

# 『日本語話し言葉コーパス』における韻律単位の認定基準について

小磯 花絵 (国立国語研究所理論・構造研究系)<sup>†</sup>

前川 喜久雄 (国語研究所言語資源研究系)

五十嵐 陽介 (広島大学)

## Criteria for Intonational Unit Identification: The Case of *the Corpus of Spontaneous Japanese*

Hanae Koiso (Dept. Linguistic Theory and Structure, NINJAL)

Kikuo Maekawa (Dept. Corpus Studies, NINJAL)

Yosuke Igarashi (Hiroshima University)

### 1. はじめに

『日本語話し言葉コーパス』(*Corpus of Spontaneous Japanese, CSJ*)は、1999年から5年間かけ、国立国語研究所・情報通信研究機構(旧通信総合研究所)・東京工業大学が共同で開発した、約650時間の日本語自発音声からなるデータベースである(前川2004,2006)。2004年に公開を開始して以降、幅広い領域で利用されており、データを修正、追加しながら、第2刷(2008年)、第3刷(2011年)と版を重ねている。

現在、CSJのうち多種多様な研究用付加情報が付されたコアと呼ばれるデータ範囲(約45時間、50万語)を対象に、各種情報を相互に関連付けて表現したRDB(小磯ほか2012)を構築しており、近日中の公開を目指している。CSJ-RDB版は、原則としてCSJ第3刷に含まれるXML文書に表現されている単位や各種情報を反映したものであるが、新たにアクセント句とイントネーション句という2種類の韻律単位の明示的に設けることとした。

CSJが採用している韻律ラベリング体系X-JToBI(Maekawa et al. 2002; 五十嵐ほか2006)およびその前身であるJ-ToBI(Venditti 1995, 2005)では、アクセント句やイントネーション句は単位としては明示的に表現されず、韻律境界の切れ目の強さを表す情報(Break Index, 以下BI情報)によって間接的に表される。概ねBI=1は語境界、BI=2類はアクセント句\*1の境界(同時に語境界)、BI=3はイントネーション句\*2の境界(同時に語とアクセント句の境界)に

---

<sup>†</sup> koiso@ninjal.ac.jp

\*1 アクセント句は、句頭第1モーラから第2モーラ付近にかけてのF0の上昇と句末への緩やかな下降を有し、かつアクセント核による下降を最大ひとつ持ちうる単位と定義される。なおX-JToBIでは、アクセント句末にポーズや上昇調などの複合境界音調(BPM)が出現する場合、BI=2とBI=3の中間値として、BI=2+p, BI=2+b, BI=2+bpを新規に導入している。

\*2 イントネーション句は、アクセント句の上位階層に位置し、アクセント句のピッチレンジを指定する単位と定義される。先行アクセント句と比較してピッチレンジが拡大した場合、そこにイントネーション句境界があるとみなされる。アクセント核が引き起こす後続アクセント句のピッチレンジの縮小効果は、イントネーション句の終端で阻止されることになる。なお、J-ToBIが立脚するPierrehumbert & Beckman (1988)の音韻理論では、アクセント句より階層的に上位の単位として「中間句(intermediate phrase)」と「発話(utterance)」の二つが認められているが、J-ToBIにおける「イントネーション句」はこれらを融合させた単位である。

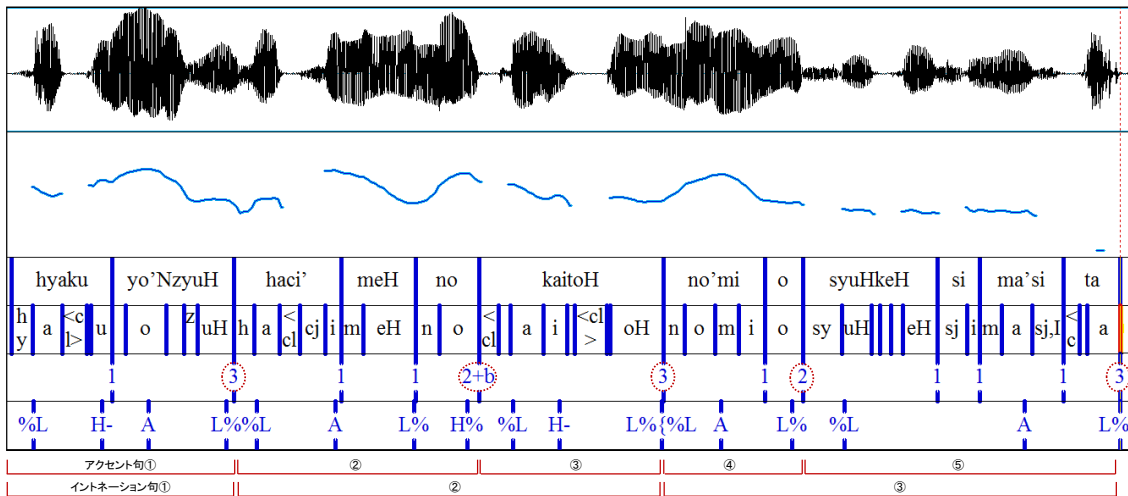


図1 非流暢現象が存在しない発話例「148名の回答のみを集計しました」

相当することから、アクセント句は両端を BI=2 か BI=3 で区切られる範囲として、イントネーション句は両端を BI=3 で区切られる範囲として、認定することができる。図1に、X-JToBIに基づくトーンとBI情報、およびそれに基づき同定したアクセント句とイントネーション句の範囲の例を示す。

このように流暢に発話されている部分では、BI情報から単純に韻律単位を同定することができる。しかしCSJが対象とするような自発性の高い発話には、図2にあるように、「アノ」「エット」「シー」などのフィラーや、「わた私が」のような言いよどみが頻出する。J-ToBIが立脚する音韻理論 (Pierrehumbert & Beckman 1988) は朗読音声に基づき構築されたものであるため、このような自発性の高い話し言葉に特徴的な現象を十分に記述できないという問題があった。X-JToBI (eXtended J-ToBI) はこの種の現象への積極的な対応が検討され様々な拡張がなされたが、その結果、体系が複雑化し、アクセント句・イントネーション句の認定が単純には行かなくなった。例えば図2にあるフィラーを一つとっても、それを独立したアクセント句とみなすべきか、それとも先行あるいは後続するアクセント句の一部とみなすべきかはすぐには決まらない。

CSJの第1刷、第2刷では、各種情報を統合したXML文書の中で、単語(短単位)の属性としてアクセント句のID情報 (APID) が表現されているが、時間的な制約もあり、非流暢現象を含めてBI情報からアクセント句を汲み上げる方法を十分に検討することができなかった。CSJ-RDB版を作成するにあたり、これら非流暢現象の扱いを中心にアクセント句・イントネーション句の認定基準を改めて検討し、またX-JToBIの認定基準についても一部変更を施すこととなった。そこで本稿では、検討の対象となったフィラー・言いよどみの扱いを中心に韻律単位の認定基準について報告する。

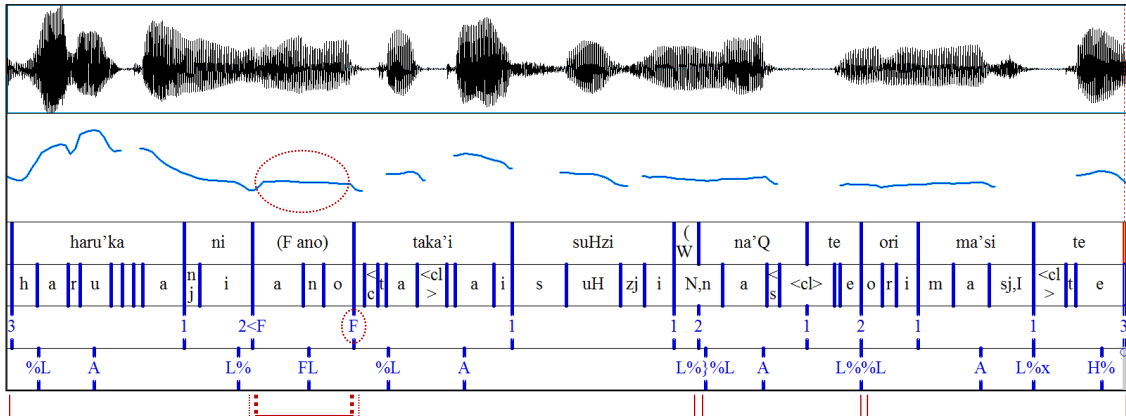


図2 非流畅現象（フィラー）が存在する発話例「遥かにあの高い数字になっておりまして」

## 2. 韻律単位の認定基準

### 2.1 フィラーの扱い

■対象とするフィラー CSJの転記テキストでは、語彙的・機能的な観点からフィラーが認定されるのに対し（小磯ほか2006）、X-JToBIでは韻律的な観点からフィラーが定義される（五十嵐ほか2006）。原則として、転記基準でフィラーと認定されるもののうち、フィラーの主観的長さが1モーラのもの、あるいは、句頭の上昇が認められずかつアクセント類似の局所的なピッチの下降が認められないものだけが、X-JToBIではフィラーとみなされる。また接続詞の「で」など、転記では機能的観点からフィラーとは認定されないものであっても、長さが1モーラで韻律的にフィラーと似た形で現れるものは、X-JToBIではフィラーとみなされる。

■フィラーに関わるBI値の認定基準の概要 フィラーは次の方針でその前後のBI値が認定されている。これはフィラーが無音区間と類似した機能を持つとし、フィラーを透過的に（存在しないものとして）扱うという立場で定めたものである\*3。

1. フィラーに先行するアクセント句（図2の例では「遥かに」）のBI値は、フィラーを透過してそれに後続するアクセント句（「高い数字に」）との関係でBI=2か3を判断した上で、フィラーの始端を示す（フィラーが後続することを示す）「<F」を付与し、BI=2<F, 3<Fとする。
2. フィラーの終端には「F」を単独で付与する。透過要素のため後続するアクセント句との韻律的接続関係は特定しない。

### ■フィラーに関わる韻律単位の認定方針

アクセント句：このようにフィラーは透過要素として扱われるため、現在付与されているBI値から、フィラーを先行あるいは後続するアクセント句の一部に含めるべきか、それとも独立したアクセント句と認定すべきかは決まらない。そこで、無理にどちらかのアクセント句に含

\*3 そのような仮定の実証的根拠としては、例えば前川（2012）がある。

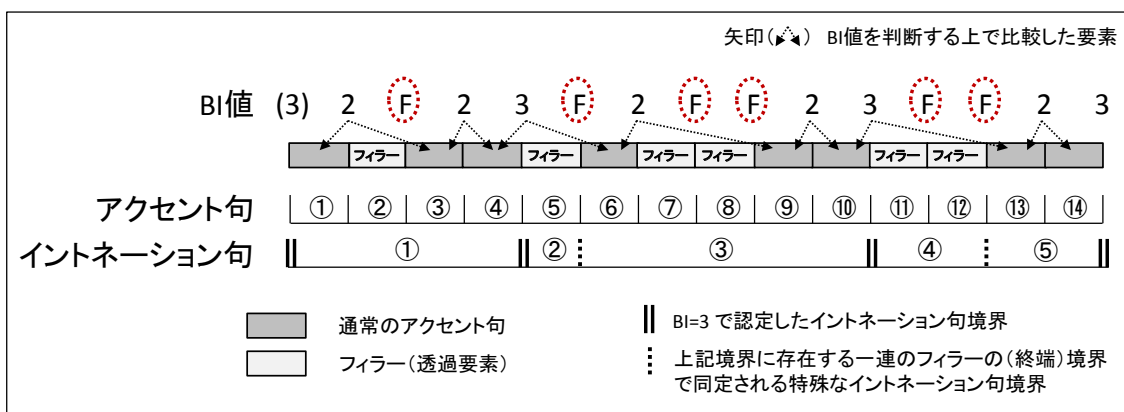


図3 フィラーが関わる場合のアクセント句・イントネーション句の認定例

めるのではなく、独立した特殊なアクセント句（フィルター句）と認定する。連続して出現するフィルターはそれぞれ別のアクセント句とする。

**イントネーション句：** フィラーを透過要素とみなして BI=3 の位置でイントネーション句を認定した上で、句境界に存在するフィルター（連鎖する場合はその全体）は無理に前後のイントネーション句には含めず独立した特殊なイントネーション句として、またイントネーション句の内部に存在するフィルターは当該イントネーション句に含める形で認定する。

以上の方針で認定したアクセント句・イントネーション句の認定例（模式図）を図3に示す。

なお、対話データに限定して、「ウン」「フーン」「シー」等の応答表現やあいづちの一部に対し BI 値を F2 としているが、韻律単位の認定基準については F と同じとする。

## 2.2 言いよどみの扱い

■**対象とする言いよどみ** 韻律ラベリングにおいて問題となるのは、「わた 私が」のように、言いよどみに伴い語の断片などが生じ、かつその境界（当該要素の前側あるいは後側境界、ないしその両方）に何らかの韻律上の不連続性が知覚される場合である。

言いよどみ部の後側境界に韻律的不連続が感じられる場合とは、アクセント句の終端を特徴付ける句末への緩やかな下降が見られない事例に相当する。「すじ 推定したものです」や「典型的もんだ 典型的事例は」のように、言いよどみに伴い発話を途中で言いやめたケースに典型的に見られる（図4（A）参照）。一方、言いよどみ部の前側境界に韻律的不連続性が認められる場合とは、アクセント句の始端を特徴付ける句頭第1モーラから第2モーラ付近にかけての F0 の上昇が見られない事例に相当する。「聞き分け易さが の 評価が」のように機能語を言い直すケースに典型的に見られる\*4（図4（B）参照）。「相互関係を い 示します」のような短い言いよどみの場合には、句頭の上昇も句末の下降も観察されないことがあり、その場合は言いよどみ部の両側に韻律的不連続性があるとみなされる（図4（C）参照）。なお語彙的・統

\*4 韻律的に不連続性が観察されるのは、言い直された要素（上記例では「が」）ではなく、後続の言い直した要素（下線部の「の」）である。なお、「前足をから」のように、2モーラ語以上の機能語で言い直した場合、言いよどみではなく通常のアクセント句として扱われる。

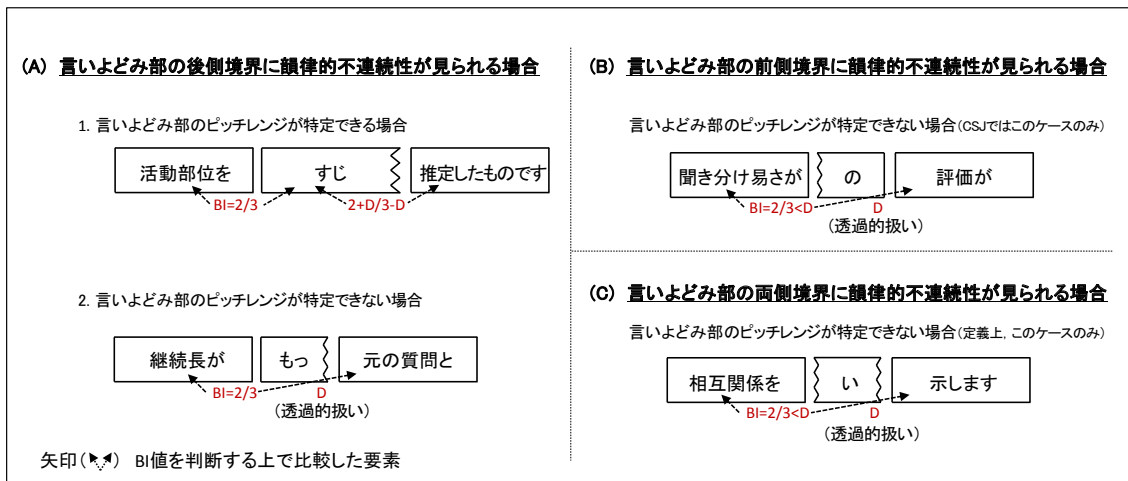


図4 言いよどみのタイプとBI値

合的な言いよどみであっても、韻律上の不連続性が全く認められない場合には、通常の語の場合と同様に扱われる。

■ 言いよどみに関わるBI値の認定基準の概要 韻律的不連続性の観察される言いよどみについては、次の方針で前後のBI値が認定される\*5。

1. 言いよどみ部に先行するアクセント句のBI値は、言いよどみ部のピッチレンジが特定できる場合\*6、言いよどみ部と先行アクセント句との比較でBI=2か3を判断するのに対し(図4(A1))、言いよどみ部のピッチレンジが特定できない場合は、フィラーと同様に言いよどみ部を透過要素とみなし、言いよどみ部の前後のアクセント句との比較でBI=2か3を判断する(図4(A2)(B)(C))。
2. 言いよどみ部の前側境界に韻律的不連続性が認められる場合、先行アクセント句のBI値に、始端の韻律的不連続性の後続を示す「<D」を付与する(図4(B)(C))。
3. 言いよどみ部のピッチレンジが特定できない場合、透過要素とみなし、言いよどみ部と後続するアクセント句との韻律的接続関係は特定せず、言いよどみ部の終端を示すBI=Dのみを付与する。一方、言いよどみ部のピッチレンジが特定できる場合、言いよどみ部と後続アクセント句とを比較し、イントネーション句境界が認められる場合はBI=3-Dを、認められない場合はBI=2+Dを付与する\*7。

\*5 第2刷までの基準では、図4(B)に示す始端境界にのみ韻律的不連続性が見られる場合、「聞き分け易さが(BI=<D)の(BI=2)評価が(BI=3)」のように、語断片の先行要素にBI=<Dを、言い直し部の終端にBI=2,3を付与していた。この方針は、言いよどみ部を先行アクセント句の一部とみなす立場と言える。しかし実例を観察すると、言いよどみが複数繰り返されたりフィラーが連鎖したりと、先行要素との間に強い韻律的不連続性を感じるが多い。そこで第3刷では、図4(B)に示すように、言いよどみ部の先行要素にBI=2,3を、言いよどみ部の終端にBI=Dを付与するよう、仕様を変更した。これにより、後述のアクセント句認定基準に従うと、言いよどみ部の前までが一つのアクセント句と認定され、言いよどみ部は独立した単位とみなされる。

\*6 言いよどみ部のピッチレンジが特定できる場合とは、言いよどみ部に句頭上昇か句末下降のいずれかが観察され、かつ、句頭音調(トーン記号H-)かアクセント核(トーン記号A)のいずれかが見られる場合である。トーンの詳細は五十嵐ほか(2006)を参照のこと。

\*7 第3刷までの基準では、言いよどみ部の終端には「D」を単独で付与し、言いよどみ部と後続するアクセント句と

## ■ 言いよどみに関わる韻律単位の認定方針

**アクセント句**：ピッチレンジが特定できる言いよどみ部は、不完全な単位ながらも通常のアクセント句と同様の手続きで BI=2 か 3 の判断がなされるため、独立したアクセント句とする。一方、ピッチレンジの特定できない言いよどみ部は透過要素とみなすため、フィラーと同じ扱いとし、独立した特殊なアクセント句とする。結果、ピッチレンジの特定の有無に関わらず、言いよどみ部は独立したアクセント句と認定される。

**イントネーション句**：ピッチレンジが特定できない言いよどみ部はフィラーと同じ扱いとする。つまり、透過要素とみなしてイントネーション句を認定した上で、その境界に存在する言いよどみ部（連鎖する場合はその全体）は独立した特殊なイントネーション句として、イントネーション句の内部に存在する言いよどみ部は当該イントネーション句に含める形で認定する。一方、ピッチレンジが特定できる言いよどみ部は、通常のアクセント句と同様の基準に従い、BI=3, 3-D の場合にイントネーション句境界があるとみなす。

以上の方針で認定したアクセント句・イントネーション句の認定例（模式図）を図 5 に示す。



図 5 言いよどみが関わる場合のアクセント句・イントネーション句の認定例

## 2.3 アクセント句・イントネーション句の認定基準

以上、フィラーと言いよどみごとに、問題の所在と認定方針を示した。これらの検討をふまえ、一般化したした形で韻律単位の認定基準をまとめる。

アクセント句の認定基準は極めて単純に、次のようにまとめることができる。

**アクセント句** BI 値が 2, 3, F, F2, D, 2+D, 3-D のいずれかで区切られる単位

の韻律的接続関係は一切認定されていなかった。しかしこの基準では、ピッチレンジが特定できる言いよどみの場合（図 4 (A1)）、後続アクセント句との韻律的接続関係は一切認定されないことになる。これではイントネーション句の認定に問題が生じるため、(A1) の事例を対象に、後続アクセント句とのピッチレンジを比較し、BI=2 か 3 の判断を人手で行うこととした。なお、タグ BI=D+2, D-3 は暫定的なものであり、今後、韻律ラベリングの最新版を公開する際に変更する可能性がある。

イントネーション句の基準をまとめる前に、簡単に整理しておこう。ピッチレンジの特定できる言いよどみ (BI=2+D, 3-D) は、通常のアクセント句 (BI=2, 3) と同じ扱いとなる。一方、フィラー (BI=F, F2) とピッチレンジの特定できない言いよどみ (BI=D) は、いずれも透過要素として扱われる。以上をふまえ、イントネーション句の認定基準を以下のようにまとめる。

### イントネーション句

- (1) BI=3, 3-D の位置にイントネーション句の境界があるとみなす
- (2) (1) で認定されたイントネーション句の境界に存在する透過要素 (BI=F, F2, D の要素, 連鎖する場合はその全体) は、前後とは独立したイントネーション句とする  
補足 イントネーション句の内部に存在する透過要素は当該イントネーション句に含める

図 6 に、この基準に従い認定したイントネーション句の例を、フィラーと言いよどみが連続するケースを含めて示す。なおアクセント句は各セルに相当するため省略する。

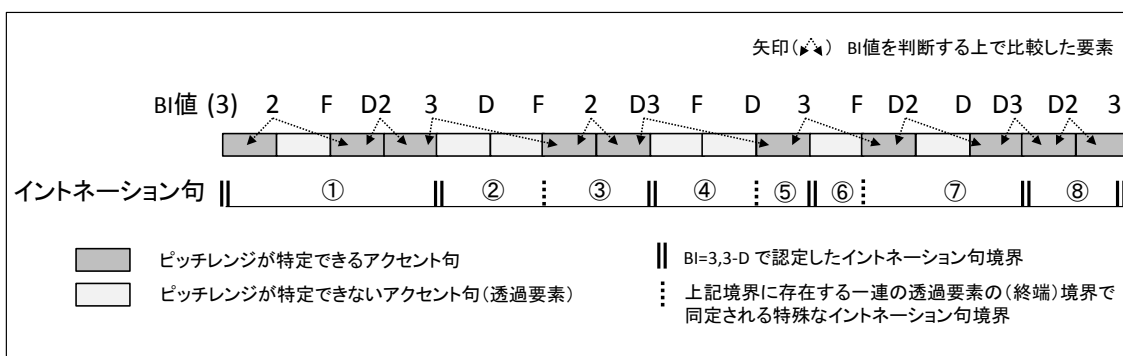


図 6 イントネーション句の認定例

### 3. おわりに

非流暢現象の音韻論的なトーン指定の仕組みや韻律構造に与える影響については、CSJ を対象とした前川 (2012) などの研究でようやく検討が始まった状況である。そのため、純粋に理論的な観点から認定基準を定めることが難しいケースも少なからず存在する。しかしこのような状況だからこそ、アクセント句・イントネーション句を含む CSJ-RDB 版を公開することに意味があると言えよう。今後、非流暢現象の韻律特徴についての研究が進展するにつれて、認定基準もより堅固な基礎に立つことができるようになることが期待される。

### 参考文献

- 五十嵐陽介・菊池英明・前川喜久雄 (2006) 「韻律情報」『日本語話し言葉コーパスの構築法』(国立国語研究所報告 124), 347-453.
- 菊池英明・塚原渉 (2006) 「XML 文書」『日本語話し言葉コーパスの構築法』(国立国語研究所報告 124), pp. 455-526.

- 小磯花絵・西川賢哉・間淵洋子 (2006) 「転記テキスト」『日本語話し言葉コーパスの構築法』(国立国語研究所報告 124), pp. 23–132.
- 小磯花絵・伝康晴・前川喜久雄 (2012) 「『日本語話し言葉コーパス』RDBの構築」『第1回コーパス日本語学ワークショップ予稿集』, pp. 393–400.
- 前川喜久雄 (2004) 「『日本語話し言葉コーパス』の概要」『日本語科学』, 15, pp. 111–133.
- 前川喜久雄 (2006) 「概説」『日本語話し言葉コーパスの構築法』(国立国語研究所報告 124), pp. 1–21.
- 前川喜久雄 (2012) 「自発音声中のフィラーの特性に関する予備的分析:位置と高さの分析」『第26回日本音声学会全国大会予稿集』, pp.115–120.
- Maekawa, Kikuo, Hideaki Kikuchi, Yosuke Igarashi & 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.
- Pierrehumbert, Janet & Mary Beckman (1988) *Japanese tone structure*, Cambridge: The MIT Press.
- Venditti, Jennifer (1997) Japanese ToBI Labelling Guidelines, In K. Ainsworth-Darnell M. D’Imperio (eds.) *Papers from the Linguistics Laboratory, Ohio State University Working Papers in Linguistics* 50, pp.127–162. (First distributed in 1995 at a web document)