

日本語音声コーパスにおける促音・非促音の判別

天野成昭[†] (愛知淑徳大学)

山川仁子 (愛知淑徳大学)

近藤真理子 (早稲田大学)

Discrimination between Single and Geminate Stops in Japanese Speech Corpus

Shigeaki Amano (Aichi Shukutoku University)

Kimiko Yamakawa (Aichi Shukutoku University)

Mariko Kondo (Waseda University)

1. はじめに

日本語において「っ」で表記される促音の主な音響的特徴は、音声波形上の無音区間の時間長にある。例えば促音を含む「いった」では無音区間の時間が長く、非促音の「いた」では無音区間の時間が短い。言い換えるならば、促音と非促音の生成範疇境界は無音区間の時間長に関係していると言える。

しかし、この無音区間の時間長は、発声速度によって変動することが分かっている。すなわち、発声速度が速ければ無音区間の時間長は短くなり、逆に発声速度が遅ければ無音区間の時間長は長くなる。したがって、無音区間の時間長だけでは、様々な発声速度における促音と非促音の生成範疇境界を正しく表すことはできない。

この発声速度による変動を吸収し、発声速度に依存しない促音・非促音の生成範疇境界の表現を求めることを目的として、Amano & Hirata (2010)は、Hirata & Whiton (2005)が録音した日本語2音節単語における促音・非促音の生成データを分析した。彼らはその分析結果に基づき、「促音・非促音の生成範疇境界は、閉鎖区間の時間長とそれを含む単語の時間長の2変数による1次式で表される」と主張した。これに対しHirata & Whiton (2005)は、「促音・非促音の生成範疇境界は、閉鎖区間の時間長とそれを含む単語の時間長の比で表される」と主張している。Amano & Hirata (2010)およびHirata & Whiton (2005)が分析に用いた生成データは、話者数が4名と少ないため、それぞれ主張の正否を判断することは難しく、またどちらの主張も一般化するにはやや難があった。

そこで、多数の話者による生成データによってAmano & Hirata (2010)およびHirata & Whiton (2005)の主張を検証することを目的として、『日本語話し言葉コーパス』RDB版(小磯・伝・前川, 2012)から複数話者が発声した促音・非促音を抽出して判別分析を行い、促音・非促音の生成範疇境界の表現を特定することにした。

なお、Amano & Hirata (2010)およびHirata & Whiton (2005)が用いた単語は全て2音節単語であるので、彼らの主張はそれぞれ「促音・非促音の生成範疇境界は、閉鎖区間の時間長とそれを含む2音節の時間長の2変数による1次式で表される」および「促音・非促音の生成範疇境界は、閉鎖区間の時間長とそれを含む2音節の時間長との比で表される」と読み替えることができる。本研究では、この読み替えた2つの主張をそれぞれ検証する。

2. 解析

2.1 解析対象

『日本語話し言葉コーパス』RDB版(小磯・伝・前川, 2012)から、以下の条件を満たす促音を含む音韻列 $C_1V_1QC_2V_2X$ 、および非促音を含む音韻列 $C_1V_1C_2V_2X$ を抽出し、解析対象とした。ただし、 C_1 : 先行子音、 V_1 : 先行母音、 Q : 促音、 C_2 : 後続子音、 V_2 : 後続母音、

[†] psy@asu.aasa.ac.jp

X: 任意の音韻である。

- V_1, V_2 は /a/, /i/, /u/, /e/, /o/ のいずれかである
- V_1, V_2 は無声化・長音化していない
- C_1 は任意の子音または空である
- C_2 は /p/, /t/, /ts/, /ch/, /k/ のいずれかである
- Q および C_2 は長単位で区切った場合の単語の語頭に位置しない
- C_1, C_2 は拗音を含まない
- X は撥音・促音ではない

この条件は Amano & Hirata (2010) が解析対象の促音・非促音を含む 2 音節単語の選択に用いた条件よりも緩和されている。緩和された点は、以下のとおりである。

- C_1 は任意の子音である
- C_2 に /ts/, /ch/ が含まれる
- V_2 に /i/, /u/ が含まれる
- 単語長が任意である
- 促音・非促音の単語内位置が語頭以外の任意の位置である
- 促音と非促音を含む音韻列がミニマルペアをなさなくてもよい

抽出の結果、得られた促音を含む音韻列の数は 6056、非促音を含む音韻列の数は 34012、合計数は 40068 であった。話者の異なり数は 201 名であった。

2.2 解析方法

抽出した音韻列の各音韻の時間情報から、促音および非促音の閉鎖区間の時間長 y (ms) と、閉鎖区間を含む 2 音節 $C_1V_1(Q)C_2V_2$ の時間長 x (ms) を計算した。時間長の計算は Amano & Hirata (2010) の計算方法と同一にするために、次の 2 条件を設けて実施した。

- C_1, C_2 に無音区間を含めない
- V_1, V_2 の終了後の声帯振動の持続時間は母音の時間長に含める

計算で得た y と x を独立変数とし、促音・非促音のカテゴリーを従属変数として 1 次式による判別分析を行った。さらに y と x の比 (y/x) を独立変数とし、促音・非促音のカテゴリーを従属変数とする判別分析を行った。

2.3 解析結果

閉鎖区間の時間長 y (ms) とそれを含む 2 音節の時間長 x (ms) を変数とする 1 次式を用いて判別分析を行った場合、得られた判別関数は、

$$y = 0.0588x + 38.4 \quad (\text{式 1})$$

であった。図 1 に、この判別関数と判別対象の促音・非促音のデータを示す。この判別関数による促音と非促音の誤判別率は 7.30% であった。

一方、閉鎖区間の時間長 y (ms) とそれを含む 2 音節の時間長 x (ms) の比 y/x を変数として判別分析を行った場合、得られた判別関数は

$$y/x = 0.255 \quad (\text{式 2})$$

であった。この判別関数による促音と非促音の誤判別率は 17.3%であった。

3. 考察

閉鎖区間の時間長 y (ms) とそれを含む 2 音節の時間長 x (ms) を変数とする 1 次式によって判別分析を行った場合の誤判別率 7.30% は Amano & Hirata (2010) が得た誤判別率 8.72% とほぼ同じである。この結果は促音と非促音の判別が、閉鎖区間の時間長 y (ms) とそれを含む 2 音節の時間長 x (ms) の 2 変数による 1 次式を用いてほぼ可能であることを示している。すなわち促音と非促音の生成範疇境界はこの 2 変数の 1 次式によって表せると言える。この結果は Amano & Hirata (2010) の主張を支持する結果である。ただし、Amano & Hirata (2010) が得た判別関数は

$$y = 0.446x - 17.1 \quad (\text{式 3})$$

であり、本研究で得られた判別関数(式 1)とは異なっている。この違いは、Amano & Hirata (2010) では「速い・普通・遅い」の 3 種の発声速度による音声を用いているのに対し、本研究では『日本語話し言葉コーパス』RDB 版(小磯・伝・前川, 2012)から抽出した普通の発声速度が大半を占める音声を用いていることに一因があると考えられる。すなわち、促音においても非促音においても、速い発声速度では y と x は小さくなり、遅い発声速度では y と x は大きくなるので、速い発声速度のデータや遅い発声速度のデータが普通の発声速度のデータと同数程度存在した場合、判別関数の傾きは大きな正の値になりやすく、かつ切片は負の値になりやすくなる。この傾向によって Amano & Hirata (2010) の判別関数(式 3)の傾きは大きくかつ切片が負となり、本研究の判別関数(式 1)の傾きは小さくかつ切片が正になったと考えられる。

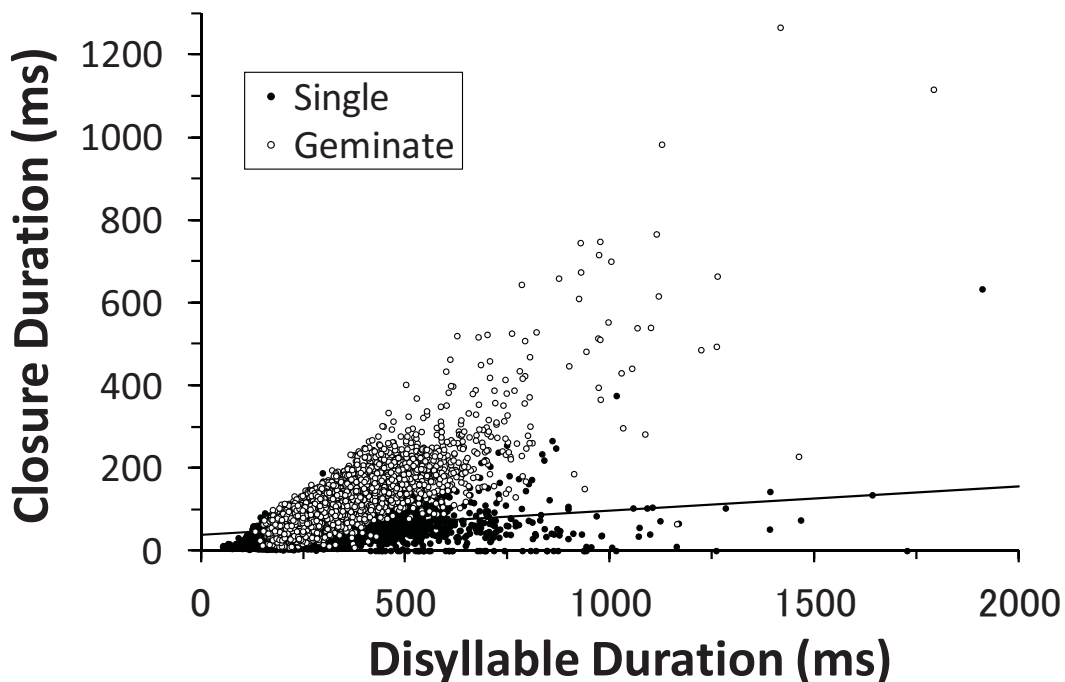


図1 閉鎖区間の時間長 y と閉鎖区間を含む 2 音節の時間長 x で表される平面上にプロットした促音(Geminate)・非促音(Single)のデータおよび判別関数($y = 0.0588x + 38.4$)

一方、閉鎖区間の時間長 y (ms) とそれを含む 2 音節の時間長 x (ms) の比 y/x を変数として判別分析を行った場合の誤判別率は 17.3% と大きい。これは、比 y/x を用いた場合、促音と非促音の判別が十分にできないことを意味している。したがって、促音・非促音の生成範疇境界は比 y/x では正確に表せないと言える。これは「促音・非促音の生成範疇境界は、閉鎖区間の時間長とそれを含む 2 音節の時間長との比で表される」という Hirata & Whiton (2005) の主張に反する結果である。本研究で用いた音声は Hirata & Whiton (2005) よりも、発声数、話者数、および母音や子音の種類が多いので、本研究の結果のほうが正しい可能性が高いと考えられる。

多数話者の生成データを用いた本研究の結果、閉鎖区間の時間長 y (ms) とそれを含む 2 音節の時間長 x (ms) の 2 変数による 1 次式を用いれば、促音と非促音の範疇境界が表せるとい結論が得られた。しかし、この結論はいくつかの制限条件の下で得られた解析結果に基づいているので、これがどのような条件においても成り立つという保証はない。特に次に示す場合に関しては、この結論が正しいか否か不明である。

- V_1, V_2 が無声化した場合
- V_1, V_2 が長音化した場合
- V_2 に撥音が後続する場合
- V_2 に促音が後続する場合
- C_1, C_2 に拗音が含まれる場合
- C_2 が摩擦音である場合

より一般性のある結論を導き出すために、これらの場合について今後さらに解析を進める予定である。

謝 辞

本研究の一部は国立国語研究所（言語資源研究系）基幹型共同研究「コーパス日本語学の創成」（リーダー：前川喜久雄）による成果である。また本研究は、愛知淑徳大学研究助成費（共同研究）「音声コーパスを用いた促音・非促音の範疇境界を表す判別関数の検証」（平成 23～24 年度）による助成を受けた。

文 献

- Amano, S., and Hirata, Y. (2010). "Perception and production boundaries between single and geminate stops in Japanese," *The Journal of the Acoustical Society of America* 128, 2049-2058.
- Hirata, Y., and Whiton, J. (2005). "Effects of speaking rate on the single/geminate stop distinction in Japanese," *The Journal of the Acoustical Society of America* 118, 1647-1660.
- 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.