

歯擦音の母音無声化・VOT 分布・促音挿入

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

Vowel Devoicing, VOT Distribution and Geminate Insertion of Sibilants

Michinao F. MATSUI

Shoin Institute for Linguistic Sciences, Osaka Health Science University

Abstract

日本語の有声破裂音における声立ち上がり時間 (Voice Onset Time, VOT) の分布は、しばしば双極的な性質を持つ。2つの分布ができることを説明する1つの可能性として、声帯振動の開始時点を決める基準点が1つではなく、2つあることが考えられる。例えば、1つは破裂音の開放時点、もう1つは子音・母音間の音韻境界を過程できよう。本稿は、この子音・母音間の音韻境界が日本語の母音無声化、阻害音の有声性、借用語における促音挿入／促音抑制といった様々な音韻現象に影響することを考察する。また、最後に、母音・子音間の音韻境界に基づくモデルが、VOTの分布をよくシミュレートできることを述べる。

The distribution of voice onset time (VOT) in voiced plosives usually has bimodal feature in Japanese. One of plausible explanation for “two peaks” is that the command of vocal fold vibration is controlled on two different reference points: one is based on the release timing of closure, the other reference point is the phonological boundary which indicates both the ending timing of the preceding consonant and the starting timing of the following vowel. This paper proposed that the consonant-vowel boundary has relation with various phonological phenomena in Japanese: vowel devoicing, voicedness of obstruents, geminate insertion/inhibition in Japanese loanwords and so on. Finally this paper shows that the model according to consonant-vowel boundary can simulate the distribution of VOT.

キーワード: VOT の双極分布、母音無声化、極周波数遷移、音韻境界、C/D モデル

Key Words: Bimodal distribution of VOT, Vowel devoicing, Pole frequency transition, Phonological boundary, C/D model

1. はじめに

日本語は、音声の有声性に纏わる様々な現象を持つ。その代表がライマンの法則で、あるいは、和語や漢語の複合語において、後続形態素が有声阻害音(いわゆる濁音)を持っている場合には、連濁が起こらない。また、日本語では借用語や「ひっどい」のような特殊な語形を除き、有声子音の促音が存在しない。借用語であっても、「ブルドック」「ハンドバック」のように、有声促音が無声化する現象がしばしば観察される。母音についても、東京方言では無声子音に挟まれた狭母音が、ほぼ義務的に無声化を起こす。

調音上は、有声性の制御は声帯の振動によってもたらされる。無声音は声帯の隙間である声門が広く空いている時の発音で、呼気流が素通りするために声帯振動が起こらない。一方、声門を軽く閉じると、呼気流がむりやり通り抜けていく力で、声帯の振動が引き起こされ、有声音が生成される。知覚の面では、声帯の振動によってもたらされる低周波数域のエネルギー分布(スペクトログラムではボイスバーとして現れる)や音の調波性が、有声性の知覚をもたらす。特に語頭の破裂音においては、破裂の瞬間と声帯振動の発生時間との関係が有声性知覚の主要な手がかりとなっていることが、Lisker and Abramson (1964) によって明らかにされている。彼らはこうした破裂の瞬間と声帯振動開始時点との関係を、VOT (voice onset time, 声立ち上がり時間) と呼んだ。

VOT の値と有声性の関係は普遍的なものではなく、言語によって若干の違いが生じる。大雑把に一般化すると、VOT がマイナスの値を取る場合、すなわち声帯振動が破裂の瞬間より前に発生している場合には有声音と知覚され、VOT がプラスの値を取る場合(声帯振動が破裂の瞬間より遅れて生じる場合)には、無声音として知覚されやすい。ただし、その境界となる VOT の値は言語によって微妙な違いが見られ、英語ではほぼ VOT=0msec 付近が有声/無声の境界となるのに対し、日本語では VOT が若干プラスの値を取っている場合でもしばしば有声音の知覚が生じる。

有声性と VOT の関係は、音声の生成においてさらに興味深い性質を持つ。高田 (2011) は、日本語の語頭有声音について様々な方言の VOT を測定し、その分布について非常に興味深く厳密な研究を行っている。Lisker and Abramson (1964) でも報告されていることであるが、有声破裂音は図 1 に示すように、マイナスの分布と若干プラスの分布に二極化することが多い。このデータは筆者自身が 23 歳~39 歳の関西方言話者男性 4 名・女性 9 名を対象に、語頭破裂音の VOT を計測したものである。実線が有声破裂音の VOT であるが、VOT=0msec の生起確率がほぼ 0% になり、双極分布を成している点に注目されたい。

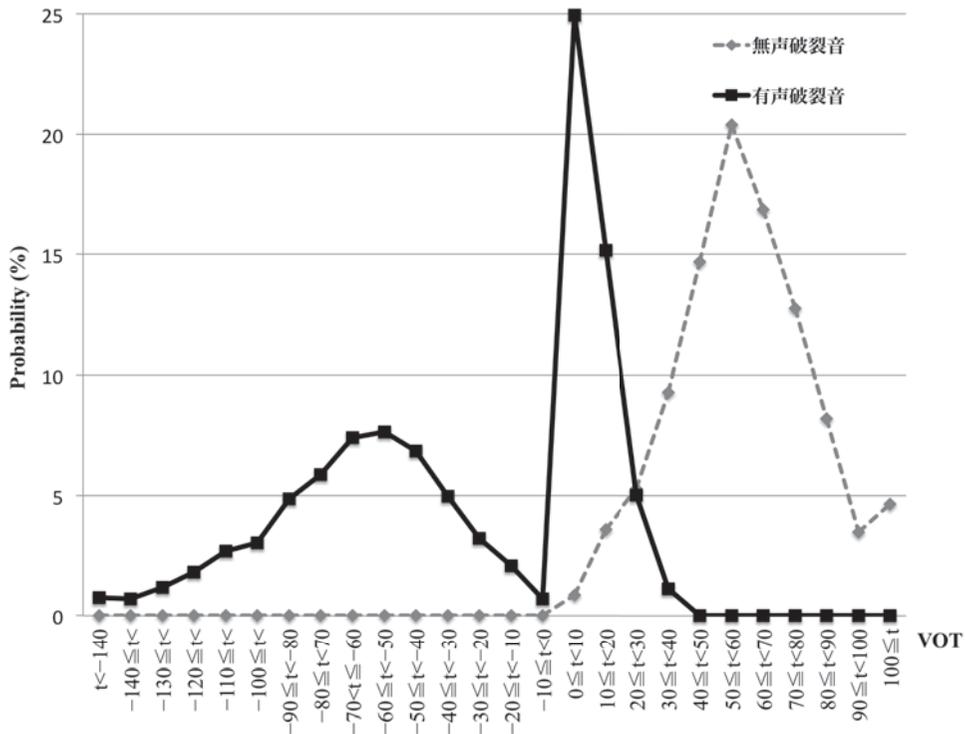


図 1: 関西方言における VOT の分布例

しかし、高田 (2011) によると、関西方言における高齢者の発音では、語頭有声音の VOT はマイナス側の分布のみが出現し、プラス側の分布はほとんど観察されないという。一方、東北方言話者の語頭有声音では、マイナス側の分布は基本的に出現せず、VOT の分布がプラス側に偏る傾向が非常に強い。このように、高田 (2011) の研究によって、マイナスの VOT 分布あるいはプラス側の分布といった片側の分布しか持たない世代・方言の存在が明らかになった点は極めて重要である。このことは、VOT の双極分布は偶然のものではなく、ある明確な制御によってもたらされるものであり、VOT=0msec における落ち込み (以後、これを 0msec-dip と呼ぶ) も偶然ではなく、何らかの理由に起因する現象であることを明確に示すものといっていよう。

しかし、VOT の分布が双極化する理由、および 0msec-dip の生じる原因は、現在のところ明確ではない。本稿では、有声性の生成に関する理論に基づいて VOT の分布をシミュレートするモデルを提案し、VOT の分布が双極化する理由に関する仮説を提出してみたい。まず次節で、VOT 分布の双極化する理由に複数のモデルが考えられ、いずれも VOT 分布を妥当に説明することを示す。その後、そのうちの 1 つのモデルを取り上げ、そのモデルの妥当性について音声知覚の面から検証を行う。

2. VOT 分布のシミュレーションモデル

2.1 モデルの基盤となる理論について

VOT 分布のシミュレーションを行うためには、まず音声生成に関する定量的な理論を採用する必要がある。音声生成の数理モデルに関しては、基本周波数に関する藤崎モデルが有名であるが、有声破裂音という分節音レベルでの生成については、未だに確固とした理論があるわけではない。Klatt Synthesizer も分節音の性質を決める極めて重要な工学モデルであるが、これは分節音のパラメータを全て決められていることを前提とした音響モデルである。VOT 分布のシミュレーションを行う場合、分節音のパラメータ自体を「定性的な言語情報」から「定量的」に決められるものでなければならない。現時点でこの目的に最も合致する理論として、Boersma (1998) による“Functional Phonology”や藤村 (2007), Fujimura (2002) による“C/D model (Converter/Distributor model)”を挙げることができよう。このうち Functional Phonology は、弁別素性を言語情報の基本とし、その生成情報と知覚的制約に基づいて定量的な性質を計算していく。例えば、VOT に関して言えば、[continuant] 素性から得られる分布と [voiced] 素性から得られる生成-知覚分布の相互作用から計算される。一方、藤村靖氏による「C/D モデル」は、音節単位で音声特性の時系列変化に関する計算を行う。各分節音における個別の特徴は、音節構造自体が出すパルス指令の分布に基づいて演繹的に生成される。

いずれの理論も、定性的な素性の集合から定量的な値を導出する演算に関してはまだ研究途上にあり、明確な性質は分かっていない。ただし、本稿の目的である VOT 値の導出に関しては、C/D model のほうがより強い制限を持つ。Functional Phonology では、[continuant] 素性と [voiced] 素性の時間的な関係はなんらかの相互作用を導入しない限り独立に計算できるため、VOT 分布特性を演繹的に導出することが難しい。それに対し、C/D model は、onset の子音に関する特性分布について一定の制限が設けられているため、VOT 値を理論的に計算することができる。したがって、本稿ではシミュレーションモデルの基盤として、VOT に関する制限がより強い枠組みである C/D モデルを採用したい。ただし、本稿における議論の多くは、Functional Phonology にも適用できるものである。

2.2 C/D model の枠組み

本節では、藤村 (2007), ? (?) にしたがって、C/D model の概略を述べる。この理論では調音の運動を、母音を中心とする音節列の上に onset や coda の子音動作を乗せて発音動作を行うものと見なす。例えば、「音(おと)」という語の場合は、母音「オ」の流れの中に子音“t”を乗せる形で発音を行う。各分節音の性質は素性で表現されるが、通常の音韻論とは異なり、素性の束は分節音単位ではなく、音節単位で集合を成す。例えば、「マン」という 1 音節の単語は、{low, back, nasal^O, labial^O, nasal^C} といった素性の集合で表される。[nasal] 素性の右肩についている O, C の記号は、各々 onset, coda の分節音を持つ素性であることを示す。なお、C/D model でどのような素性を用いるかについては明確でない点も多い。過小指定 (underspecification) を導入するか否かについても議論の余地を残す。

C/D model では、こうした音節ごとの素性集合と共に、その音節が強く（あるいは高く）発音されるか否かといった韻律的な性質や、発話のスタイル（肯定文／疑問文といった違いのみならず、感情表現といったパラ言語情報も含む）といった性質も加味しながら、言語情報の定性的な性質を「シラブル・パルス」という時間軸上に広がる定量的な性質に変換 (convert) する。このシラブル・パルスの大きさによって、その音節の時間長や強さ・高さなどの性質が決まっていく。さらに、音節境界に子音の性質を「周縁パルス」として分配 (distribute) する。こうしたパルスの形状に基づいて、各素性の時間特性が IRF (impulse response function) によって定量的に決まり、最終的な調音運動に結びつく。Fujimura (2002) による C/D モデルの概念図を、図 2 に引用しておく。

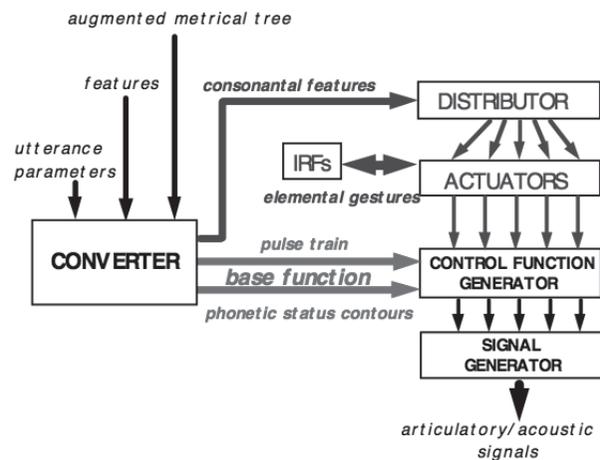


図 2: C/D model の概念図

シラブル・パルスは、図 3 に示すように、左右対称形のシラブル三角形 (syllable triangle) として表現される。1 つの発話内では、このシラブル三角形の高さは変化し得るが、三角形の角度は変わらない。すなわち、シラブル・パルスの形状は同一発話内では相似形を成す。本稿では、このシラブル・パルスによって決まる境界を「音節核境界 (音節核開始境界および音節核終止境界)」と呼んでおく。潜在的には、この音節核開始境界から音節核終止境界までが音節の中心である母音の性質がマップされる範囲となる。ただし、実世界における実現値は、筋肉の応答時間や運動制御の遅れといった諸条件の影響により、音節核境界の範囲からずれてしまう。

音節の中心を担う母音に対し、子音は音節の onset, coda といった周辺に分布する分節音である。こうした子音の性質を決定する「周縁パルス」を生み出す指令は半三角形で表され、これを「句境界三角形」と呼ぶ。Onset の句境界三角形は音節核開始境界を基準に生成され、coda の句境界三角形は音節核終止境界を基準として周縁パルスの分布を成す。なお、Onset と coda の句境界三角形は形状が異なる。通常、英語では coda の子音は弱いため、coda の句境界三角形における底辺の長さが onset の句境界三角形よりも短い。

C/D model では、これらのパルスの性質によって発話の時間的な特性が完全に決定される。また、子音の持つ様々な素性の性質は、句境界パルスの強さに基づいて、各素性ごとに固有に持つ IRF によって決められていく。このように、C/D model では、分節音単位ではなくシラブル単位で時間軸上に広がった定量的性質を計算していく点に特徴がある。

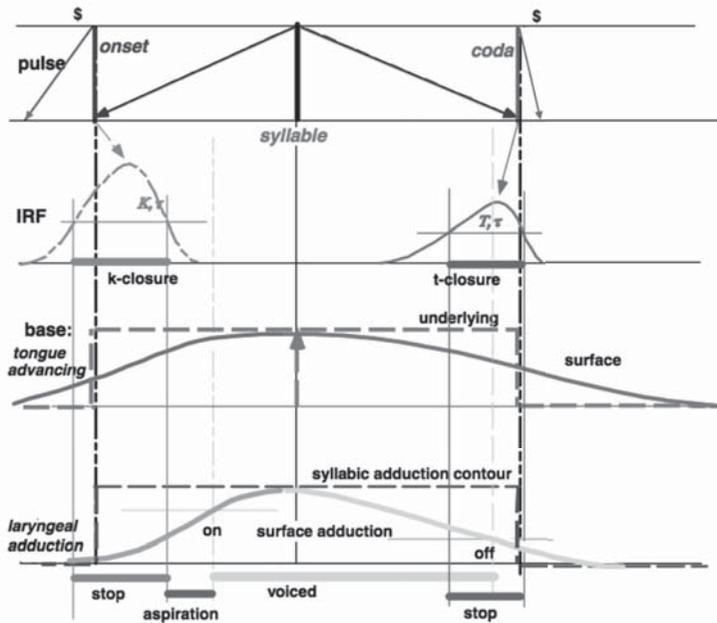


図3: 'kick' [kɪk] の CD ダイアグラム (Fujimura (2002) より引用)

前述したように、シラブル三角形の大きさは、アクセントやリズム、開口度に依存する。図4に、藤村(2007)によって分析されている "It's an echo" という発話の CD ダイアグラムを示す。この発話では、'echo' という単語が最も強く発話され、その中でも [ɛ] が最も強い。すなわち [ɛ] を含む音節のシラブル三角形が最も高く、シラブル三角形が相似形を成すという性質により、時間軸上での広がりも一番長くなる。

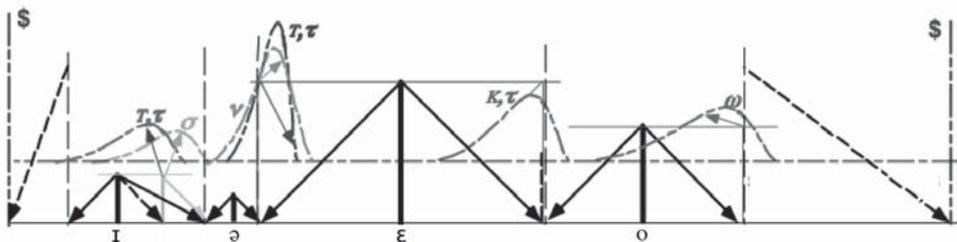


図4: 'It's an echo' の CD ダイアグラム (Fujimura (2002) を一部改変)

2.3 本稿における周辺パルスの扱い

C/D model におけるパルスの性質については、オハイオ州立大学で Matlab 上の定量モデルが開発されているようであるが¹、現在のところ一般には公開されていない。調音方法により子音パルスの性質は大きく異なるであろうが、阻害音についていえば時間的に非対称なインパルス応答を想定するのが妥当であろう。そこで本稿では、阻害音のインパルス応答はガンマ分布に従うと仮定する。また、藤村 (2007) では、周縁パルスが決定する各素性の時間特性に対し、実際の調音運動は一定の予測を行いながら行われると考えられているが、本稿では周縁パルスが決める時間特性が「調音運動の指令」が行われるタイミングであると仮定しておく。²

図 5 は、高さの異なる句境界三角形によって生成された 2 種類の周縁パルスを示したものである。縦軸は分節音の持つ各素性が実現されるレベルを、横軸は音節核開始境界を基準 (ゼロ) とした時間分布を表す。例えば図 5 では、縦軸に閉鎖性実現レベルが取られており、このレベルが“1”を超えた範囲で破裂音の完全閉鎖が起り続けることを意味する。したがって、周縁パルスが強い (句境界三角形が高い) 場合には完全閉鎖が -76msec から +8msec の範囲で起り、84msec の閉鎖時間を持つ破裂音が生成される。一方、弱い周縁パルス (低い句境界三角形) では、音節核開始境界を基準にして完全閉鎖が -42msec から -2msec の範囲で起り、この子音はわずか 40msec の閉鎖時間しか持たない。このように、句境界三角形の高さから計算される周縁パルスの強さは、最終的に子音の定量的な強度を生じさせるのである。

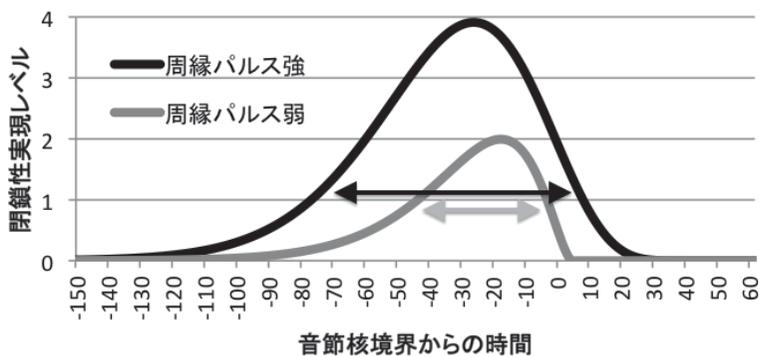


図 5: 破裂音のインパルス応答

周縁パルスがもたらすこうした性質は、子音の弱化といった現象にも有効である。例えば、アメリカ英語では語中の [t] 音や [d] 音がしばしば弾音化するが、C/D モデルに基づくなら、これは子音のインパルス応答が弱くなることによって生じる閉鎖時間の短縮と見なせよう (図 6)。また日本語でも、語中の有声破裂音や有声破擦音が接触性を失い、

¹川原重人氏からの情報による。

²この結果、Fujimura (2002), 藤村 (2007) におけるダイアグラムと本稿の計算結果は、予測の行われる時定数分だけズレが生じる。

[b] 音が [β] 音に、あるいは [dz] 音が [z] 音に摩擦化する現象が観察される。前川 (2010) によると、これは語中という音韻環境に動機づけられた弱化ではなく、発話速度に依存した音声的な弱化であるという。すなわち、発話速度が遅く、調音の時間を十分に確保できるのであれば、語中という環境であっても、有声破裂音や有声破擦音が摩擦化を起こすことは少ない。こうした現象も、C/D model では、発話速度が句境界三角形の大きさに影響を与え、速度が速いほど句境界三角形の高さは低くなると考えることで説明がつく。したがって周縁パルスも弱くなり、発話速度がある程度以上に速くなった時、閉鎖性実現レベルが完全閉鎖の閾値である“1”に到達し得ない (図 7)。これが、前川 (2010) のいう発話速度に依存した日本語の子音弱化の生成過程を C/D model の観点から見た時の捉え方である。

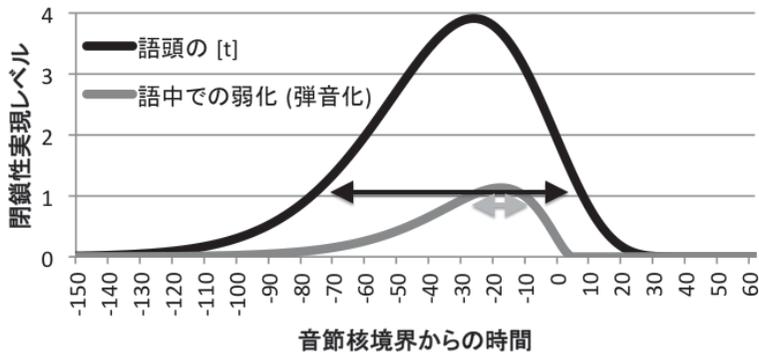


図 6: 英語における [t] 音の弱化 (破裂音の弾音化)

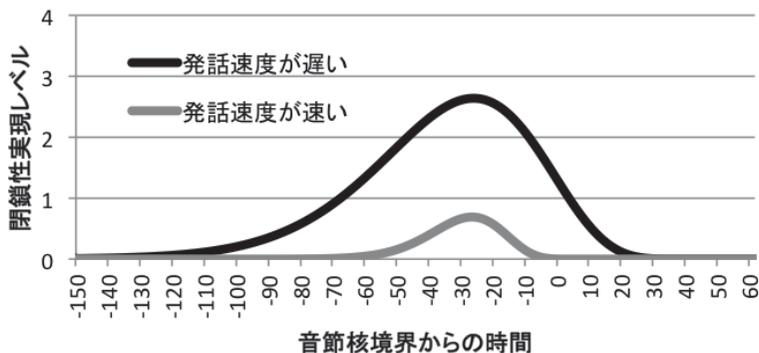


図 7: 発話速度に依存した日本語における語中有声阻害音の弱化

2.4 声帯駆動の矩形パルス指令

次に、本稿のテーマである有声／無声に関する C/D model の表現方法を述べておく。まず本稿では、有声性をもたらす声帯振動の性質を矩形パルス列によって表現する。これは、子音のような急速に変化する性質とは異なり、声帯振動は一定時間持続する性質を

持つという理由に基づく。これは、藤崎 須藤 (1973), Fujisaki and Hirose (1984) によって提案された音調動態の数理モデル (いわゆる「藤崎モデル」) でも採用されている前提である。藤崎モデルでは、アクセントとイントネーションの総合的な音調動態を「フレーズ指令」と「アクセント指令」との合成によって生成するが、前者はデルタ関数のインパルス応答、後者は声帯振動の矩形パルス列を用いた表現を行う。本稿でも、こうした考え方を援用することにしたい。

図 8 に、声帯振動を引き起こす矩形パルス指令と、実際に起こる有声音現との関係を示す。音響的な有声音は、有声音現レベルが 1 を超えている時間帯で続く。パルスの立ち上がり時定数を $x\text{msec}$ とすると (図 8 では 30msec としている)、ステップ指令の開始から約 $x/2\text{msec}$ 遅れて有声音の振動が現れ、ステップ指令の終了から約 $x/2\text{msec}$ 以上遅れる形で無声音になる程度の声門開口度となっている。この結果、有声音の指令自体は $y\text{msec}$ の時間長を持っているとすると、実際に実現される有声音は $y\text{msec}$ よりも少し長く続くことになる。こうした指令と実現値とのズレも、シラブル・パルスや周縁パルスの強さに依存して決まっていく。

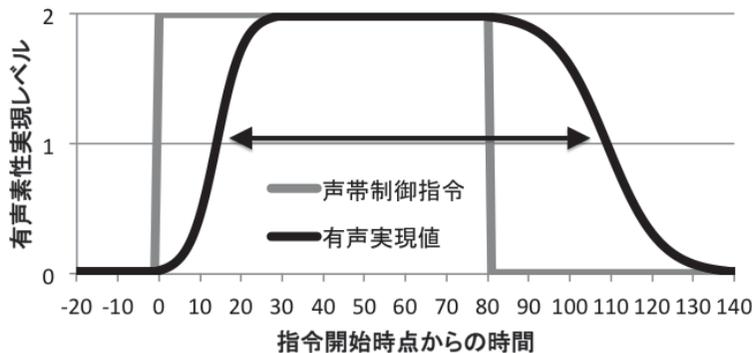


図 8: 声帯振動に関わる矩形パルス指令と実際の有声音区間

2.5 矩形パルス指令における開始および終止のタイミング

声帯振動を引き起こす指令が矩形パルスによって行われるとするなら、この指令の開始および終了するタイミングについても問題となる。このうち、終止時点については、次のように考えることにしよう。まず、発音において声帯振動が抑制されるきっかけは、無声音が存在する時か、あるいは発話の終了時点 (より正確には、Pierrehumbert and Beckman (1988) のいう intermediate phrase (major phrase) の終了時点) に限られる。したがって C/D model では、「声帯振動指令」の終了する時点 (1a) か (1b) のいずれかであると考えるよくだらう。

- (1) a. 「無声音」の周縁パルスが始まるタイミング (図 9 の A 点)。
- b. intermediate phrase 末に設定される音節核終止境界。

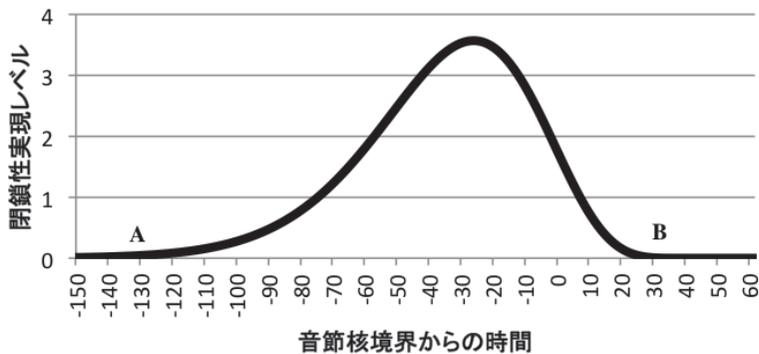


図 9: 周縁パルスの開始点 (A) と終了点 (B)

一方、声帯振動の始まるポイントは、無声子音が終わるタイミング (すなわち後続母音の始まるタイミング) か、有声子音区間中のタイミングが代表的なものである。あるいは、intermediate phrase が母音で始まる場合には、その母音の始まる時点も声帯振動の始まるタイミングを担う。また、本稿では他にも重要な基準点があると考えますが、これについては 3 節で議論する。本節では、まずは「声帯振動指令」の始まるタイミングとして (2a)~(2c) のいずれかが用いられるとしておく。後述するが、この特徴のうち (2b) の基準が、「有声子音の VOT は基本的にマイナスの値を取る (すなわち、開放時点以前の任意のタイミングで声帯振動が始まる)」という性質と密接に関わると考えられる。

- (2) a. 「無声子音」の周縁パルスが終了するタイミング (図 9 の B 点)。
- b. 「有声子音」の閉鎖区間中における任意のタイミング (図 5 の閉鎖区間を参照)。
- c. Intermediate phrase が母音で始まる場合のみ、初頭に設定される音節核開始境界。

2.6 周縁パルスの強さと無声破裂音の VOT

以上の性質を用いて、周縁パルスの強さが無声破裂音の VOT および帯気性に与える影響について見てみよう。「無声」破裂音であるので、声帯振動開始指令の出るタイミングは (2a) に従う。図 10 に強く子音が発音される場合、すなわち周縁パルスが強い (句境界三角形が高い) 場合の結果を、また図 11 に周縁パルスが弱い場合の結果を示す。無声子音の周縁パルスの終了時点が声帯振動の矩形パルス列「指令」の開始タイミングとなるため、「実際の声帯振動」はそこから一定時間遅れて始まることに注意されたい。なお、以後のグラフで、右向きの矢印は「VOT の値がプラスである」ことを、左向きの矢印は「VOT がマイナス値である」ことを意味する。グラフから周縁パルスと VOT の関係を見てみると、周縁パルスが強く出ている場合は完全閉鎖の区間が長く、また VOT もより長い値を持つ。一方、周縁パルスが弱くなるにつれて完全閉鎖の区間が短く (図では約 40msec)、VOT も短くなっていく。このように様々な周縁パルス強度から、無声子音の VOT 分布をシミュレートした結果が図 12 のグラフである。これは、完全閉鎖区間が僅

かに生じる周縁パルスの強度を基準値とし、この基準値から9倍の強度を持つ周縁パルスまでを正規分布に従って様々に発生させたもので、図1における無声子音のVOT実測値とよく似た形状を持ち、VOT分布のパターンをある程度模擬できているといえよう。

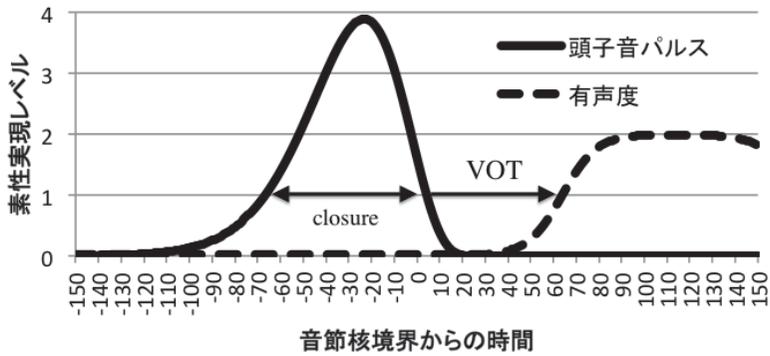


図 10: 強い無声子音の周縁パルスがもたらす閉鎖区間と VOT

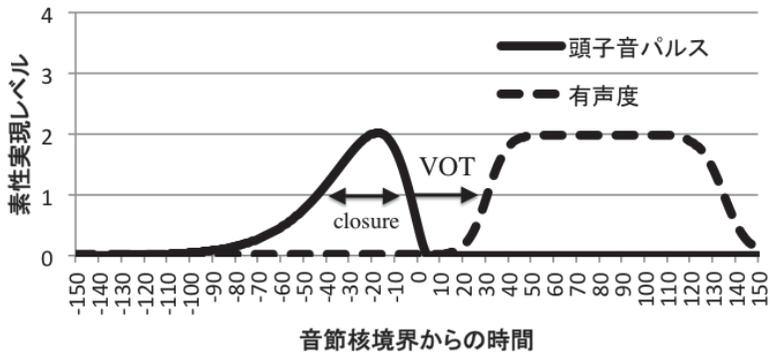


図 11: 弱い無声子音の周縁パルスがもたらす閉鎖区間と VOT

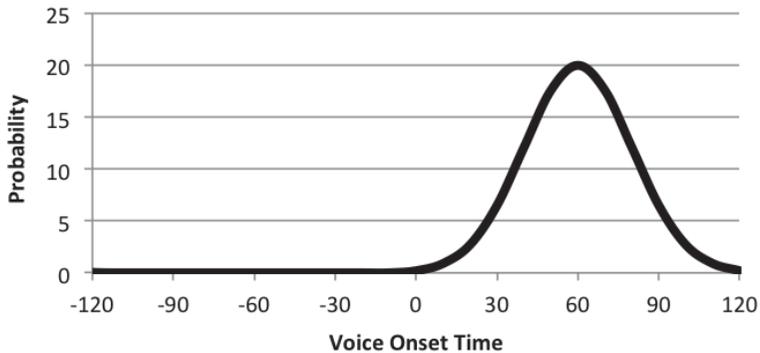


図 12: 無声破裂音の持つ VOT 分布のシミュレーション

図から分かる通り、本稿のモデルでは、Lisker and Abramson (1964) の実測値と同じく、無声破裂音の VOT は必ずプラスの値を持つ。すなわち、これは (2a) の定義から導かれる性質で、母音が持つ (正確には周縁パルスの終了時点に基づく) 有声性の開始時点は、必ず閉鎖区間が終わった後であることが保証されるのである。すなわち、C/D model において無声子音がプラスの VOT を持つ理由は、以下の要因の相互作用に基づく。

- (3) a. 完全閉鎖区間は必ず周縁パルス中に存在する (図 5 を参照)。
- b. (2a) の定義より、無声子音に後続する母音の有声性は、無声子音の周縁パルス終了時点が出る指令に基づく。
- c. この有声性の実現は、矩形パルス指令より数十ミリ秒遅れる (図 8 を参照)。
- d. この結果、無声子音の VOT がマイナスの値を持つことはない。

3. 有声破裂音における VOT のシミュレーション

3.1 一基準モデルのシミュレーション

多くの言語で、典型的な有声破裂音の VOT はマイナスの値を持つ。すなわち一般的には、閉鎖開放時点よりも早いタイミングで声帯振動が始まる。(2b) で述べたように、この性質を C/D model に当てはめるなら、有声破裂音における閉鎖区間のみを「声帯振動開始指令」の基準点と見なすことになる。これは、Lisker and Abramson (1964) による VOT の定義に類似しており、こうしたアプローチを「1 基準モデル」と呼ぶことにしよう。

1 基準モデルに従って有声破裂音の VOT を模擬してみると、VOT がマイナスの値から若干プラスの値まで広がるのが分かる。また、子音の強度が強いほど、VOT はマイナス値を取りやすい。その様相を少し詳しくみてみよう。まず (2b) の性質より、声帯振動の始まる許容範囲は図 13 における“closure”の時間帯となる。

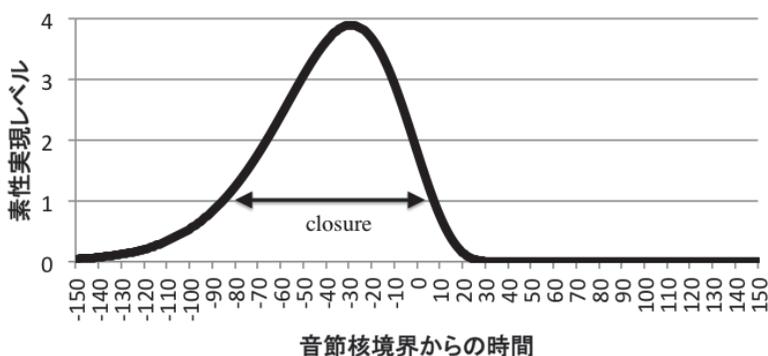


図 13: 有声破裂音における声帯振動開始時点の許容範囲

次に、有声子音の周縁パルスが強い場合を考えてみよう。図 14 に示す通り、最も VOT が大きくマイナス値を取るの、当然ながら完全閉鎖開始時点を生帯振動指令の開始時点とした場合である。また、完全閉鎖区間中に声帯振動が起こった場合にも、マイナス

の VOT 値 (図 15) を持つ。完全閉鎖の終了開始時点を声帯振動の開始時点とした時に、初めて VOT 値はほぼゼロか、あるいは若干プラスの値を取る (図 16)。

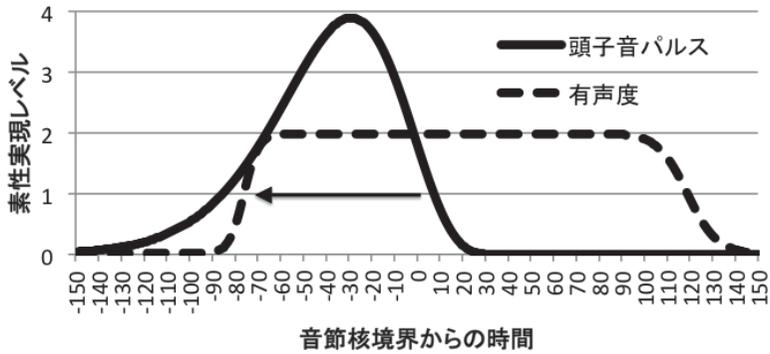


図 14: 周縁パルスが強い時の VOT 値 (VOT が大きくマイナス)

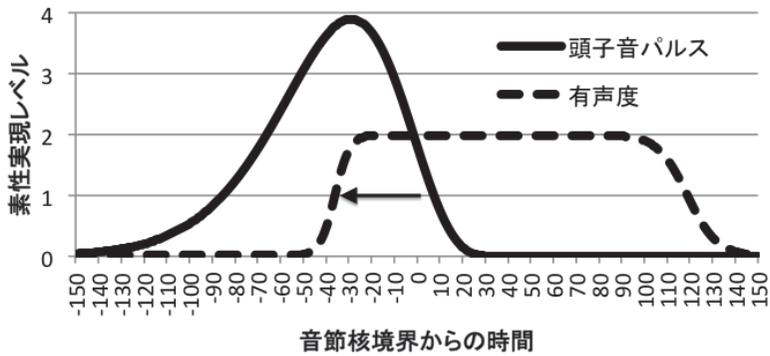


図 15: 周縁パルスが強い時の VOT 値 (VOT がマイナス)

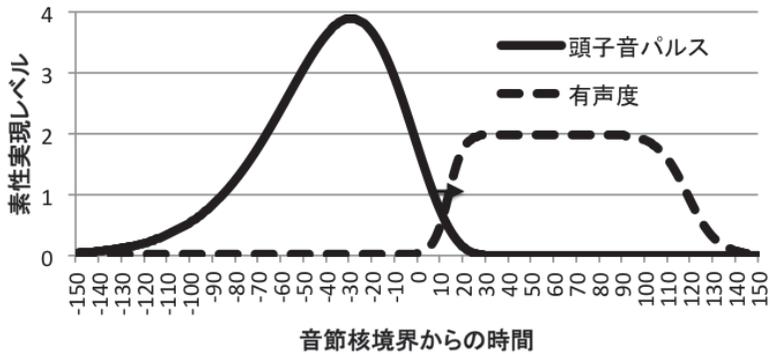


図 16: 周縁パルスが強い時の VOT 値 (VOT が若干プラス)

また、有声子音の周縁パルスが中程度の場合には、図 17～図 19 に示すように、VOT 値はある程度マイナスの値から若干プラスの値にまで分布する。ただし、いずれの場合も、周縁パルスが強い時よりも VOT 値がプラス方向にシフトしていく。

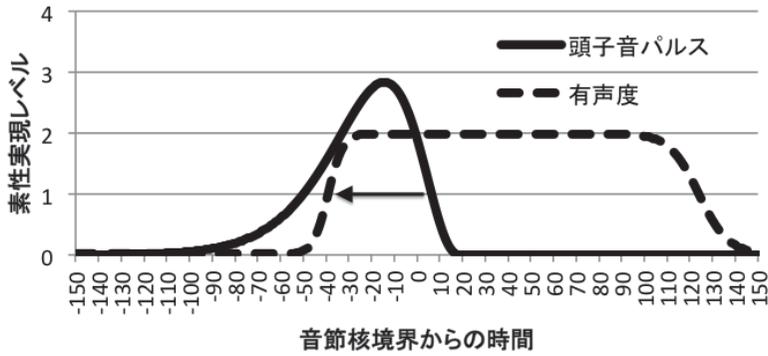


図 17: 周縁パルスが中程度の VOT 値 (VOT がマイナス)

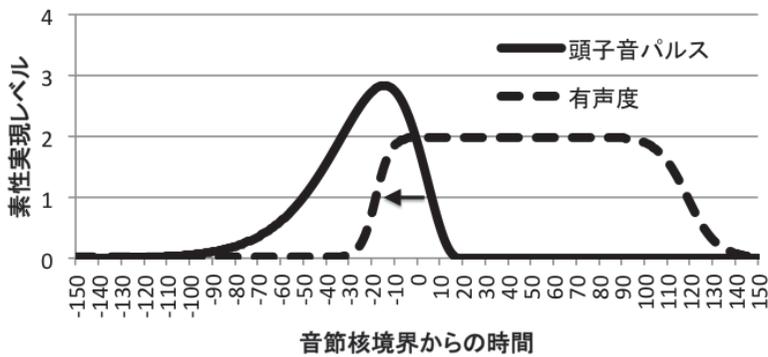


図 18: 周縁パルスが中程度の VOT 値 (VOT が若干マイナス)

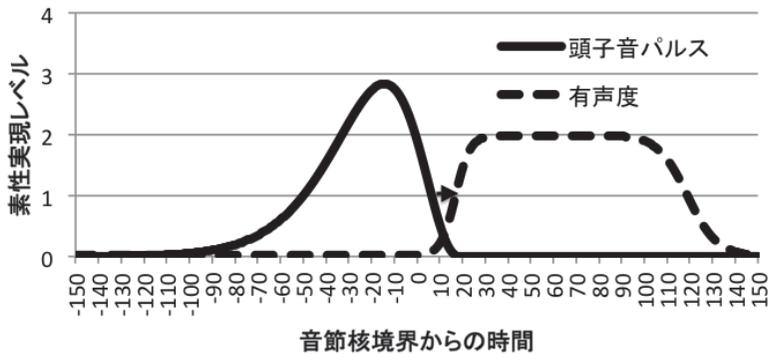


図 19: 周縁パルスが中程度の VOT 値 (VOT が若干プラス)

一方、有声音の周縁パルスが弱い時には、VOT 値はある若干マイナスの値から若干プラスの値にまで分布する。その様子を、図 20～図 22 に示す。

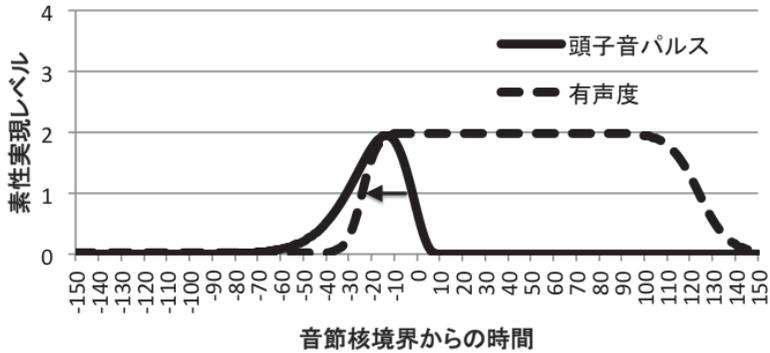


図 20: 周縁パルスが弱い場合の VOT 値 (VOT が若干マイナス)

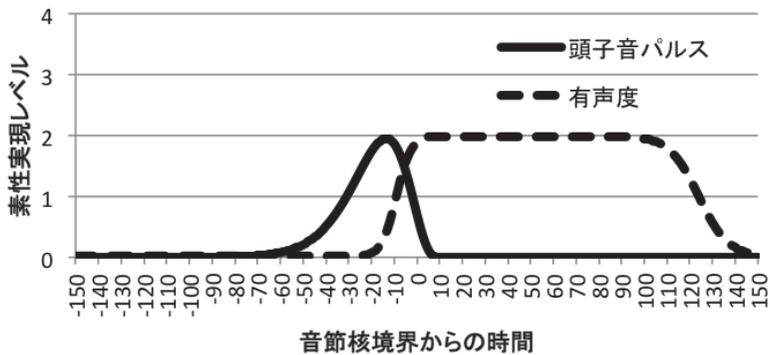


図 21: 周縁パルスが弱い場合の VOT 値 (VOT がほぼゼロ)

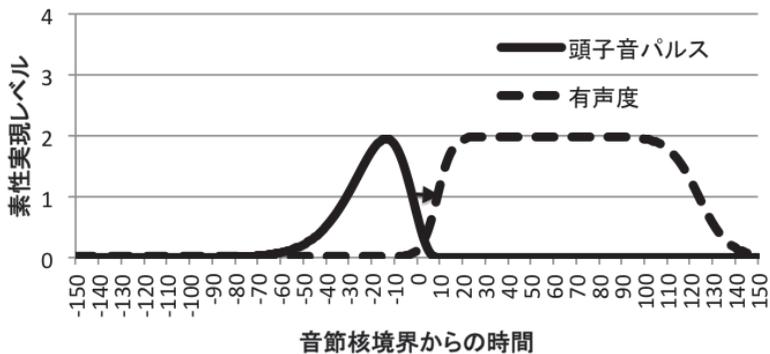


図 22: 周縁パルスが弱い場合の VOT 値 (VOT が若干プラス)

3.2 1 基準モデルに基づく VOT 分布

この1基準モデルにおいて、子音強度を様々に変化させ、またそれによって生じる閉鎖区間内で声帯振動が一様に生じると仮定して VOT の分布をシミュレートした結果が、図 23 である (-120msec 以上のデータは -120msec にまとめておく)。図 1 における有声子音の VOT 分布と異なり、図 23 のシミュレーション結果では、有声子音の VOT 分布が持つ最も重要な特性である双極化が見られない。言うまでもなく、0msec-dip も生じていない。論理的に言って、1 基準モデルで各条件が均等に生じる (一様分布に従う) と仮定する場合、あるいは正規分布に従って生じると仮定する場合に、分布の双極性が生じることはない。

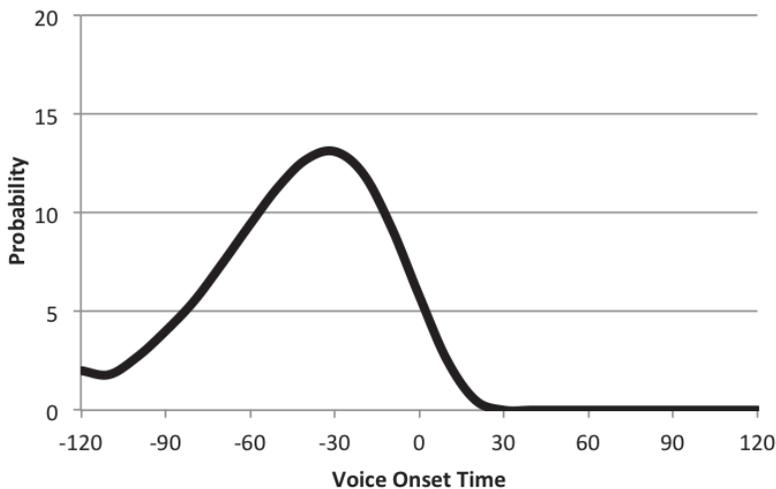


図 23: 1 要因モデルによる VOT 分布

また、高田 (2011) が主張しているように、VOT 分布は必ず双極化するとは限らない。方言によっては、東北方言のように VOT がプラスの分布に偏って出現することもあれば、年配の関西方言話者のように VOT がマイナスの分布に偏ることもある。1 基準モデルを単純に仮定するだけでは、こうした方言間の変異も説明できない。

1 基準モデルで「現実」の VOT 分布が 0msec-dip を起こすことを説明するためには、声帯振動が開放時点とほぼ同時に生起する (すなわち VOT がほぼ 0msec になる) ことを妨害する何らかの要因を仮定するしかない。今のところ、こうした妨害要因の実在自体が明らかではないが、もし妨害要因が存在するのであれば、それは妨害要因は声帯振動の開始時点が閉鎖解放時点に近いほど妨害を起こしやすく、また声帯振動の開始時点が閉鎖解放後であれば妨害を起こさないという性質を持つはずである。一応、本稿ではこうした条件に合う可能性のあるものとして、声門上下圧のエネルギーロスに比例する性質が声帯振動の生起を妨害していると考えしておく。他にも、破裂の瞬間に声帯振動の情報をピックアップすることが難しいといった知覚的要因や、破裂の制御と声帯の制御

を同時に行うことが困難であるといった生理学的要因も考えられ、こうした要因については今後の検討課題としておきたい。

こうした妨害要因を考慮し、周縁パルスの強さを強いもの・少し弱めのもの・弱いものという3つの段階で変化させた結果を、図24～図26に示す。薄い実線が妨害要因の関わらない時のVOT分布を、薄い点線が妨害の関与する相対的な強さ(実際には妨害の起こる頻度)を、濃い実線が「現実に観察」されるであろうVOT分布の形状を表す。周縁パルスの強さによって、分布の形状が大きく異なることが分かるであろう。周縁パルスが強い場合にはVOT値がマイナス側に偏って出現し、VOTがプラスになることはほとんどない。これに対し、周縁パルスが少し弱い場合には、VOT分布の特徴ともいえる双極性の分布が生み出されている。この双極性は周縁パルスが弱くなるほど特徴が消えていき、周縁パルスがかなり弱い場合にはVOTの分布はプラス側に偏っていく。

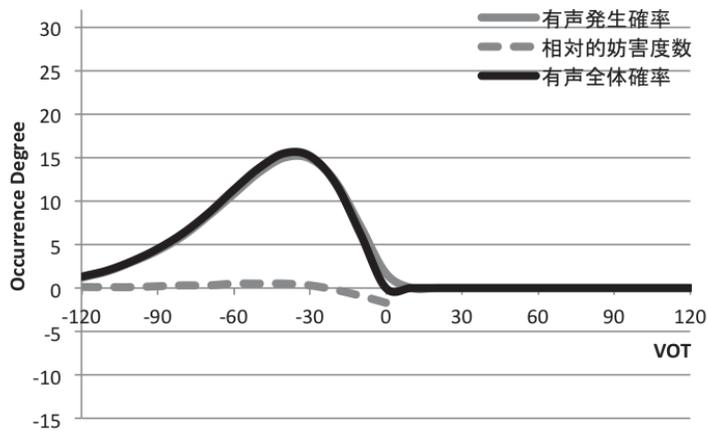


図 24: 周縁パルスが強い 1 要因モデルの VOT 分布

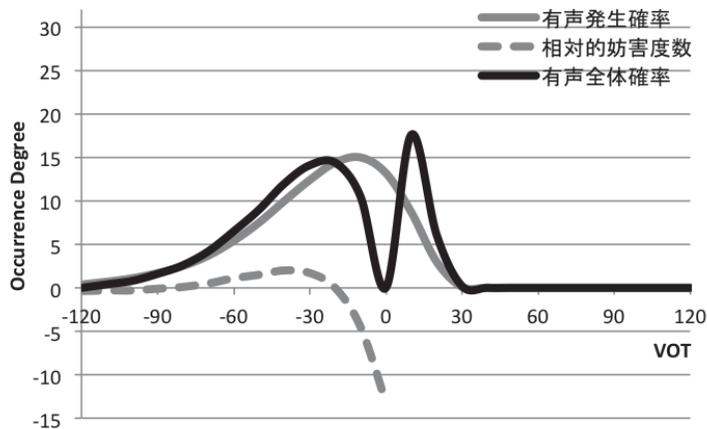


図 25: 周縁パルスが若干弱く、妨害要因が関与した VOT 分布

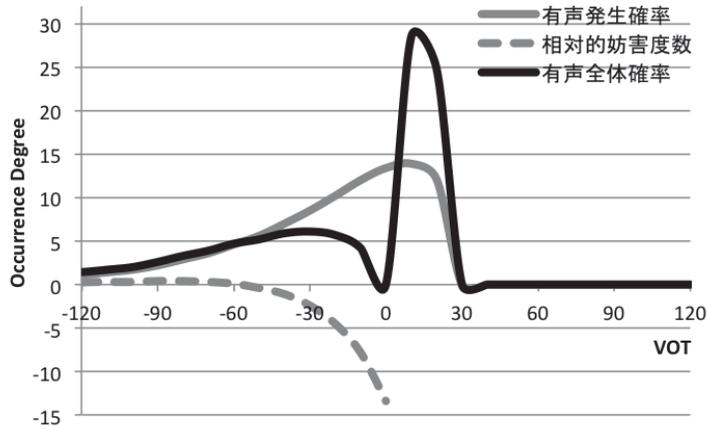


図 26: 周縁パルスが弱く、妨害要因が関与した VOT 分布

このように、1 基準モデルに妨害要因を加えた場合、高田 (2011) が報告している方言や世代間に見られる VOT 分布の変異パターンを、周縁パルスの強さによって一応模倣できそうである。しかし、そのためには、声帯振動の開始時点が閉鎖区間終了時に近くなるほど強度を増す妨害要因を仮定しなければならない。しかし、声帯振動を引き起こす声門上下圧のことを考えると、振動の開始時点が閉鎖開始時に近い時には妨害要因が働く可能性があるが、VOT がゼロに近いほど機能するような妨害要因の存在は想定しにくい。そこで次節では、VOT 分布の双極性や 0msec-dip を説明する別のアプローチを考えてみよう。

3.3 2 基準モデルのシミュレーション

双極化を起こす分布に対する最も自然な解釈は、その分布を引き起こす要因が単一のものではなく、複数存在すると考えることであろう。すなわち、VOT 分布が双極化するということは、有声子音における声帯振動の制御が 2 つの独立した基準点によって行われている可能性を示唆する。こうした 2 つの基準点によって VOT の生起を考えるアプローチを「2 基準モデル」と呼んでおく。このアプローチが正しいのなら、2 つの基準は各々 (4) を満たすようなポイントでなければならない。この基準を満たしたときに、はじめて 0msec-dip の効果を生み出せるからである。

- (4) a. 有声子音の VOT としてほぼ必ずマイナス値をもたらすポイント。
- b. 有声子音の VOT として若干のプラス値をもたらすポイント。

(4a) を満たす基準として最も可能性のあるものは、閉鎖区間の開始時点であろう。この閉鎖区間の始まるポイントを声帯振動開始時点の基準にした時に起こる VOT 値の概念図は図 14, 図 17, 図 20 に示されている通りであるが、これらの図を見ると、周縁パルスが弱くなるにつれて VOT は短くなっていくが、それでも VOT 値は基本的にマイナス値をとり続けることが分かる。何度も述べているように、周縁パルスの強さは潜在的に句

境界三角形によって決定される性質であり、句境界三角形は音韻過程によって決定される子音の強さを反映している指令である。したがって、(4a)を満たす基準点として閉鎖の開始時点を設定することは、有声性の基準として句境界三角形を用いることと同義であり、根源的には子音の強さが VOT の値に反映されるといってよい。

また、閉鎖開始時点を生帯振動の励起タイミングとした時に、VOT 値が 0msec になるのは、図 27 に示す通り、周縁パルスの強度がほぼ“1”になる時に限られ、この時の閉鎖持続時間は 15msec を切るために、もはや破裂音としての性質を保持できていない。以上のことから、閉鎖区間の開始時点は (4a) を十分に満たす基準であることが分かる。

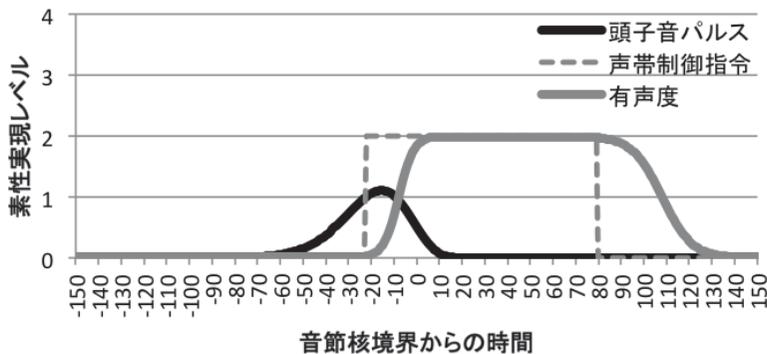


図 27: 閉鎖開始時点を基準とした時に VOT=0msec となる時の周縁パルス強度

一方、(4b)を満たす基準として有力な候補は、音節核開始境界であろう。このタイミングを生帯振動指令の開始基準点とすると、図 28 および図 29 に示すように、VOT は若干プラスの値を取る (図では周縁パルスが強い時には +4msec 程度、弱い時には +25msec 程度の値となっている) ことが言えるからである。

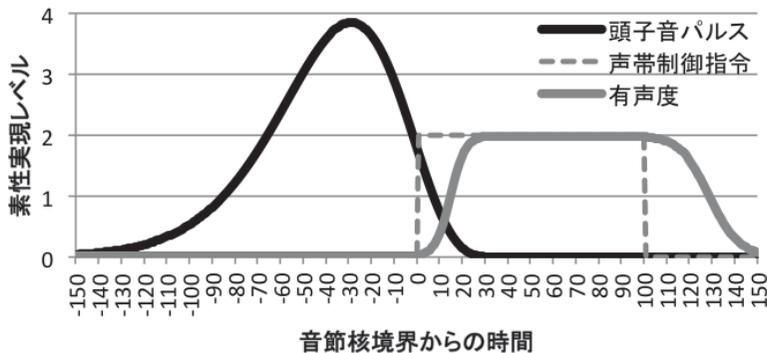


図 28: 周縁パルスが強く、かつ音節核開始境界を基準とした時の VOT

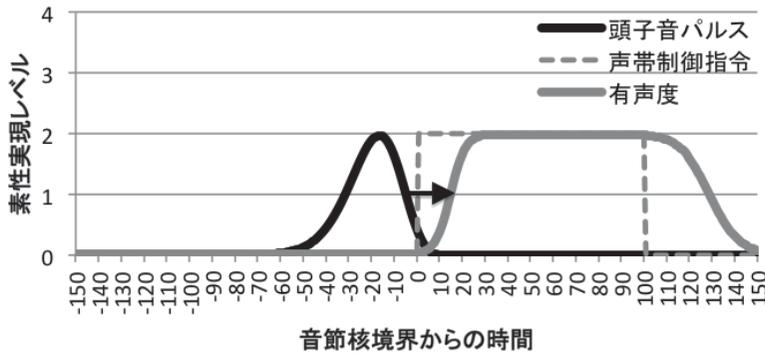


図 29: 周縁パルスが弱く、かつ音節核開始境界を基準とした時の VOT

また、周縁パルスが弱いほど VOT の値はプラス側にシフトしていくが、それでも図 12 に示す無声子音の VOT ほど大きな値にはならない。この有声・無声における VOT 値の違いは、無声子音に後続する声帯振動指令が周縁パルス終了時を基準にするのに対し、有声子音の(2つ目の)声帯振動指令は音節核開始境界を基準にするという点に起因する。すなわち、周縁パルスの終了時点は音節核開始境界よりも後にずれ込むことが一般的であるため、無声子音の VOT は有声子音の VOT より大きな値を持つのである。

以上の議論から、本稿では有声子音の有声性(声帯振動)が1つは声帯振動開始時点、もう1つは音節核開始境界を基準にして制御されていると考える。また、この2つの基準点は独立に用いられ、ある話者は声帯振動開始基準点を、別の話者は音節核開始境界をよく用いると仮定しよう。方言による VOT 分布の違いも、この2つの基準点を用いる割合が異なることが原因であると見なす。この時、2基準モデルでは妨害要因を仮定することなく、各方言における VOT の分布をシミュレートできる。様々な強度の周縁パルスを発生させながら、2基準モデルによって VOT 分布のシミュレーションを行った結果を、図 30、図 31、図 32 に示す。いずれのシミュレーション結果でも、前述したように無声子音の VOT 分布と大きく重複することはない。また図から分かる通り、音節核開始時点を基準点としてよく用いた場合には VOT がプラス側に偏った分布が生じ、閉鎖開始時点をよく用いた場合には VOT 分布がマイナス側に偏る。これらの条件下では、VOT 分布はいずれも単極性の分布を成す。高田(2011)のデータと照らし合わせるなら、前者は東北方言の、後者は年配の関西方言話者の VOT 分布に近い。

一方、多くの方言で見られる双極性の分布は、閉鎖開始時点と音節核開始時点のどちらの基準点も用いる場合に生じる。例えば、図 31 は閉鎖開放時点のほうを若干多く用いた時のシミュレーション結果であるが、若年層関西方言話者の実測値である図 1 に比較的近似していることが見て取れるであろう。若年の関西方言話者は高齢の話者ほど VOT がマイナス側に偏らないことが高田(2011)によって報告されているが、このシミュレーション結果は若年層の関西方言話者もどちらかという閉鎖開放時点をよく用いようとしていることを示唆している。

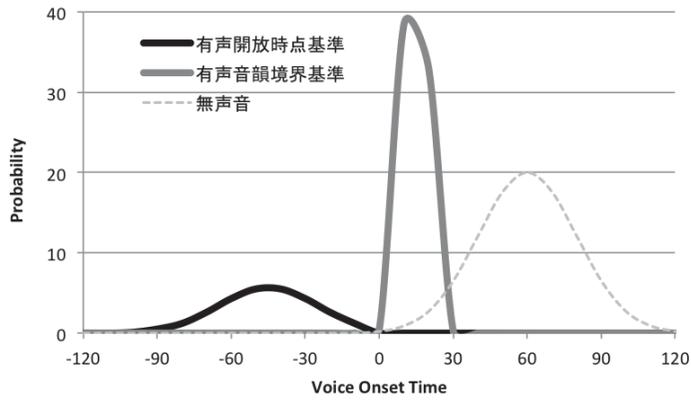


図 30: 音節核開始時点を基準値としてよく用いる場合

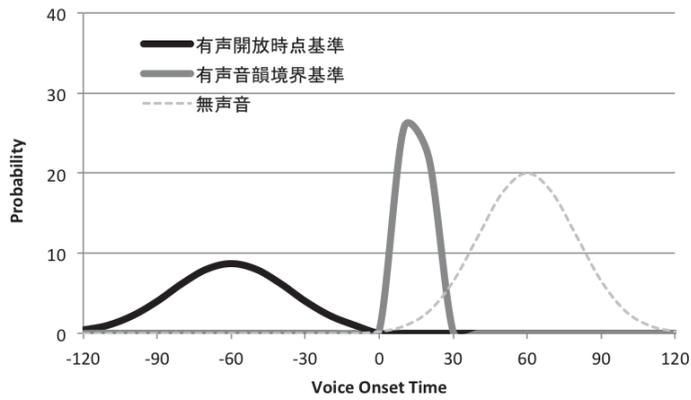


図 31: 両基準値を均等に用いる場合

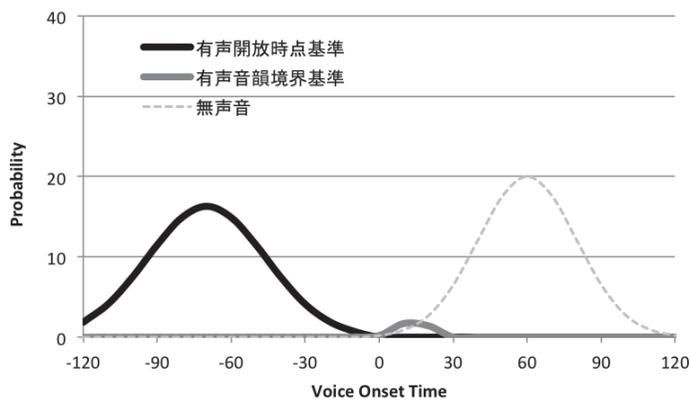


図 32: 声帯振動開始時点を基準値としてよく用いる場合

この2基準モデルが妥当であれば、子音の有声性判断における知覚方略について1つの示唆が得られる。声帯振動を生み出す矩形パルス指令の開始基準のうち、時間的により遅い基準点である音節核開始境界は潜在的な音韻境界—子音の終端境界であり、母音の母音の開始境界—を成す(以後、この境界をCV音韻境界と呼ぶ)。実際のCV音韻境界は、この音節核開始境界より後に生じるので、知覚上はCV音韻境界より前に声帯振動の物理特性であるボイスバーが生じていれば、その子音は有声子音と判断できることになる。言い換えるなら、有声性判断に関する知覚の手がかりは、VOTの定義である「破裂開放時点」ではなく、「CV音韻境界」である可能性が高い。2基準モデルの妥当性を確かめるために、次節では日本語の母音無声化や子音有声性判断における知覚の手がかりを探る実験を行い、CV音韻境界について検討を行う。

4. CV音韻境界と有声性

4.1 2種類の無声化母音

日本語の東京方言では、声帯振動が止まる環境に挟まれた狭母音—すなわち無声子音に挟まれた狭母音や単語末にある無声子音に後続する狭母音—がほぼ義務的に無声化を起こす。なお、非狭母音でも無声化を起こすが、その環境はきわめて限定されるため、本稿では狭母音の無声化にテーマを限っておく。前川(1989)は、こうした無声化を起こした狭母音の音声的実体について、「し」は[ɕi]ではなく、[ɕ]が延長したものであるとし、子音のみであるにも関わらず、無声化母音の認識が可能になる理由を以下のように説明している。これは、馬場(1997, 1998)や吉田(2006, 2008)らの見解とも整合性を持つ。

- (5) a. イ段とウ段で子音の音質が異なり、硬口蓋化が全てのイ段音節に認められる。したがって、子音だけでイ段音節をウ段音節と聞き分けることが可能である。
- b. 「シ・チ・ヒ」と「シュ・チュ・ヒュ」でも、実際の調音に違いがあり、拗音の子音のほうが後ろ寄りになっているため、混同の生じる恐れは大きくない。

これに対し、川上(1977)は、無声化を起こした狭母音について次のような観察を行っている。(6b)は前川の見解と同一であるが、(6a)に母音の存在を認めている点に注目されたい。

- (6) a. 無声子音の直前の「き、ぴ」「く、ぶ、しゅ、ちゅ」は無声母音 [i], [u] を持つ。
- b. 無声子音に先行する「し、ち、ひす、つ、ふ」は摩擦子音の延長である

前川(1989)と川上(1977)の見解は、「し」音と「す」音における無声化母音の実態が摩擦音の延長であるという点では共通点を持つ。しかし、「しゅ」と「ちゅ」における無声化母音の解釈は、川上が無声化「母音」を認めているという点で一致しない。ここで、実際の無声化母音を伴う単独単語発話「ニス」「二紙」「二種」の特性を見てみよう。図33では、前川(1989)の主張の通り、いずれも摩擦音が延長しているだけであり、母音の音響特性である調波構造は観察されない。また、「ニス」と「二紙・二種」の間では特性の

違いが存在するのはもちろんのこと、「二紙」と「二種」の間でも特性の違いがあり、「二種」の発音では特に摩擦成分の後半部において顕著な極周波数遷移が観察される。一方、「ニス」や「二紙」ではこうした明確な遷移が観察されない。

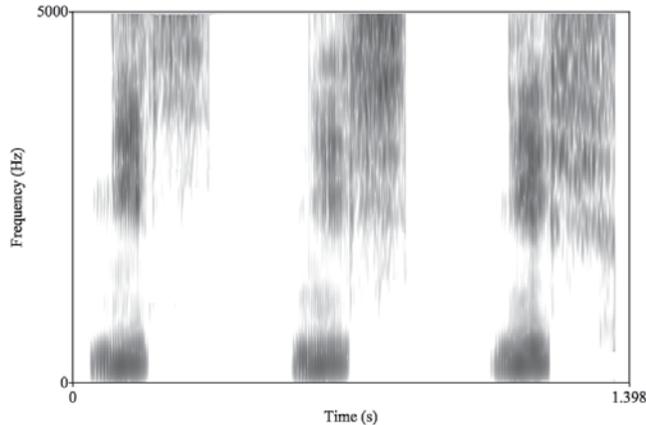


図 33: 無声化母音の音響特性 (左: ニス 中: 二紙 右: 二種)

この特性は (5b), (6) のいずれの見解にも矛盾しない。両者の妥当性を判断するには、無声化母音を伴う「シ／ス」の違いが極周波数のエネルギー帯域自体にあるのか、それとも極周波数の時間的遷移にあるのかを判断する必要がある。前者なら (5c) の見解が妥当であるし、後者なら (6) が妥当と言ってよいだろう。この点を検証するため、摩擦成分の極周波数とその時間遷移が無声化母音の知覚に与える影響について 3 種類の知覚実験を実施した。

4.2 実験 1 : 歯擦音に伴う無声化母音の知覚実験

はじめに、摩擦成分の極周波数の影響を調べるため、持続時間 120msec の母音 [o] と持続時間 160 msec の定常的な歯擦音で終わる合成音 [os]~[oç]~[of] を formant synthesizer によって作成した。歯擦部分の条件は、一番低い極周波数の中心周波数を 5200Hz~2400Hz まで低下させ、エッジ極周波数を 2800Hz~1400Hz まで低下させた S1~S8 までの 8 系列で、いずれも定常的な時間特性を持ち、極周波数遷移を持たない。摩擦部の特性を図 34 に示す。知覚実験は「オス・オシ・オシュ」の強制選択課題で、健聴者 23 名 (男性 4 名、女性 19 名) に対して、完全上下法を用いて実施した。

実験結果を図 35 に示す。摩擦成分が時間軸上の遷移を持たない場合、[s] 音と知覚された場合には無声化母音として /u/ 音が、[ç] 音と知覚された時には /i/ が一貫して認識されていることが分かる。極周波数が極めて低い [f] の範疇に入る音でも、時間的に安定した特性を持つ摩擦成分からは「シュ」の知覚がほとんど生起しない点に注目されたい。

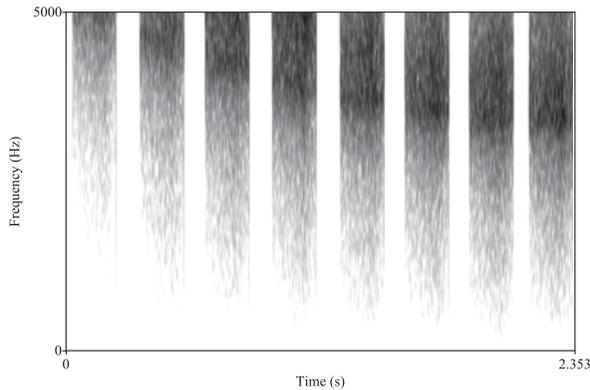


図 34: 歯擦部の音響特性

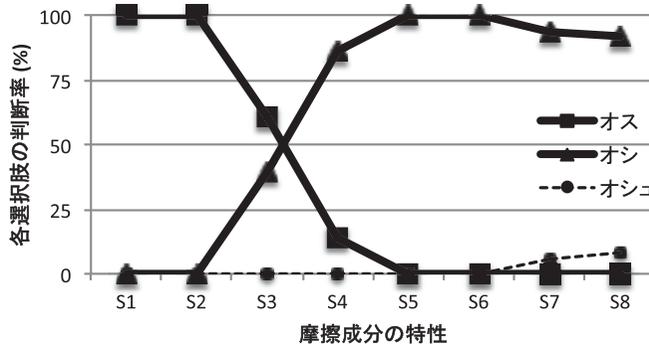


図 35: 定常的な歯擦音の知覚様式

4.3 実験 2：摩擦成分の遷移と CV 音韻境界

次に、[c] 音の極周波数が時間的な遷移を持っている場合に、無声化母音の知覚がどのように変化するかについて実験を行った。刺激音は実験 1 と同じく合成音 [oc] を用い、摩擦音の極周波数に時間遷移を付加した。極周波数の時間遷移の条件は以下の通りである：(a) 遷移開始時点は摩擦音開始後 100msec (遷移時間 60msec)。 (b) 遷移開始周波数は 3200Hz~2000Hz まで 2000Hz 刻みで変化 (7 条件)。 (c) 遷移終端周波数：遷移開始周波数から 1200Hz まで 200Hz 刻みで変化 (したがって時間遷移を持たない刺激音も含む)。図 36 に定常的な刺激音と極周波数遷移を持つ刺激音の例を示しておく。

実験は実験 1 と同一の被験者に対し、遷移開始条件を 1 系列とした完全上下法によって実施し、「オシ・オシユ」の強制選択を行わせた。実験結果を図 37 に示す。グラフ中の各線は遷移開始周波数を表し、横軸が遷移終了周波数となっている。

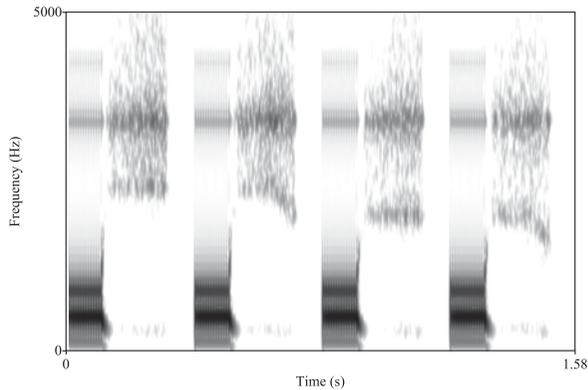


図 36: 刺激音の例

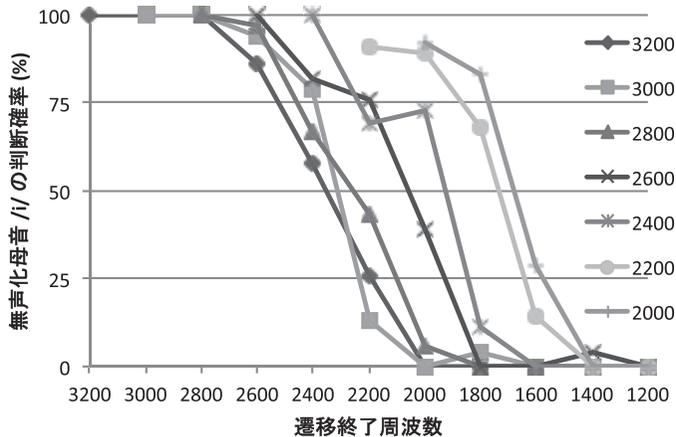


図 37: 時間遷移の影響に関する実験結果

実験 1 と同じく、極周波数の時間遷移がない場合には、「シュ」の知覚(すなわち無声化母音 /u/ の知覚)がほとんど生じていない。また、時間遷移があっても、遷移の終端周波数が 2400Hz 以上であればやはり「シ」と知覚されている。一方、終端周波数が 2200Hz 以下になると「シュ」の知覚が生起してくるが、それは極周波数遷移の開始周波数と終端周波数の関係に依存しており、今回の実験条件下では約 600Hz 以上の極周波数遷移がなければならない。ただし、周波数下降幅が重要であるのか、600Hz/60msec という傾斜が手がかりなのかという点は、今回の実験からは明らかでなく、今後の検討課題である。

ここで実験 1 と実験 2 の結果をまとめておくと、次のようになる。まず、倍音構造といった母音の音響特性を持たない音声であっても、摩擦成分の情報から無声化母音の知覚が可能である。これは、無声化母音の実体が摩擦成分の延長であるという先行研究を支持する結果と言ってよい。また、基本的に [s] 音からは無声化母音として /u/ が、[ç] 音

や [j] 音からは /i/ が知覚される。これは摩擦成分が時間的に安定しており、フォルマント遷移に代表される調音結合などの母音を示唆する音響特性が全くない音声であっても成り立つ。一方、[ɛ], [j] 音から /u/ 音の認識が生起するためには、低い極周波数という手がかり(すなわち円唇性の後続母音を予測させる手がかり)だけでは不十分であり、極周波数が必ず一定以上の時間遷移を持っていなければならない。言い換えるなら、音響情報自体に、周波数遷移という後続母音の存在を示唆する情報が物理的に残存していなければならない。

この結果は、川上(1977)のいう(6)と整合性を持つ。すなわち、川上が「シュ」に「無声化母音」を認めたのは、音響情報に後続母音を示唆する情報が必ず残存していることが影響したためだと考えられる。これに対し、「ス・シ」における無声化母音を「母音」と認定しなかったのは、これらのモーラにおいて母音の残存が要求されないからであろう。これは、国立国語研究所(1990)による日本語発音の X 線映画資料からも裏付けられる。図 38 を見ると、「ス」の発音ではこのモーラ全体で舌の運動がほとんど認められない。「シ」に関しては、舌の狭広の運動が若干起っているが、摩擦成分の極周波数遷移をもたらす前後の動態はやはり観察されない。しかし、「シュ」に関しては舌の明確な前後運動が起こっており、これは明らかに母音「ウ」が調音上も残存している証拠である。

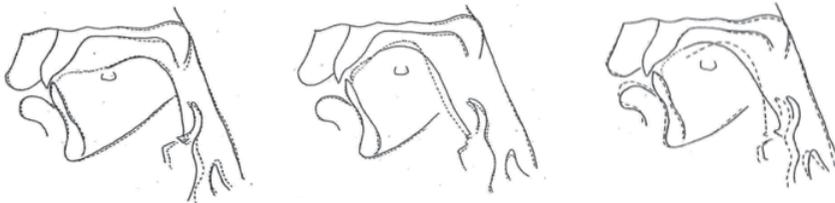


図 38: 日本語歯擦音の X 線資料(左:ス 中:シ 右:シュ)

国立国語研究所(1990)から引用

母音無声化現象において母音が残存するか否かという問題は、C/D model の観点からも興味深い。「シ」において母音が残存しない理由は、[ɛ] 音の持つ硬口蓋性と [i] 音の持つ前舌性が同化しているからであろう(このことから、[Palatal] 素性と [Front] 素性/[−Back] 素性は同一の素性で表した方がよいことになるが、C/D model における基底の素性表現については、また稿を改めて議論を行う)。こうして母音の自律性がほぼ失われた結果、シラブル三角形は非常に弱いものとなり、無声子音の句境界三角形によって生じる周縁パルスによって潜在的な母音指令全体が完全に覆われてしまう。図 39 に、母音の自律性が失われた場合の CD-diagram を示す。図からも分かる通り、無声子音後の母音については、周縁パルスの終了後に声帯振動の指令が出されるため、周縁パルスに比べてシラブル・パルスが極めて小さい場合には、声帯振動についても指令自体が全く出されない。これが「ス」や「シ」において母音が無声化する理由であり、かつ母音が残存しない理由である。

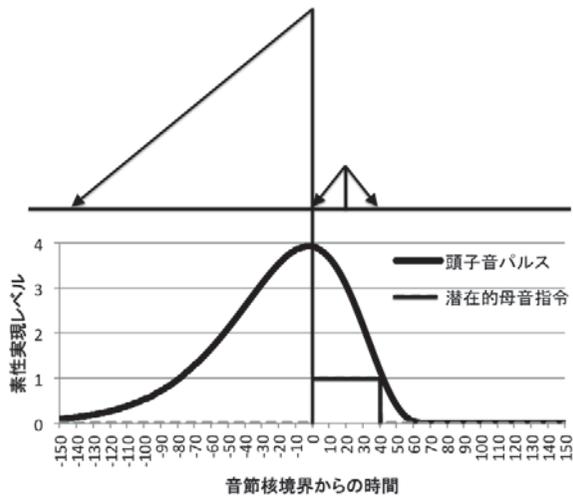


図 39: 無声化したス・シの CD diagram

これに対して、「シュ」の場合は子音の素性と母音の素性がぶつかり合うことがない。したがって、母音の自律性が保たれ、シラブル三角形は一定の大きさを保つ。しかしそれでも、語末における「シュ」はシラブル三角形自体の高さ（すなわち母音の強さ）が十分に大きくはならないため、周縁パルス終了時点で発令される声帯振動励起の指令時間は短くなってしまふ。語末の「シュ」ではない場合にはシラブル三角形は通常の高さを持つが、後続子音が無声音であれば、その周縁パルスがシラブル三角形の区域に浸食してくるため、やはり声帯振動励起の指令時間が十分に取れない。声帯振動の指令時間が長くない場合、図 40 の示されている通り、声帯振動の実現値が“1”に到達せず、収束していく。すなわち、物理的な声帯振動は起こらない。これが「シュ」において、母音が残存しつつも無声化の起こる過程であると考えられる。

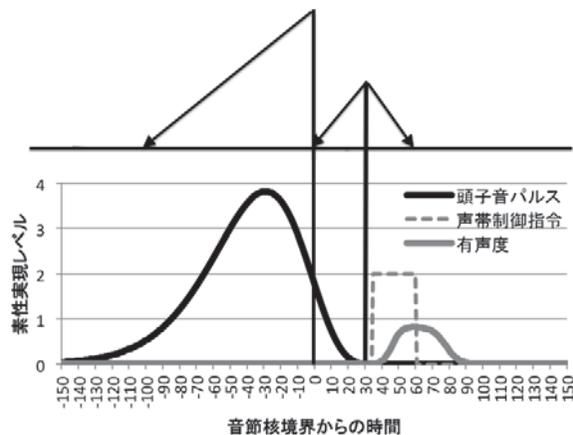


図 40: 無声化したシュの CD diagram

音声知覚という観点から言うなら、「シュ」における無声化母音の知覚は調音結合の手がかりとなる音響情報に基づいており、ボトムアップ処理に基づく通常の音声知覚過程と違いはない。これに対し、「ス・シ」における無声化母音の知覚は、音響情報自体に後続母音を示唆する情報がなくても可能であるので、これは音声知覚過程におけるトップダウン処理、すなわち音韻知識に基づいた音韻修復処理の結果であることを示唆している。こうしたタイプの無声化母音を、知覚的補間母音と呼んでおく。知覚的補間母音の決定過程については、(松井, 2012)などを参照されたい。

4.4 実験3：極周波数遷移と CV 音韻境界

前節で述べた「ス・シ」と「シュ」に関する CD diagram の説明が正しいなら、「シュ」における極周波数遷移中に子音と母音との音韻境界 (CV 音韻境界) が「知覚的」にも存在するはずである。そうした知覚的、この CV 音韻境界よりも前に存在する音響特性は摩擦子音に関する情報と判断され、境界後に存在する音響特性は、摩擦成分中に存在する情報であるにも関わらず、母音の情報と判断され、「摩擦子音」に関する情報としては用いられないであろう。例えば、CV 音韻境界以前に声帯振動が存在すれば「有声」摩擦子音の知覚が生起するが、声帯振動が CV 音韻境界以後に起こっている場合は、例え摩擦成分中に発生した声帯振動であっても、それは子音の情報とは見なされず、結果として「無声」摩擦子音に知覚されるはずである。この点を検証するために、図 41 に示すような刺激音を用いた知覚実験を行った。各々、遷移開始時点から声帯振動を持つ条件、遷移途中で声帯振動を起こす条件、遷移終了時に声帯振動が起こる条件の例である。

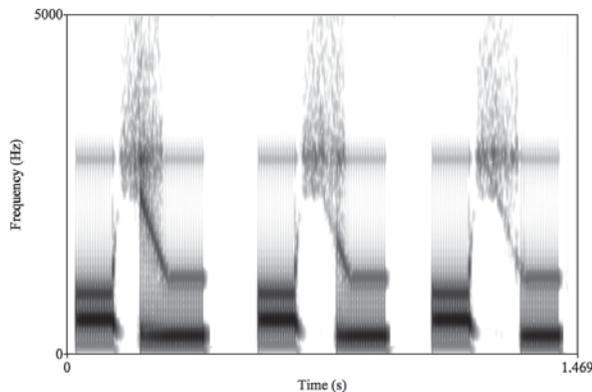


図 41: 刺激音の例

刺激音は、各音韻時間長が 120msec である [oεu]～[ozu] の連続体を合成した。摩擦部の条件は、2400Hz のエッジ周波数を持った定常的な特性が 40msec 続いた後、40msec～120msec の間で後続母音 [u] の第 2 フォルマント (1200Hz) に向かって極周波数遷移を起こすものとなっている。また、摩擦音中における声帯振動の開始時点を、摩擦開始後 40msec (すなわち極周波数遷移の開始時点) から摩擦音終了時点 (極周波数遷移の終了

時点) である 120msec まで 10msec 刻みで変化させた。知覚実験は健聴者 39 名 (男性 6 名、女性 33 名) に対して完全上下法を用いて実施し、「オシュ・オジュ」の強制選択を行わせた。実験結果を図 42 に示す。

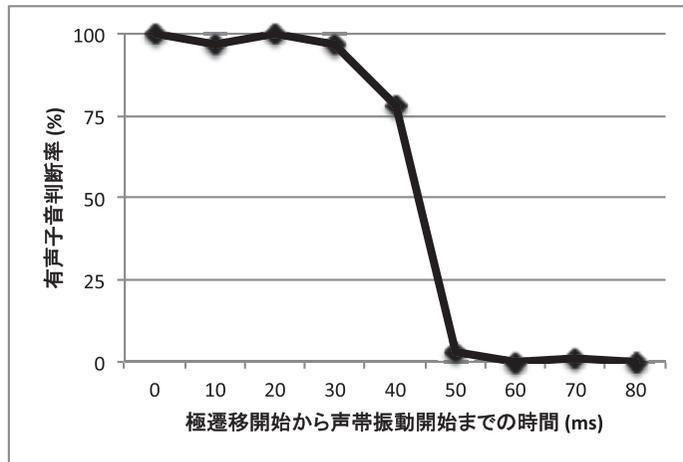


図 42: 時間遷移の影響に関する実験結果

図 42 から、有声摩擦音 [z] と無声摩擦音 [c] の判断境界は、遷移開始後 43.4msec (摩擦子音開始後 83.4msec) であることが分かった。これは、摩擦成分中に起こった声帯振動であっても、その開始時点によっては摩擦子音の性質に帰着されないことがあることを意味する。この知覚判断を左右する CV 音韻境界は、本実験条件下でいうなら約 1750Hz の周波数帯域で設定されており、一定以上の極周波数遷移は音声知覚過程において後続母音の音響情報と判断されていることを示す。したがって、極周波数遷移が伴う場合の「シュ」における無声化母音 /u/ の知覚は、トップダウンの補間ではなく、摩擦成分中に実在する音響情報に基づくボトムアップ処理の反映と見てよいであろう。

なお、本実験とほぼ同条件で、後続母音のない刺激や破擦音・破裂音を用いた刺激でも類似した実験結果となる。ただし、破裂音ではスパイクノイズとの相互作用が観察され、知覚様式が複雑であるため、詳細な結果は稿を改めて議論を行うことにしたい。

4.5 実験 4：借用語における促音抑制と CV 音韻境界

摩擦成分の極周波数遷移中に母音開始境界が設定されるという点については、他の音韻・音声現象と興味深い相関を持つ。本節ではその例として、借用語の促音挿入を取り上げてみたい。

借用語ではしばしば促音の挿入が観察される。促音挿入を引き起こす最も典型的な子音は、(7) に示すように、原語で抑止母音 (checked vowel) に後続し、語末の coda 子音となる無声破裂音・無声破擦音・無声摩擦音である。

- (7)
- | | | | | |
|----|-----|--------------|---|---|
| a. | p# | lip [lip] | → | [.ri <u>p</u> ^ˈ . <u>pi</u> .] ([rip:ui]) |
| b. | t# | mitt [mit] | → | [.mi <u>t</u> ^ˈ . <u>to</u> .] ([mit:oi]) |
| c. | k# | pick [pik] | → | [.pi <u>k</u> ^ˈ . <u>ku</u> .] ([pik:ui]) |
| d. | tʃ# | pitch [pitʃ] | → | [.pi <u>tʃ</u> ^ˈ . <u>tei</u> .] ([pitʃ:ei]) |
| e. | ts# | Mets [mets] | → | [.me <u>tʃ</u> ^ˈ . <u>tsu</u> .] ([metʃ:ui]) |

Coda position の子音促音化ほど典型的ではないが、強勢を持つ抑止母音に後続し、かつ consonant cluster を形成しない ambisyllabic な (Kahn, 1998) 無声子音についても、(8) に示すように、促音挿入が非常に起こりやすい。

- (8)
- | | | | | |
|----|------|-------------------|---|--|
| a. | [p] | happy [hæpi] | → | [.ha <u>p</u> ^ˈ . <u>pi</u> .] ([hap:i:]) |
| b. | [t] | motto [mátou] | → | [.mo <u>t</u> ^ˈ . <u>to</u> .] ([mot:oi:]) |
| c. | [k] | lucky [lÁki] | → | [.ra <u>k</u> ^ˈ . <u>ki</u> .] ([rak:i:]) |
| d. | [tʃ] | kichen [kítʃən] | → | [.ki <u>tʃ</u> ^ˈ . <u>tein</u> .] ([kitʃ:ei:n]) |
| e. | [ts] | Kratzer [kɾætsə] | → | [.kɾa <u>tʃ</u> ^ˈ . <u>tsa</u> .] ([kɾatʃ:sa:]) |
| f. | [ʃ] | cushion [kóʃən] | → | [.ku <u>ʃ</u> ^ˈ . <u>ɕon</u> .] ([kuʃ:ɕon:]) |
| g. | [s] | essay [éseɪ] | → | [.es <u>s</u> ^ˈ . <u>se</u> .] ([es:e:]) |
| h. | [f] | buffalo [bÁfəlou] | → | [.ba <u>f</u> ^ˈ . <u>fa</u> .ro:] ([baʃ:arou:]) |

ところが、(7), (8) に類似した環境であるにも関わらず、促音の挿入が阻止されてしまう現象が存在する。その代表的な例は「mag → マグ / *マッグ」などの有声阻害音によって引き起こされる現象で、有声阻害音は (7) と同様の環境であっても促音挿入が抑制されるか、あるいは「bag → バッグ / バック」のような揺れを持つ。これについては、日本語が有声促音を阻止する制約を持っており、和語や漢語で有声促音が基本的に許されないことが原因であり、Kawahara (2006) が興味深い分析を行っている。同様に、「beat → ビート / *ピーツ」のように長音の後にも促音挿入が起こらないが、これも和語・漢語自体が持っている制約 (1 音節 3 モーラ以上の超重音節を避ける * $\sigma_{\mu\mu}$ 制約) が影響しているためであろう。

しかし、借用語の促音阻止には、こうした日本語全体の音韻制約からは説明のつかない奇妙な現象も観察される。例えば、和語や漢語には「欲する」や「渴水」のように「ス音」の促音を持つ単語があることから分かる通り、日本語の音韻制約は「ス音」の促音を抑制する制約を持っていない (同様に、「muff → マフ」のように「フ音」でも促音が抑制されるが、ハ行子音は和語や漢語でも促音抑制の制約を持っているため、議論の対象から外す)。しかし、借用語では挿入母音 /u/ を伴う語末の /s/ 音において、(9a) に見られるように、(7) と同一の環境であっても促音の抑制が観察される。興味深いことに、挿入母音 /i/ を伴う [ʃ] 音も (9b) のように促音抑制を引き起こす傾向を持つ (「サッシ」のみが例外)。逆に、挿入母音 /u/ を伴う [ʃ] 音ではほぼ 100% の確率で促音挿入が起こる。

- (9) a. s# mass [mæs] → [.ma.su.] / * [.mas'.u.]
 class [klæs] → [.ku.ra.su.] / * [.ku.ras'.su.]
- b. sh# brush [bɹʌʃ] → [.bu.rɑ.ɕi.] / * [.bu.rɑɕ'.ɕi.]
 radish [ɹædɪʃ] → [.rɑ.de.ɕi.] / * [.rɑ.deɕ'.ɕi.]
- c. sh# mash [mæʃ] → [.mɑɕ'.ɕu.]
 clash [klæʃ] → [.ku.rɑɕ'.ɕu.]
 brush [bɹʌʃ] → [.bu.rɑɕ'.ɕu.]
 radish [ɹædɪʃ] → [.rɑ.dʲiɕ'.ɕu.]

Kubozono, Itô, and Mester (2009) は、こうした促音挿入／促音抑制に関わる現象を説明する際に、音節の重さに基づく制約である PROSODIC FORM (PROSFoRM) と、韻律外指定の考え方を導入した。PROSFoRM は (10) のような制約であり、借用語のみならず、漢語の「詩歌」や「夫婦」という軽音節＋軽音節の連鎖 ([.ci.ka.] や [.ɸu.ɸu.]) を持つ単語が、実際には [.ci:ka.] や [.ɸu:ɸu.] という重音節＋軽音節の連鎖として発音されるといった現象や、「女王 ([.dzo.o:.]）」という軽音節＋重音節の連鎖が、しばしば [.dzo:ɔ:] として発音される現象、また幼児語やズージャ語などにも働く制約であることが知られている。

- (10) PROSFoRM 制約：単語は、重音節-重音節 (HH) の連鎖か、あるいは重音節-軽音節 (HL) の連鎖で終わらなければならない。

Kubozono et al. (2009) の分析では、この PROSFoRM 制約が借用語の促音生起と促音抑制の両面に本質的な役割を果たす。まず (7), (8) で見た促音が起する場合の分析を示しておく。

- (11) a.
- | pick [pɪk] | *VoiGEM | *σ _{μμ} | PROSFoRM | *GEM |
|--|---------|------------------|----------|------|
| .pi.ku. | | | *! | |
|  .pik.ku. | | | | * |
- b.
- | peak [pi:k] | *VoiGEM | *σ _{μμ} | PROSFoRM | *GEM |
|--|---------|------------------|----------|------|
|  .pii.ku. | | | | |
| .piik.ku. | | *! | | * |
- c.
- | happy [hæpi] | *VoiGEM | *σ _{μμ} | PROSFoRM | *GEM |
|---|---------|------------------|----------|------|
| .ha.pi:. | | | *! | |
|  .hap.pi:. | | | | * |

ここで、(11) と同じ制約を、*mass* や *mash* という入力に対して掛けてみよう。結果は、(11a) と全く同様の計算により、促音の挿入された候補が解として選択される。これは、*mash* の借用形 [.mɑɕ'.ɕu.] としては正しいが、*mass* については促音のない借用形 [.ma.su.] を導き出せない。同様に、「ブラシ」に対しても、「ブラッシ」という促音形を予測する。

(12) a.

mass [mæs]	*VoiGEM	* $\sigma_{\mu\mu\mu}$	PROSFORM	*GEM
.ma.su.			*!	
*  .mas.su.				*

b.

mash [mæʃ]	*VoiGEM	* $\sigma_{\mu\mu\mu}$	PROSFORM	*GEM
.ma.s ^j u.			*!	
 .mas ^j .s ^j u.				*

c.

brush [b.rʌʃ]	*VoiGEM	* $\sigma_{\mu\mu\mu}$	PROSFORM	*GEM
.bu.ra.s ^j i.			*!	
*  .bu.ras ^j .s ^j i.				*

ここで Kubozono et al. (2009) は、単語末における /su/ の挿入母音 /u/ が韻律外要素となるという仮定を置く。この結果、語末の /su/ が音節性を喪失し、/s/ 音が先行する音節に組み込まれ、*mass* は (13a) のように促音の抑制される音型が最適であると正しく予測できる。Kubozono et al. (2009) では述べられていないが、韻律外指定というアプローチが正しいのであれば、/s^j/ の挿入母音が /i/ である時にも、(13c) のように挿入母音 /i/ が韻律外要素として指定されていると考えるべきであろう。逆に、/s^j/ に対して /u/ 音が挿入される場合には、この挿入母音 /u/ に対して韻律外指定が適用されてはならない。もしこの /u/ 音が韻律外指定を受けると、(13b) に示した通り、促音の入らない誤った音形を予測してしまう。

(13) a.

mass [mæs]	*VoiGEM	* $\sigma_{\mu\mu\mu}$	PROSFORM	*GEM
 .mas.<u>			*	
.mass.<u>		*!	*	*

b.

brush [b.rʌʃ]	*VoiGEM	* $\sigma_{\mu\mu\mu}$	PROSFORM	*GEM
*  .bu.ras ^j .<u>			*	
.bu.ras ^j s ^j .<u>		*!	*	*

c.

brush [b.rʌʃ]	*VoiGEM	* $\sigma_{\mu\mu\mu}$	PROSFORM	*GEM
 .bu.ras ^j .<i>			*	
*  .bu.ras ^j s ^j .<i>		*!	*	*

このように Kubozono et al. (2009) の分析において促音抑制に関わる根本的な要因は、韻律外要素の指定にある。彼らの議論では、この韻律外指定がいかなる動機で決まるのかが明確に述べられていないが、4.2 節・4.3 節で述べた現象が動機であると見てよいだろう。すなわち、極周波数遷移が存在せず、音響的に母音が残存していない「ス音」「シ音」では母音が韻律外指定を受け、調波複合音という音響的性質はないにせよ、極周波数

遷移という形で母音が残存している「シュ音」では母音の韻律外指定が回避されるということである。C/D model の観点から言うなら、「ス音」や「シ音」では図 39 のようにシラブル・パルスに基づく母音の特性が子音の周縁パルスによって完全にマスクされてしまい、調音としても・音響情報としても母音が残存しないために、その母音は韻律外指定を受けてしまう。一方、図 40 に示した「シュ音」のように、完全な母音情報ではないにせよ、シラブル・パルスの情報が何らかの機能を果たしている場合には、母音情報が残存するため、その母音は韻律外指定を受けない。このように解釈すると、Kubozono et al. (2009) の分析における韻律外指定は根拠のある妥当なアプローチとあってよいだろう。

彼らの分析でもう 1 点検討の必要な部分は、最終的に促音抑制を生じさせる原因を、超重音節を禁じる $*\sigma_{\mu\mu}$ 制約に求めている点である。確かに、「beat → *ビート」のような長音後の促音抑制には、 $*\sigma_{\mu\mu}$ 制約が重要な役割を果たす。しかし、「mass → *マッス」の抑制が、超重音節 “.mass.<u>” を回避するためであるという確証はない。この点を検証するため、定常的な摩擦部を持つ [os], [oɕ] および摩擦部に極周波数遷移を持つ [oɕ] を刺激音に用いて、促音の知覚様式を検討した。刺激音の例を図 43 に示す。

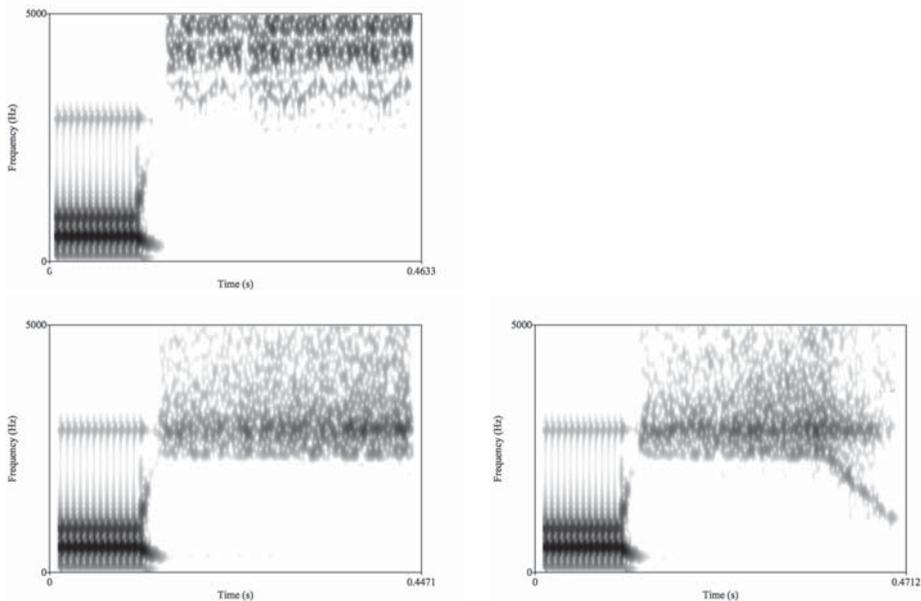


図 43: 疑似借用語の知覚実験における刺激音の例

左上：定常的な [os] 左下：定常的な [oɕ] 右下：極周波数遷移を持つ [oɕ]

先行母音 [o] の持続時間は 120msec で固定し、摩擦部の持続時間は 180msec～360msec まで変化させた。摩擦部に後続する調波複合音は存在しないため、これらの音声は日本語母語話者にとって全て「無声化母音」を持っているように聞こえる。そこで被験者 39 名に対し、「オス・オッス・オスー・オシュ・オッシュ・オシュー」の中から、最も適切

なものを強制選択させた。実験の結果を図 44 に示す。いずれも実線は促音に知覚された傾向を、点線は長音に理解された傾向を表す。

実験結果から、摩擦部の持続時間が長くなったときに、促音を知覚できるのは摩擦極周波数遷移を持った [oc:] のみであることが分かる。極周波数遷移を持たない定常的な [os:], [oc:] では、促音ではなく「無声化母音が長音化したもの」として理解される傾向が強い。すなわち、いずれの場合も刺激全体の時間長は正しく知覚されており、重音節のカテゴリに入ることは判断できているが、極周波数遷移がない場合には、それが「長子音 (促音)」にマップされるのではなく、「長母音」にマップされていく。しかし、極周波数遷移が存在する場合には、正しく「長子音 (促音)」として理解されるのである。

この結果は、「mass → マス/*マッス」「brush → ブラシ/*ブラッシ」における促音抑制が $*\sigma_{\mu\mu}$ 制約によるものではないことを示唆している。真の原因は促音知覚の困難さにあり、こうした知覚的制約が促音生起を抑制していると考えるのが妥当であろう。

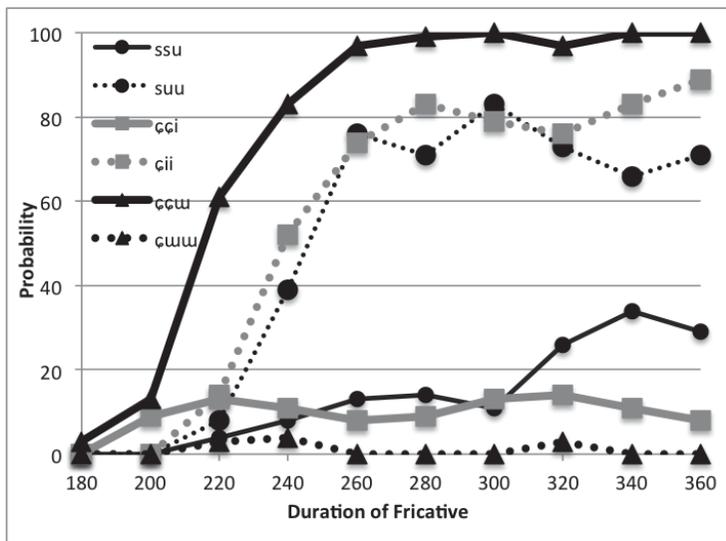


図 44: 疑似借用語の実験結果

Hirata (2007) をはじめとして、促音の知覚の手がかりには重子音部の持続時間と単語長との比率や重子音部の持続時間と先行母音との比率など、様々な指標が提案されている。しかし、そのような指標を取るにせよ、まず、促音が知覚されるためには、第一次情報として重子音部の持続時間が検出されなければならない。この持続時間検出の可能性こそが、本節で述べた実験結果や借用語の促音抑制を説明する原因と思われる。すなわち、イ音挿入を伴う [e] 音やウ音が挿入された [s] では、極周波数が定常的になりやすいため、図 45 に示すように CV 音韻境界を設定できない。

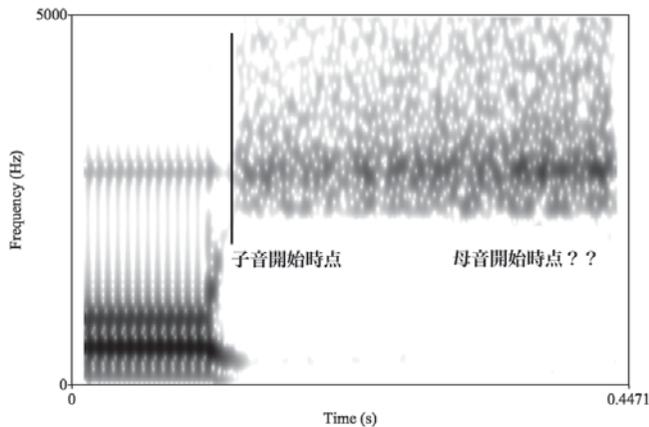


図 45: 語末「シ」音における CV 音韻境界設定の不可能性

CV 音韻境界が知覚されない限り、促音知覚に必要な子音持続時間を検出することもできない。この結果、「キス」や「ブラシ」において促音が抑制されてしまう。一方、「シュ」音の場合は CV 音韻境界の知覚の手がかりである摩擦成分の極周波数遷移が必ず存在し、重子音の開始部と終了部が明確になるため、図 46 のように重子音持続時間が確実に検出可能となるため、借用語の促音生起が可能になると考えられる。

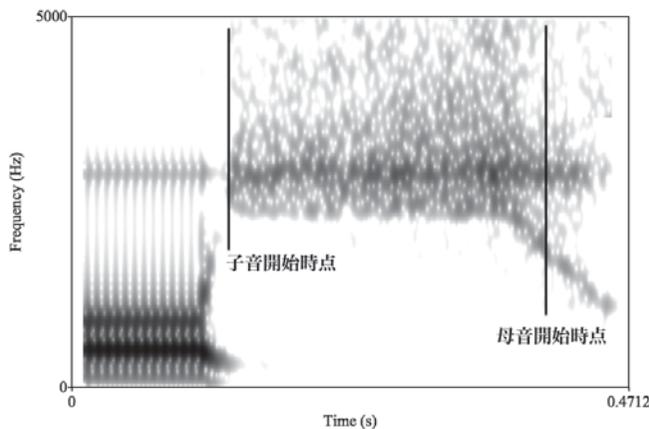


図 46: 語末「シュ」音における CV 音韻境界設定の成功

以上の議論から分かる通り、借用語における促音抑制と母音の無声化は深い関係を持つ。挿入母音が無声化を起こすこと、そしてその無声化母音の音響的実態が先行子音の持つ摩擦成分の延長であり、かつ摩擦成分の周波数遷移が起こらないこと、それによって促音知覚が困難になること—こうした要因が促音抑制を引き起こす。

この点に基づいて、松井 (2012) では、促音知覚の可能性に関する制約 CUEGEM によって、借用語の促音抑制と促音挿入の現象を分析している。この分析では、例え「ス音」であっても、「lesson → レッスン」「essay → エッセー」のような場合には促音化が可能である点も自然に説明がつく。こうした例では、「ス音」が無声化せず、その結果 CV 音韻境界がはっきりと設定されるため、促音知覚の手がかりが明確なのである。

(14) a.

miss [mis]	* \check{V}	* $\mu F\#$	CUEGEM	* $LH_c\#$	* $LL\#$	FtD	*GEM
$\text{mis}(\text{.mís}_\mu\text{.})$							
$(\text{.mí}_\mu\text{.s}_\mu\text{.})$					*!		
$(\text{.mís}'\text{.})\text{s}_\mu\text{.}$			*!			*	*

b.

dash [dæʃ]	* \check{V}	* $\mu F\#$	CUEGEM	* $LH_c\#$	* $LL\#$	FtD	*GEM
$(\text{.dá.}\xi\text{u.})$					*!		
$\text{dác}'\text{.})\xi\text{u.}$						*	*

c.

lesson [lesən]	* \check{V}	* $\mu F\#$	CUEGEM	* $LH_c\#$	* $LL\#$	FtD	*GEM
$\text{.re.}(\text{.s}_\mu\text{N.})$		*!		*		*	
$\text{.ré.}(\text{.s}_\mu\text{N.})$				*!		*	
$(\text{.rés}'\text{.})\text{.s}_\mu\text{N.}$							*

以上の議論から、子音の極周波数遷移中に子音・母音間の音韻境界 (CV 音韻境界) が設定されると見なしてよいだろう。また、実験 3 の結果から、この CV 音韻境界に基づいて、摩擦子音の有声性が決定されていることも明確となった。この結果は、声帯制御のタイミングとして CV 音韻境界が機能していることを示唆しており、2 基準モデルの妥当性を支持する間接的な証拠である。

5. 総合論議

本稿では、VOT 分布における双極性について、いくつかの側面から考察を行った。まず、この現象には 2 通りの説明が可能であることを述べた。1 つは、声帯振動の開始するポイントが閉鎖区間という 1 つの基準のみに依存しているという説明である。ただし、この考え方を取った場合、双極性のグラフを得るためには、閉鎖の開放時点に近づくにつれて何らかの妨害要因が機能するという仮定を置かなければならない。

もう 1 つの説明は、声帯振動の開始ポイントに 2 種類の基準が用いられている可能性である。C/D model に基づいて考えるなら、この 2 つの基準点閉鎖開始時点と音節核開始境界である可能性が高い。この立場を取れば、妨害要因を考慮することなく、各方言における VOT 分布の特徴を模擬できる。

そこで後者の妥当性を検証するために、いくつかの知覚実験を行った。実験結果から、(i) 摩擦子音の極周波数遷移は後続母音の実在であり、その知覚の手がかりであること、および (ii) それ故、遷移中に母音開始境界が設定されることが明らかになった。無声化母

音の音響的実体については、先行研究で議論されている通り、先行子音に含まれる摩擦成分の延長であることが知覚面からも検証された。ただし、「ス・シ・シュ」の無声化母音が区別可能な点については、川上(1977)の見解(6)が支持されたといつてよい。すなわち、「ス」と「シ」については、調音結合としての母音の手がかりさえ実在しない音響情報であっても無声化母音の知覚が可能である。一方、「シュ」の知覚においては、音響的には摩擦成分であるが、その音響情報に後続母音の情報が何らかの形で残存していなければならない。すなわち、「ス・シ」の無声化母音は知覚的補間母音であり、「シュ」の無声化母音は音声的に実在する母音である。ここで、「シュ」において厳密な意味で「後続」母音の情報が残存していなければならないという点に注意されたい。例えば今、母音の弁別素性が知覚過程で(15)のような音響特性と対応するとしよう。

- (15) a. [-High] : 約 370~900Hz に最初の音響特性が存在。
- b. [-Back] : 約 1800~3500Hz に音響特性が存在。
- c. [+Low] : 約 720~900 Hz に最初の特性が存在。

[s], [ç], [ʃ] の摩擦成分はいずれも(15a)を満たさないため、音響的に摩擦成分しか存在していない場合でも、その情報からボトムアップ処理で後続母音が[-High]の性質を持たないことが言える。また、[s]音は(15b), (15c)の特徴も持たないため、[s]音の音響特性は[-High], [-Back], [+Low]のいずれの性質も満たさない補間母音を導く。したがって、phonotactics等の音韻知識に駆動されるトップダウン処理に寄らず、ボトムアップ処理のみでも[s]音から後続母音/u/を同定できる。

しかし、[ç]音ではこの説明が成立しない。一般的な[ç]音は(15b)の性質を満たす。したがって、そうした[ç]音から導出される母音は[-High], [+Low]を満たさず、[-Back]を満たすもの、すなわち/i/音となる。これは本研究の実験結果と合う。しかし、非常に低い極周波数を持つ[ç]・[ʃ]音では、(15b)の特徴が満たされない。したがって、ここから導出される母音は[-High], [-Back], [+Low]を満たさない/u/音になるはずである。しかし、本実験の結果では、そうした[ç]音であっても、極周波数遷移がなければ/i/音の知覚が起り、/u/音は知覚されない。これは円唇性に伴う摩擦子音の極周波数低下が、後続母音を同定する手がかりとして使われていないことを意味すると共に、無声化母音の同定が単なる「ボトムアップ処理」で行われているわけではないことを示唆している。これに対し、「時間と共に」変化する極周波数遷移は、「音韻的に後続する」情報が音響情報にも実在していることを示す。だからこそ、こうした極周波数遷移を持つ[ç]音では、後続母音として/u/の知覚が可能となる。

こうした仮定は、図39および図40に示したCD-diagramによって説明がつく。また、借用語における促音抑制現象も、CV音韻境界の存在と深く関わっており、これも図39および図40に示したCD-diagramによって解釈が可能であり、いずれも2基準モデルの妥当性を補強するものといえるだろう。

2基準モデルのメリットは、声帯振動の開始時点が破裂時点のみならず、CV境界によっても制御されていると考える点にある。これは一見Liskerの定義よりも複雑になる

ように見えるかもしれないが、現実の VOT 分布を説明するためには、Lisker の定義を採用したとしても、別個に妨害要因を考慮しなければならない。二基準モデルでは、妨害要因を導入する必要はない。また、二基準モデルに「破裂の開放時点」が存在しない摩擦音の有声性についても破裂音と共通した基盤で議論を行えるようになるのである。これは子音の有声性について、これまでにない見方を与えることになるだろう。

言うまでもなく、CV 境界という新たな基準点を導入することは、2 基準モデルの重大な問題点でもある。したがって、このモデルの有効性を検証するためには、大規模音声データベースなどで子音・母音間の音韻境界を基準としたタグを振り直し、どの程度まで阻害音の有声・無声を区別できるかという点を検討しなければならない。また、Kawahara (2006) および 松浦 (2012) で取り上げられている有声促音の制御についても、2 基準モデルの考え方でシミュレーションができることを確認しなければならない。有声促音における half voicing, full voicing の現象は声帯振動の制御にとって本質的な問題の 1 つであり、これはモデルの正しさを検証する適切な材料であろう。

2 基準モデルにとって、もう 1 つの興味深い現象は、破裂子音の有声性判断に破裂開放時のスパイクノイズが影響を与えている可能性が見られる点である。これは現在進行中の実験の途中結果に過ぎないが、スパイクノイズが弱い場合には、VOT がある程度プラスの値になっても以前として有声子音の判断が続く。もしこの現象が確かなことであれば、知覚においても有声性判断の基準が複数用いられている可能性がある。この点についても、今後の研究課題としたい。

なお、そうした検証によって、もしも 2 基準モデルの妥当性が検証されたとしても、1 基準モデルが誤っているということにはならない。実際の声帯制御は、基準点を 2 種類使いつつ、かつ何らかの妨害要因が働いている可能性もある。また、方言によっては 1 基準モデルのような形で声帯制御を行っている可能性も否定できない。これらの点については、モデルの改良を含め、今後の課題としたい。

謝辞：本研究は 2013 年 6 月に国立国語研究所で行われた日本音声学研究会研究例会における発表原稿が元になっています。発表時に、貴重なご意見をいただきました窪園晴夫先生、松浦年男先生、高田三枝子先生、上野善道先生、Timothy J. Vance 先生に感謝いたします。また、本研究の一部は科学研究費・基盤研究 (C) 「認知的関連性のモデル化と文理解実験に基づく実証的研究」の援助を受けました。

参考文献

- 馬場良二 (1997). 日本語における無声化母音の音価. 『日本音響学会聴覚研究会資料』, pp. H-97-83.
- 馬場良二 (1998). 無声化母音は摩擦音である. 『熊本県立大学文学部紀要』, **50**, 1-22.
- Boersma, Paul (1998). *Functional Phonology: Formalizing the interactions between articulatory and perceptual drives*. The Hague: Holland Academic Graphics.

- Fujimura, Osamu (2002). Temporal Organization of Speech Utterance: A C/D Model Perspective. *Cadernos de Estudos Linguisticos, Instituto de Estudos da Linguagem, Campinas*, **43**, 9–36.
- 藤崎博也 須藤寛 (1973). 音声の音調的法則. 比企静雄 (編), 『音声情報処理』, pp. 123–142. 東京大学出版会, 東京.
- Fujisaki, Hiroya. & Hirose, Keikichi. (1984). Analysis of voice fundamental frequency contours for declarative Sentences of Japanese. *Journal of Acoustical Society of Japan (E)*, **5** (4), 233–242.
- Hirata, Yukari. (2007). Durational variability and invariance in Japanese stop quantity distribution: Roles of adjacent vowels. *Journal of the Phonetic Society of Japan*, **11** (1), 9–22.
- Kahn, Daniel (1998). Syllable-based Generalizations in English Phonology. Indiana University Linguistics Club.
- Kawahara, Shigeto (2006). A faithfulness ranking projected from a perceptibility scale: The case of voicing in Japanese. *Language*, **82**, 536–574.
- 国立国語研究所 (1990). 『日本語の母音, 子音, 音節: 調音運動の実験音声学的研究』. 秀英出版, 東京.
- Kubozono, Haruo., Itô, Junko., & Mester, Armin (2009). Consonant Gemination in Japanese Loanword Phonology. In *Current Issues in Unity and Diversity of Languages. Collection of Papers Selected from the 18th International Congress of Linguists*, pp. 953–973. Dongam Publishing Co.
- Lisker, Leigh & Abramson, Arthur S. (1964). A Cross Language Study of Voicing in Initial Stops: Acoustical Measurements.. *Word*, **20**, 384–422.
- 前川喜久雄 (1989). 母音の無声化. 杉藤美代子 (編), 『講座 日本語と日本語教育第2巻』, pp. 135–153. 明示書院, 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.
- 松井理直 (2012). 借用語における促音生起の抑制要因. *Theoretical and Applied Linguistics at Kobe Shoin*, **15**, 49–102.
- 松浦年男 (2012). 宮古諸方言の音声実現に関する予備的検討. 『「消滅危機方言の調査・保存のための総合的研究」宮古島方言調査報告書』. 国立国語研究所.
- Pierrehumbert, Janet. & Beckman, Mary. (1988). *Japanese Tone Structure*. The MIT Press, Cambridge.

川上 稔 (1977). 『日本語音声概説』. 桜楓社.

高田三枝子 (2011). 『日本語の語頭閉鎖音の研究』. くろしお出版.

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.

吉田夏也 (2006). 日本語母音無声化の音声学的研究. 『音韻研究』, **9**, 173–180.

吉田夏也 (2008). /p/に後続する無声化母音を知覚する音響上の手がかり. 『音声研究』, **12** (3), 52–58u.

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

(受付日: 2014.1.10)