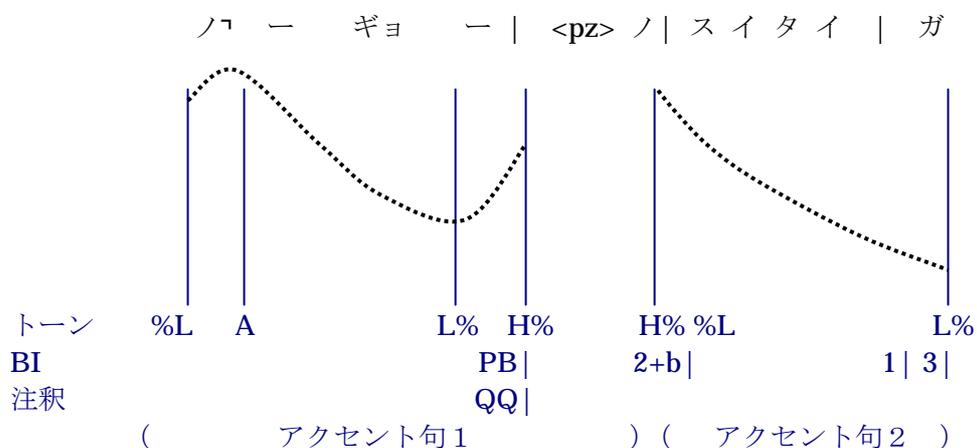


しかし、上記 A)-C)のような合理的説明が不可能なケースも、稀にはあるが生じる。例えば講演の冒頭や、長いポーズに続く発話が、いきなり「というわけで」、「ってことで」等で始まり、その際、「ト」や「ッテ」のピッチが顕著に高いことがある。これはちょうど上図の右半分が先行文脈なしに生じるケースである。このような発話は、句頭にアクセントがなく、また、先行するイントネーション上の文脈も存在しないので、発話冒頭に高いピッチとして実現される句頭境界音調が存在するとみなさざるをえない。CSJのイントネーションラベリングでは、上例のように、1)講演の冒頭ないしポーズの直後に生じる発話が、2)韻律的には高いピッチで始まっており、3)形態的には機能語（上例では助詞）で始まっている場合に限り、トーン層に境界音調”（H%）”を付与し、プロミネンス層にラベル”HBP”を付与している。

なお、ラベル”（H%）”が句頭境界音調でありながら”（%H）”ではないのは以下の理由による。上記 1)-3)の条件にかなう発話が生成されるのは、やはり何らかの先行発話が（たとえ表層的には実現されていなくても）話者に意識中に存在しているからであると考えられる。その潜在的発話の影響によって当該発話が高いピッチで始まったという解釈がラベル”（H%）”によって含意されている。

**QQ**：“Quasi-question”の略。俗に「半疑問」「半クエスチョン」と呼ばれるイントネーションが生じたことを示すラベルである。「半疑問」の認定基準は明示化されておらず、ラベリング作業者の直感に依存している。その位置に上昇句末境界音調（”H%”ないし”LH%”）が生じていることが必要条件である。ラベル”QQ”の付与位置は半疑問と判断される句末境界音調が生じたアクセント句の終端である。？で半疑問の音調、|でアクセント句境界を示せば、「スマート？と|言うのかな」の場合助詞の「と」の終端である。

QQ はしばしば PB と共に生じる。アクセント区内部に（半疑問の）BPM が生じ、後続する機能語が高いピッチを有する場合である。例として下図に「農業？<pz>の|衰退が」（<pz>はポーズ）の模式図を示す。この種の発話は下図のようにラベリングするが、PB と共起する場合にかぎり、ラベル”QQ”はラベル”PB”と同じ BI 位置に付与している。



**AYOR** : "At your own risk"の略。いわゆる引用の発話では、**X-JToBI** によるラベリングに困難が生じることがある。典型的には「～と言う (ことです)」「～って (ことです)」などの発話形式である。引用表現には、先に説明した **PB** と **QQ** が多く生じるが、その他に「という」「ってこと」などの引用句部分が低く (稀には高く) 平らなピッチで (あたかもフィラーのように) 発音されることもある。これら、引用にかかわるイントネーション特徴の解釈には現在でも十分な成案を得ていないので、イントネーションラベルの利用目的によっては、引用の発話を分析対象から除外する方が賢明な場合もある。そこで、引用とみなされる発話に"**AYOR**"ラベルを付与することにした。なお、ここで「引用の発話」と呼んでいるのは、文法的に正確な用法ではなく「典型的には引用の表現に生じる韻律上の特徴を有する発話」という意味である。上に挙げた例のほかに「～と聞いた (んです)」「～みたいな」などの発話にも"**AYOR**"が付与されていることがある。"**AYOR**"ラベルは対象となる発話の最後の **BI** 位置に付与している。

## 8 . 非流暢性の BI ラベリングについての補遺

自発音声を朗読音声から区別する顕著な特徴のひとつが非流暢性(**disfluency**)である。非流暢性については、ここまでの説明で折に触れて言及しているが、もっとも頻繁に生じる非流暢性要素であるフィラーと語断片に関わる **BI** の処理についてまとめて説明する。

### 8 . 1 フィラー

自発音声を特徴づける非流暢性(**disfluency**)現象のうち、もっとも頻度が高いのは、「ア(ー)」「エ(ー)」「アノ(ー)」「アノデスネ」等、言い淀みによって生じるフィラー(**filled pause**)である。**CSJ** の転記作業では、フィラーないし感情表出系感動詞 (アー、オー、ホー、ギャ、フーン等) を(**F**)タグで囲って表記している (対話では応答表現類も (**F**) で囲われている。**transcription.pdf** 参照)。**X-JToBI** では、転記テキストにおいて(**F**)タグを付与された発話のうち、以下の基準によって韻律的なフィラーを認定し、通常のアクセント句とは異なる方法でラベリングしている。

- 1) フィラー候補の韻律的長さが1モーラの場合は、すべて韻律的フィラーとみなす。
- 2) 長さが2モーラ以上のフィラー候補については、**A)**句頭の上昇が認められず、かつ、**B)**発話中にアクセント類似の局所的なピッチ下降が認められないものだけを韻律的フィラーとみなす。

上記の基準にかなう発話は、全体を通してピッチが平坦であり、これが韻律的フィラーの必要条件となっている。なお、アクセント句末で最終モーラの母音が顕著に延長される発音は、機能的にはフィラーの役目を果たしていると考えられるが、韻律的フィラーとは認定していない。(ただし転記テキストでは"<H>"タグが付与されており、**X-JToBI** ラベルの単語層にもそのまま反映されている **transcription.pdf** および **segment.pdf** 参照)。

韻律的フィラーはその前後を特殊な **BI** で区切る。フィラーの始端にはラベル"<F"を、終端にはラベル"F"を付与している。このうち始端ラベルは、**2<F**, **3<F**, **2+bp<F** のように、フィラー区間の直前・直後に出現する通常のアクセント句間について判定した **BI** 値と共に記入されている。例えば以下の例は、「サキニ」と「モーシマシタヨーニ」の間の **BI** 関係が、**3**であることを示している。

サキニ | エー | モーシ | マシ | タ | ヨー | ニ

3<F    F            1        1 2        1

フィラーのピッチは、全体を通して高いか低いかを判定し、それぞれ”FH”、”FL”というトーンラベルで表現する。このラベルは対象とするフィラー区間の内部に位置していればよい。複数のフィラーが継起している場合は、下の例のように、各フィラーの終端に BI ラベル”F”を記入し、各フィラーの高さを”FH”ないし”FL”で判定する。この場合も、数字”3”は「サキニ」と「モーシマシタヨーニ」の間の関係を示している。

	サキニ		エー		マー		モーシ		マシ		タ		ヨー		ニ
BI	3<F		F		F		1		1		2		1		
トーン			FL		FL										

## 8.2 語断片

語断片(word fragment)とは、「ワ、私が」における「ワ」のように、語(短単位)を発音している途中で発話が中断されることによって生じた語の断片のことである。自発音声には語断片が頻繁に現れ、これに対して消極的な立場をとると、場合によってはほとんどラベルが付与できなくなることがある。そこで、語断片が生じていてもできるだけ多くの情報を記録するよう努めた。

語断片の認定にあたっては、転記テキストにおいて(D)タグが付与されている発話を語断片候補とした。また第三刷ではタグ<FV>、”(?)”で表現される音声区間も語断片候補とした。ただし転記テキスト上の語断片と韻律上の語断片とは必ずしも一致しない。統語的・文脈的には断片化している要素であっても、そのイントネーションは滑らかに連続していることがある。その場合、韻律上の語断片を認定する必要はない。

CSJ のラベリングでは、語断片候補の前側境界あるいは後側境界のいずれか一方あるいは両方に韻律的不連続性が知覚される場合、BI 層に語断片ラベル”D”を付与することにした。また語断片の前側境界に韻律的不連続性が認められる場合、その前側境界に”<D”を付与している。ここで前側境界に韻律的不連続性が認められる場合とは、具体的には当該要素に句頭境界音調%Lが見られない事例に、後側境界に韻律的不連続が感じられる場合とは、当該要素に句末境界音調 L%が見られない事例にそれぞれ相当する。韻律的語断片が連続している場合、韻律的語断片の境界には”D”を付与するのみであり、”D<D”とはしない。

ちなみに転記テキストでは、(D)タグの他に(D2)タグも用いられている。これは、「話者ごと(D2)の)に」、「言語(D2)に)では」のように、機能語(助詞・助動詞・接頭辞・接尾辞)が機能語で言い直された場合に用いられるタグである(transcription.pdf 参照)。転記テキストでは上例のように先に出現した機能語に対して(D2)タグを付与するが、韻律的にはむしろ転記テキストで(D2)タグを付与された機能語とそれに後続する機能語との間に不連続性が生じることが多い。そこで、イントネーションラベリングでは、後続要素(上例では「に」と「では」)を語断片候補としている。

なお CSJ 第三刷(2011 年)では、ラベル”D”の付与基準を「後側境界に韻律的不連続が感じられる場合」から「語断片候補の前側境界あるいは後側境界のいずれか一方あるいは両方に韻律的不連続性が知覚される場合」に変更した。語断片の扱いについては『日本語話し言葉コーパスの構築』(本文書末尾参照)7.4 節にその詳細が記されているが、上記変更は反映されていないため注意が必要である。

ラベリングの例を次頁に示す。

**語断片に関する注意：**

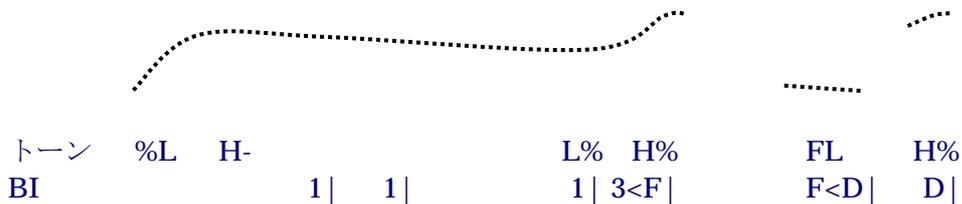
トーン層のラベリングにおいては、語断片に対しても極力トーンラベルを付与している。そのため語断片の周辺では、トーンの文法に従わないラベル系列が現れることがある（例えばアクセント句頭の境界音調が存在しても句末の境界音調が存在しない）。

**(1) 前側境界にのみ韻律的不連続性が見られる場合**

（助詞「ワ」を、フィラーを挟んで「ノ」に言い直し）

エ キ メー | ト | テン メー | ワ | <pz> | エー | ノ |

↓ %L の欠落

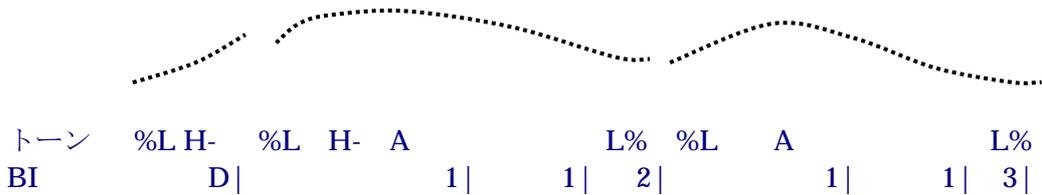


**(2) 後側境界にのみ韻律的不連続性が見られる場合**

（冒頭の「ムイ」が語断片）

ム イ | ユー セー | オン | ト | ム セー | オン | ガ |

↓ L% の欠落

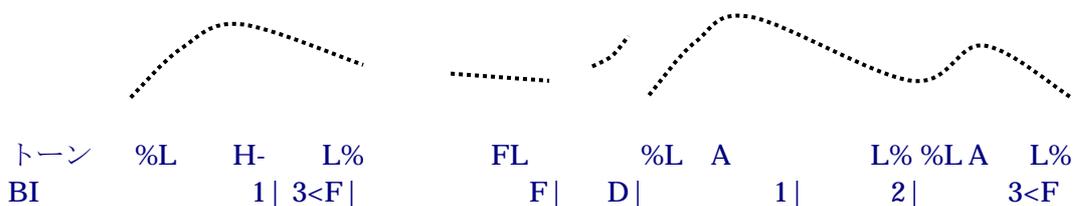


**(3) 両方の境界に韻律的不連続性が見られる場合**

（フィラー後の語断片「ン」）

ギャク | ニ | <pz> | エー | ン | ワカ イ | モノ | デ | モ |

↓ ↓ %L, L% の欠落



## 9 . ラベルファイルの構造と形式

X-JToBI の情報は、各層ごとに独立したテキストファイルとして提供される。ラベリングの対象となった講演の ID が A01M0001 である場合、X-JToBI 関係ファイルには以下の拡張子が与えられ、講演 ID 名と同じ名称のディレクトリ (/A01M0001) 直下のサブディレクトリ PLABEL に格納される (overview.pdf の付録 3 参照)。なお、対話音声については主話者 (overview.pdf の付録 1 において「元講演」をおこなっている話者) についてのみ X-JToBI ラベルを提供している。

A01M0001.seg	分節音層のラベルファイル (segment.pdf 参照)
A01M0001.tone	トーン層のラベルファイル
A01M0001.break	BI 層のラベルファイル
A01M0001.word	単語層のラベルファイル
A01M0001.prm	プロミネンス層のラベルファイル
A01M0001.misc	注釈層のラベルファイル

テキストファイルの形式は Entropic 社の waves 形式である。以下に例を示すように、冒頭にヘッダー部があり、その後にラベル部が続く。ラベル部のレコードは、「時間情報(sp)フォント情報(sp)ラベル情報」のように 3 個のフィールドがスペースで区切られている。フォント情報は Entropic 社のソフトウェアのための情報であり、X-JToBI にとっては意味のない情報である。

```
signal A01M0001
type 0
comment created using xlabel Mon Nov 17 03:53:55 2003
comment created using xlabel Tue Apr 1 07:46:41 2003
color 115
font *-times-medium-r-*-20-*-*-*-*-*
separator ;
nfields 1
#
0.413481 115 FL
0.599277 115 %L
0.701934 115 H-
1.906648 115 A
2.403810 115 L%
2.586214 115 H%
2.827159 115 %L
3.750207 115 A
4.080238 115 L%
以下省略
```

ヘッダー部

ラベル部

segment.pdf の 4.1 で説明しているように、waves 形式のファイルは wavesurfer 等のフリーウェアで直接利用することができる。ただし数十分におよぶ音声ファイルを、メモリーが充分でない PC で利用しようとする、ラベルの総数が過大であるために、ソフトウェアの動作が緩慢化したり異常終了したりすることがある。そのような場合、もし必要がないのであれば、分節音層のラベルの表示を中止することによって症状を軽減できることがある。

なお第二刷 (2008 年) から、音声分析ソフトウェア Praat で利用可能な TextGrid 形式に変換した音声ラベルデータの提供を開始した。TextGrid 形式では、上述の六つの層の情報がすべて TextGrid という拡張子をもつ 1 個のファイルに集約される。他の X-JToBI 関係ファイルと同様、講演 ID 名と同じ名称のディレクトリ直下のサブディレクトリ PLABEL に格納される。

## 謝辞

X-JToBI を設計する過程で、J\_ToBI の設計者である Jennifer Venditti 氏から有益なコメントを頂戴しました。また 2000 年に国立国語研究所で開催した J\_ToBI ラベリングワークショップ参加者の皆様からも貴重なデータとコメントをいただきました。記して感謝いたします。

## 文献\*

川上 夔(1963). 「文末などの上昇調」国語研究, 16, p.25-46.

菊池 英明・前川 喜久雄・五十嵐 陽介・米山 聖子・藤本 雅子(2003). 「日本語話し言葉コーパスの音声ラベリング」音声研究, 7 (3), pp.16-26.

前川 喜久雄・菊池 英明・五十嵐 陽介(2001). 「X-JToBI : 自発音声の韻律ラベリングスキーム」電子情報通信学会技術報告(NLC2001-71, SP2001-106), pp. 25-30.

Kikuo Maekawa, Hideaki Kikuchi, Yosuke Igarashi, and Jennifer Venditti (2002). “X-JToBI: An extended J\_ToBI for spontaneous speech.” *Proceedings of the 7th International Conference on Spoken Language Processing (ICSLP2002)*, pp. 1545-1548, Denver, USA.

Jennifer Venditti (1997). “Japanese ToBI Labelling Guidelines.” In K. Ainsworth-Darnell and M. D’Imperio (Eds.) *Papers from the Linguistics Laboratory. Ohio State University Working Papers in Linguistics*, Vol. 50: 127-162. [First distributed in 1995 at a web document].

\*ここに示した X-JToBI 関連文献は設計の過程で発表したものであり、その内容は最新の仕様を反映していない。X-JToBI についての最終的典拠となるのは本文書および以下の文献である。ただし 8.2 に記したように CSJ 第三刷における変更については以下の文献に反映されていない。

五十嵐 陽介・菊池 英明・前川 喜久雄(2006). 「第 7 章 韻律情報」『日本語話し言葉コーパスの構築』(国立国語研究所報告書 124), pp.347-453, 2006.