



## 日本語音声の時間特性に関する基礎的研究

著者	松井 理直
雑誌名	トークス = Theoretical and applied linguistics at Kobe Shoin : 神戸松蔭女子学院大学研究紀要言語科学研究所篇
巻	23
ページ	49-56
発行年	2020-03-05
URL	<a href="http://doi.org/10.14946/00002166">http://doi.org/10.14946/00002166</a>

# 日本語音声の時間特性に関する基礎的研究

松井 理直

神戸松蔭言語科学研究所・大阪保健医療大学

michinao.matsui [at] ohsu.ac.jp

---

## A preliminary study of the temporal features of Japanese speech sounds

Michinao F. MATSUI

Shoin Institute for Linguistic Sciences, Osaka Health Science University

### Abstract

本研究では、C/Dモデルの観点から日本語音声の時間特性について概観する。結論として、基底状態におけるモーラの重要性および基底状態に対するある固有の子音が持つ等時間性の性質について述べる。

This study proposed some temporal features of Japanese speech sounds from the view points of the Converter/Distributor model. The results of two experiments show that the mora-timing is the most basic feature of the ground state and the equi temporal feature of the unique consonant against the ground state.

キーワード: 日本語音声, 時間特性, C/Dモデル

**Keywords:** Japanese phonetics, temporal feature, the Converter/Distributor model

## 1. 研究の目的

意味部門や統語部門と密接な関係を持つ音韻情報と、現実の音響特性と関わる音声情報は、実時間性という点で大きな違いを持つ。一般に音韻情報は線条性と階層構造を持つが、実時間情報は捨象されていて良い。しかし音声情報は、それが詳細なものになればなる程、実時間性を無視できなくなる。

例えば、日本語音素の異音を考えてみよう。音素 /t/ の異音については、狭母音に先行するという音韻環境において必ず破擦音化が起こる。一方、音素 /z/ の異音が破擦音になるか摩擦音になるかを決める要因が音韻環境に依存したもののか否かは議論の余地を残す。

音声学の入門書では、句頭では破擦音になり、母音間では摩擦音になると記述されていることが多い。この解釈を取るなら、音素 /z/ の異音は音韻環境に依存した音韻過程の結果として生じるものと考えてよいだろう。一方前川 (2010) は、音素 /z/ の破擦音化が当該子音を発音するために必要な準備時間に依存することを見いだした。すなわち、音素 /z/ の異音が句頭で破擦音化が起こるのは子音準備時間が長く取れるからであり、母音間で摩擦音になるのは子音準備時間が十分に取れないことが原因というわけである。この解釈では、音素 /z/ の異音は音声過程によって生じるものと見なせるであろう。

前川の解釈は極めて妥当なものと考えられるが、一つ難しい問題が生じるのは、撥音に後続する音素 /z/ の異音である。少なくとも筆者がエレクトロパラトグラフィを用いて調べた限りでは、撥音に後続する音素 /z/ の異音は、句頭における異音よりも破擦音化する可能性が高い。実時間の特性としては、撥音の調音時間より句頭における調音準備時間のほうが短いとはいえないであろう。したがって、撥音に後続する音素 /z/ の異音が音声過程によって生じたものと断言するのは難しい。

こうした例が示すことは、音韻過程と音声過程を常に明確に切り分けることができるとは限らないということである。このことは、多くの音韻理論が「あらゆる」レベルにおいて実時間性を捨象している点に疑念を生じさせよう。もちろん、実時間性を考慮した枠組みもないわけではない。そうした枠組みとして、Browman (1992) による *articulatory phonology* や Fujimura (2002), 藤村 (2007) による C/D モデルを挙げることができよう。

本稿では、特に C/D モデルの枠組みに基づきながら、日本語の音韻・音声情報を扱う上で必要となる実時間性について簡単に考察を行ってみたい。

## 2. C/D モデルの概要

### 2.1. 音節：C/D モデルにおける基本的な実時間構造

まず初めに、藤村が提案した C/D モデルの枠組みについて簡単に見ておく。この理論では、調音運動を音節の大きな流れ上に局所的な子音動作を乗せて行われるものと見なす。例えば「甘美」という語の調音運動は、母音 “e, i” が作る大きな調音変化に子音 “k, m, b” が局所的に影響を与える形で計算される。分節音の情報である素性の束も、音素単位ではなく音節単位で「甘美：<{low, stop, dorsal, nasal<sup>oda</sup>}, {high, palatal, stop, labial, voiced}>」といった集合のリストを成す。一般的な音韻論では音韻素性を二値的な情報と考えるが、C/D モデルにおける音韻情報は原子的素性と捉えられている。したがって、音韻素性は *Particle Phonology* や *Government Phonology* における音韻要素のようなものであってもよい。

C/D モデルと *articulatory phonology* の大きな違いは、この音節構造に関する点にある。*Articulatory phonology* は、音節という韻律構造を特に重要視するわけではない。調音に関する様々な素性は *score* という形で実時間上に表現されていくが、それは韻律構造に支配されたものではなく、線状的に表現されていく。一方、C/D モデルは実時間上の構造として、音節構造を極めて重視する枠組みとなっている。

## 2.2. C/D モデルにおける母音と子音の実時間性

さらに C/D モデルでは、音韻素性・韻律・発話スタイルといった定性的情報を「シラブル・パルス」を中心とした定量的性質に変換し (convert), その中に素性の定量的特徴を分配 (distribute) する. 図 1 に、この変換/分配の様子 (C/D ダイアグラム) を示す. 太線で示された 2 箇所の縦棒がシラブル・パルスで、全ての計算の元となる。

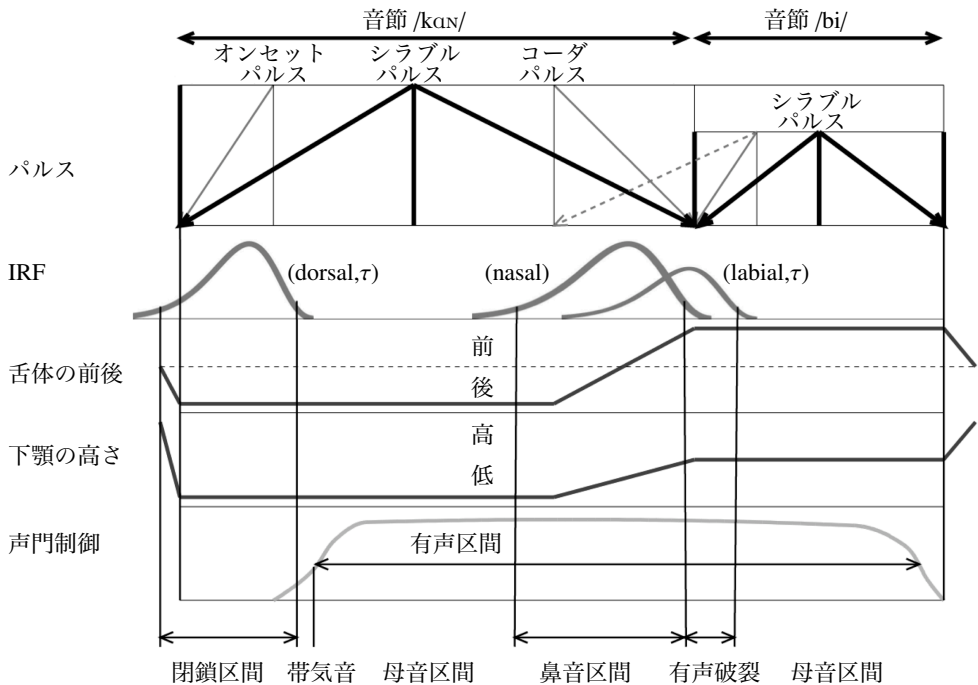


図 1: 「甘美」の C/D ダイアグラム

このシラブル・パルスの強度によって、音節およびその中核領域 (通常は母音) の「実時間領域」や強さが決まってくる. 例えば、「甘美」の場合は最初のシラブル・パルスが強いため、第 1 音節の持続時間が長い. こうした抽象的な音節時間領域は、シラブル・パルスの頂点から左右に伸びる「シラブル三角形」(図 1 の太い矢印) で示される. なお C/D モデルでは、理論の自由度を制限するために一発話内のシラブル三角形は相似形を成すという仮定を置いている点に注意されたい.

さらに音節中核領域の周縁には、子音に関する周縁パルス (オンセット・パルスおよびコーダ・パルス) が出力される. 周縁パルスはシラブル・パルスと同一の高さを持ち、周縁パルスの強度はこの高さでパルスの頂点から出る半三角形の傾斜角によって決まってくる. この時、子音の各素性は、周縁パルスの強度に基づく impulse response function (IRF) として局所的に分配される. 図 1 における 3 つの IRF は各々「軟口蓋破裂」「鼻音性」「両唇破裂」の情報を表し、各素性の時間特性を決めていく. こうした子音の局所

的性質に対し、V-to-V Coarticulation 特性などから母音の前後・狭広の情報や声帯振動の特性は音節全体に跨がる大局的な運動と考えなければならない (Fujimura, 1992; Ohman, 1966). C/D モデルが音節単位で素性束を指定し、母音素性に添え字を付けない理由がここにある。言い換えるなら、“low” という素性は広母音 [e] の性質という以上に、音節全体自体が持つ口腔の状態を表している。

### 2.3. C/D モデルにおける実時間性に関する不明点

現在のところ、C/D モデルにおける詳細な研究が少ないため、このモデルにおける実時間情報の扱い方については不明な点が多い。例えば、基底構造を形作る母音の時間特性、子音に関わる IRF における時間積分の特性、そして何よりも音節に関する時間特性といった問題が、未解決のまま残されている。

## 3. 日本語における韻律構造の時間特性

### 3.1. モーラの等時性

日本語では、少なくとも心理的にはモーラがほぼ等間隔のリズム単位を担う。では、物理現象である調音として、モーラが等時性は言えるのであろうか。ここで、日本音響学会が出している音声データベースから 500 文を選択し、 $C_1V_1C_2V_2C_3V_3$  の連鎖について分析を行った。まず、図 2 に  $C_2$  と  $V_2$  間に見られる時間補償効果の関係を示す。

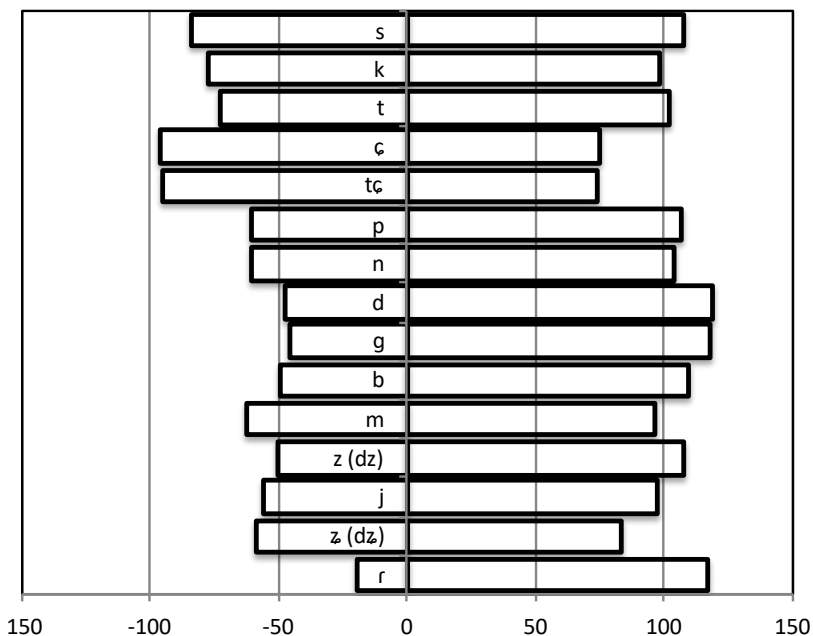


図 2: 語中のモーラにおける先行子音・後続母音間の時間補償効果

図2のグラフからも分かる通り、同一モーラ内における先行子音が長ければ後続母音が短く、先行子音が短ければ後続母音長くなるような補償作用が観察される。この点について、 $V_2$ の伸縮値に最も強い影響を与える要因を分散分析で解析したところ、 $C_2$ の影響が圧倒的に強い。次いで、 $C_3, V_3, V_1$ の順で影響が見られるが、これは $C_2$ の影響力とは有意に差がある。この結果は、同一モーラ内において子音と母音が補償効果を持っていることを示す。なお、 $V_2$ と後続子音 $C_3$ との間に隣接効果が、また $V_3, V_1$ と $V_2$ との関係においては、V-to-V coarticulationの関係が存在するといつてよい。

C/Dモデルにおいては、既に述べたようにV-to-V coarticulationはモデルとして表現されている。また、母音と後続子音との隣接効果は、後続子音のIRFが先行母音の領域に食い込むという形で表現できるであろう。一方、 $C_1$ が $V_2$ にほとんど影響を及ぼさないことも、 $C_1$ の属する音節と、 $V_2$ の属する音節は独立した時間制御を受け、 $C_1$ のIRFは次の音節に食い込むことができないことから理解される。

### 3.2. モデルにおけるモーラの必要性

問題となるのは、撥音・促音などが自立モーラに後続する重音節の時間制御であろう。日本語における重音節が軽音節よりも長い持続時間長を持つことは、C/Dモデルにおけるシラブル三角形の大きさから導き出せる。しかしこの条件のみでは、日本語の重音節が軽音節の約2倍(厳密には2.1~2.4倍ほどの範囲に及ぶ)の持続時間長を持つことを保証できない。

したがって、C/Dモデルの表現の中にモーラという単位を組み込むことが妥当であろう。数学的性質として集合は{a, b,...}といった要素と共に、{{a}, {a, b},...}のように集合自体を元として取り得る。ただし、要素を取る集合は多重集合でない限り{a, b}と{a, a, b}を区別しないが、集合を元とする{{a}, {b}}と{{a}, {a, b}}は同一の集合ではない。同様にC/Dモデルの入力情報である音節集合も、原子要素のみならず「原子要素の集合」も要素として取るとしよう。この音節集合の要素となる「原子要素の集合」が、モーラに相当する単位を表す(これを「モーラ集合」と呼んでおく)。

例えば「甘美」の第1音節「カン」の入力情報は、{low, stop, dorsal, nasal<sup>coda</sup>}の代わりに、{{low, stop, dorsal}, {nasal}}と表現される。この{{low, stop, dorsal}, {nasal}}は集合として一体となった情報であり、集合なので要素の順序は意味を持たない。ただし、素性[low]が大局的動作を決めるのに対し、[stop], [dorsal]は子音の特性を決める局所的な要素であり、図1が示すように、シラブル・パルスに基づくシラブル三角形の左端を基準として、子音要素の特徴がインパルス応答関数(IRF)によって局所的に実現される。この結果、頭子音と母音の時間順序と性質が自然に決まっていく。同様に、モーラ集合{low, stop, dorsal}と{nasal}も音節集合の要素に過ぎないため、音節として一体であり、モーラ集合の順序関係は存在しない。ただし、前者は[low]という大局的な運動情報を含むモーラ集合だが、[nasal]は子音要素のみのモーラ集合であることから、後者の鼻音性はcodaとしてシラブル三角形の右端に局所的に分配されることが計算され、撥音としての時間順序とその特性が決まる。

こうしたモーラ集合の導入は、「活気」における頭子音の [k] 音が開放を持つのに対し、促音の [k] 音が開放を持たないことの説明にも役立つ。すなわち、頭子音の [k] 音は同一モーラ集合に含まれる母音要素によって開放が引き起こされる(日本語の破裂音は厳密には閉鎖音である)のに対し、促音の [k] 音はモーラ集合に母音要素を含まないために開放部を持たず、音響上は一定時間の無音が続く。このことは、言い換えると日本語の [t] 音や [k] 音が「破裂」音ではなく、「閉鎖」音として表現されていることを示唆するものである。

#### 4. 基底状態に起因する時間特性

##### 4.1. 基底状態における下顎動作

次に、基底状態を決める母音の調音に関する時間特性について簡単なデータを示す。基底状態は、下顎の位置とそこからの舌体挙上動作によって決まっていく。ここで時間特性として興味深いのは、下顎運動に関するものである。

今、パ行子音の調音における口唇の垂直開口度について見てみよう。男性 10 名、女性 3 名の被験者を対象に、上唇中央部と下唇中央部および左右口角部に光学センサーを取り付け、サンプリング周期 2ms で口唇の垂直方向および水平方向のデータを取得した。被験者によって口唇動態の範囲が異なるため、ア音における口唇の開きを基準にし、被験者のデータを標準化した上で集計を行った結果を図 3 に示す。パ行子音は、両唇閉鎖を必ず伴う。したがって、図 3 の口唇垂直開口度は下顎の動態と強い相関を持つと考えてよいだろう。

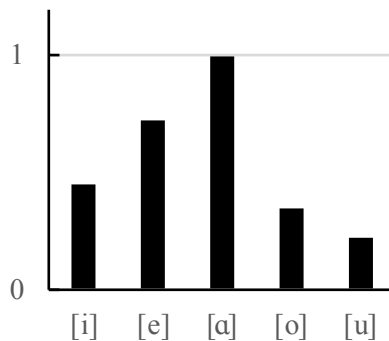


図 3: パ行調音における母音部の口唇垂直開口度

図 2 の結果から、もし下顎の運動がどのような環境であってもほぼ同じ速度で行われるのであれば、パの下顎下降時間が長く、次いでペ・ピ・ポの順となり、プの下顎下降時間が一番短くなるはずである。ところが、実際の下顎下降時間はパ・ペ・ピ・ポで有意差がなく、プのみで有意な差が生じるという結果であった。

#### 4.2. 下顎動作の時間特性

この結果は、子音における下顎の調音動態速度が一定ではなく、むしろ調音動態時間がほぼ一定であり、下顎の調音動態速度自体は後続母音に依存して調整されていることを示す。これは、IRF の特性の積分特性が時間依存なのか速度依存なのかを考える上で興味深いデータの一つと思われる。EMA を用いた調音動態のグラフを見る限り (精密な数値は示されていないため、厳密な議論はできないが)、英語では後続母音が広母音であるほど同一の子音であっても下顎下降時間が長くなる傾向にあるようだ。もしこれが正しいデータだとすると、英語はある子音が持っている固有の特性が基底状態に対して速度一定の拘束条件で決まっていくのに対し、日本語においてはある子音における固有の特性が時間を束縛条件として決まっていくのかもしれない。この点については、今後より精密な研究を行う予定である。

#### 5. まとめ

本稿では、日本語の調音動態における実時間特性についていくつかのデータを元に議論を行った。結果として、日本語はある程度の分散はあるものの物理的にもモーラを基本とした時間特性を持っていること、またモーラ内における子音の特性は、同一の子音である限り、母音が構成する基底状態に対して類似した時間特性を持つ。

このことから、C/D モデルは音節を基本単位とするため、音節の下部構造として音節量を決定するモーラ集合を導入する必要があるだろう。これによって、同一モーラにおける先行子音と後続母音間における持続時間長の補償効果が説明できると共に、重音節と軽音節の時間長関係についても適切に取り扱えるようになる。

一方、局所的な情報である子音の IRF については、時間束縛条件と時間変異条件 (速度束縛条件) を考える必要があるかもしれない。この点については、詳細な研究が必要であるが、もし英語と日本語の子音において IRF の拘束条件が異なるとすれば、拘束条件が何によって生じるかの説明も必要となるだろう。1つの可能性として、日英語における音節構造の違いに原因を求めることができるかもしれない。英語は音節内における母音に対して onset 子音の独立性が比較的高いのに対し、日本語は先行子音と後続母音がモーラによって強く拘束される。日本語ではモーラが時間特性の基本単位であるため、子音がモーラに束縛されていることが、日本語における IRF の時間束縛条件を引き起こしている可能性があるだろう。この点については、CV 音節を基本とする言語の調音動態を調査することによって、その妥当性を検討しなければならない。また、各種のデータベースを分析して、方言も含め、日本語音声の持つ時間特性についてより詳細な研究も必要であるが、これも今後の課題としたい。

#### 謝辞

本研究は、日本学術振興会科学研究費補助金・基盤研究 (C) 「日本語音声の調音協調運動に関する総合的研究」(2018 年度～2021 年度、研究代表者：松井理直、課題番号：



18K00597), および国立国語研究所プロジェクト「対照言語学の観点から見た日本語の音声と文法」(プロジェクトリーダー: 窪菌晴夫教授)の援助を受けました。

### 参考文献

- Browman, Catherine and Goldstein, Louise (1992) Articulatory phonology: an overview. *Phonetica*. 49 (3-4): 155–180.
- Fujimura, Osamu (1992) Phonology and phonetics — a syllable-based model of articulatory organization. *The Acoustical Society of Japan (E)* 13(1): 39–48.
- Fujimura, Osamu (2002) Temporal organization of speech utterance: A C/D model perspective. *Cadernos de Estudos Linguísticos, Instituto de Estudos da Linguagem, Campinas* 43: 9–36.
- 藤村靖 (2007) 『音声科学原論—言語の本質を考える』岩波書店, 東京.
- 前川喜久雄 (2010) 「日本語有声破裂音における閉鎖調音の弱化」『音声研究』14 (2): 1–15.
- Öhman, S. E. G. (1966) Coarticulation in VCV utterances: spectrographic measurements. *Journal of Acoustical Society of America*, 39, 151–168.

**Author's web site:** <http://sils.shoin.ac.jp/~matsui/>

(受付日: 2019年12月10日)