

名古屋方言話者における破裂音の発音実相

—— Voice Onset Time の分析を通して ——

城 哲 哉

0. はじめに

個別言語における音韻対立が音声としてどう具現化されるのか、その生成メカニズムを探ることは音声研究の重要な研究課題であり、言語音のうち特に破裂音の生成に関しては、Voice Onset Time (Lisker & Abramson, 1964; 以下 VOT) を中心として盛んに研究が進められている。VOTは、有声/無声の範疇化を決定づける音響要因として研究が始められたものであり、具体的には、声道閉鎖の開放から声帯振動開始までの時間差を計測することで、言語固有の調音タイミングの異同を表示したものである。VOTの計測値によって、(i) 無気有声音、(ii) 無気無声音、(iii) 有気無声音の3つの範疇化が可能であることが知られており、多くの言語においてその有効性が検証されてきた。

VOTは範疇化に関しては、普遍的な特性をもつ反面、男女差や加齢あるいは方言の差などにより、言語内部で変動性の属性をもつことも分かっており、母語の獲得や障害(失語症等)また第二言語における習得段階のスケールとして用いられることも多い。最近では、Lisker & Abramson (1964) で提案された初期の範疇化モデルに修正を迫る研究(Cho & Ladefoged, 1999)も現れており、音声言語の普遍性と個別性に関してVOTを使った研究は非常に多い。

日本語におけるVOTは、Homma (1980, 1981), Shimizu (1996) などの研究によって

広く知られているが、近年、従来の研究とは異なる日本語のVOTの実態を報告したもの(高田, 2004)や、他言語と比較して日本語の無声破裂音の特異性を強調した興味深い論考(Riney et al., 2007)なども現れてきている。筆者の知る限り、名古屋を中心とする東海地方で話されている日本語方言に関してのVOT研究はほとんど行われておらず、その実相は不明である。従って、本稿では、(i) 名古屋地区に在住する大学生から収集した音声の分析を通して、この地区で話されている日本語のVOTの実相を提示し、それに基づいて(ii) 今回のデータがもつ意義を関連論文に触れながら論じる。

1. 本稿の論点

本稿は名古屋方言(広義での)のVOTの実相を報告することを第一の目的とするが、特に次の3点を論点とする。

(1) 東京において、大学生世代を中心として有声破裂音の調音タイミングが、従来の声帯振動先行型(−VOT)から、破裂の開放後に声帯振動が始まる型(+VOT)に変わりつつあることが報告(高田, 2004)されているが、名古屋方言話者の場合はどうか。

(2) もしそのような変動が見られるならば、無声破裂音の調音タイミングはどうなっているか。これまで日本語の無声破裂音は短い形式

(+VOT) であると報告されてきたが、有声破裂音が新しい形式(+VOT)を取ることで、VOTの時間尺度上での2つの範疇の棲み分けはどのような意味をもつことになるか。

(3) これまで日本語/p, t, k/は有気性(aspiration)がなく、短いVOTをもつ(short-lag)無声無気破裂音とされることが多かったが(Tsujimura, 1996など)、他言語との比較でこれに修正を求める論考が発表されるようになった(Riney et al., 2007)。今回の名古屋方言のデータやこれまでの研究から見て、この修正案をどう考えるべきであるか。

データ分析の結果を示したあとで、以上の点を中心に考察する。

2. データの収集と分析

分析に用いた音声データは、名古屋市内にある大学の医療福祉学部言語聴覚コース2年次に在籍する41名の学生(女性36名、男性5名)から収集した。34名が東海3県(愛知、岐阜、三重)の出身で、残りの7名が北海道、東京、静岡、滋賀、徳島、島根の各県の出身者であった。録音は2年次の学年末に行ったことから、遠方から来た学生も2年間はこの地域に居住していたことになる。言語生育歴のアンケート調査も合わせて行い、方言の影響などにもできる限り注意を払うようにしたが、今回は均質な言語グループと仮定して論を進める。なお、今回の発話者の中には海外居住経験者はいない。

データ収集は、破裂音を語頭にもつ「ぱす、ばす、たす、だす、かす、がす」の6個の語句を「そして_____と言います。」というフレーム文に入れて発音することで行った。特に注意点として、直前の母音の影響を回避し、先行有声化による声帯振動の時間長を適切に計測する

ために、「そして」の後ろにポーズを取るように事前に指示した。録音は一語句につき一回のみ録音したが、言いよどみや音声として不明瞭な場合は、直後に録音のし直しを行った。

録音は大学の防音室で行い、単一指向性のコンデンサーマイクロフォン(Sony, ECM-959DT)を被験者の口元から15cm程度離れた位置に置き、DAT(Sony, TCD-D10)を使って録音した。収集した音声は、USB接続のA/D変換ボックス(Roland, Edirol UA-3)を通して、サンプリング周波数22,050Hz、量子化16ビットでコンピュータに取り入れ、SFS/WASP(Mark Huckvale at University College London, ver. 1.41)を利用して音響分析を行った。VOTの計測は、スペクトログラムを同時表示させながら音圧波形より行い、破裂のスパイク(transient)より声帯振動を示す規則的な周期波形の開始までとし(計測は、最初の山の振幅最大地点まで)、またVOTがマイナス値を取る場合は、周期波形の出現点より破裂のスパイクまでを計測した。

計測上の問題点として、前稿(城, 2006)でも論じたことであるが、有声破裂音の調音などにおいて、口腔の開放前の閉鎖継続中に声帯振動が出現し終息する場合がある。実際に今回のデータにおいても有声破裂音の11%(123個中13個)にこの形式の先行有声化(prevoicing)が観察された。本稿は、特にVOTの先行研究との比較を目的とするため、よりオーソドックスな計測方法に従い、これらの場合は声帯振動の開始を破裂の後にもつ無声破裂音と判断して計測を行った。また軟口蓋音/k/に時折見られる複数破裂(double spike等)は、最初の破裂を開放の瞬間とした。

3. 結果と考察

3.1 分析結果

VOTの測定値結果をまとめたものが表1である。古くから指摘されている通り（Fischer-Jørgensen, 1954; Peterson & Lehiste, 1960），調音点が口腔の奥に下がるにつれて破裂から声帯振動までの時間が延びる（VOTの上昇）という言語一般の傾向が今回のデータからも確認ができる。また，上昇の幅の大きい軟口蓋音とは異なり，両唇音と歯音ではVOTの差が小さいこともよく知られており（Lisker & Abramson, 1970; Cho & Ladefoged, 1999），今回のデータでもその差は僅かであった。

今回最も顕著な結果として判明したことは，表1から分かるように，有声破裂音のVOTが+/-の2系列に分かれることであった。41名のうち，有声破裂音を一貫して+VOTの形式で生成する発話者が19名と最も優勢であり，現在まで一般的と考えられていた破裂の前に声帯振動が伴う-VOTによる発話者は6名に過ぎず，残りの16名はVOT +/-を調音点によって混在させていることが判明した。分析を行った123個（3個×41名）の有声音が，具体的にはVOT +/-のどちらの値を取ったかを調音点別に表2としてまとめた。

VOTの一から+への移行は無声化の傾向が強いことを意味する。従って表2から，破裂音 /b, d, g/ の中では /g/ が最も無声化の傾向を受け易く， /b/ が最も無声化の影響を受けにくいことが予想される。この /g/ が /b/ と比べて有声

表2 有声破裂音における + VOT / - VOT の生起個数（各子音 N = 41）

	/b/	/d/	/g/
+ VOT	23	25	32
- VOT	18	16	9

性と結びつきにくいことの例証として，清水（1998）は，タイ語で有声軟口蓋音 /g/ が欠如していることや，諸言語の音声レパートリーを網羅したデータベース UPSID（UCLA）でも /g/ の発生率が /b/ に比べて低いことなどを指摘している。また破裂前の声帯振動の終息に関して，特に /g/ は軟口蓋閉鎖に伴う舌の後方動作が空気の逆の流れ（*ingressive or negative flow*）を生じさせ，空気力学的にも声帯振動の持続に影響を与えることも分かっており（Isschiki & Ringel, 1964），このことも /g/ の無声化と関連していると考えられる。

有声破裂音を + VOT の形で調音する話者が，どのような調音タイミングで一方の無声破裂音の生成を行うかは，特に日本語のような短い VOT の言語では興味深い。表3は，今回そのような調音形式を取った19名のVOT値の平均と無声/有声ペア群（/p/-/b/, /t/-/d/, /k/-/g/）の差を検定した結果である。また，図1に平均値および標準偏差を棒グラフで視覚的に示した。

軟口蓋音 /k/ を例外として，今回のデータは，他の5つの音すべてがVOTの短い+の領域（0～30 ms）に収まっている。一般に *short lag* と呼ばれるこの領域は，無気無声破裂音が占める位置と考えられており，音韻的対立をもつ

表1 VOTの平均値（msec）と標準偏差

	/p/	/t/	/k/	/b/	/d/	/g/
Mean				12.4 (5.9)	13.6 (4.3)	23.5 (5.1)
(SD)	23.8 (8.0)	25.2 (8.4)	47.8 (11.4)	-62.6 (29.9)	-61.6 (34.0)	-69.8 (22.8)

表3 +VOT話者グループのVOT平均値 (msec) とt検定

Plosives	Mean (SD)	t-test (two-tailed)
/p/	25.3 (7.0)	$t(18) = 5.89, p < 0.001$
/b/	13.6 (5.7)	
/t/	26.3 (8.6)	$t(18) = 6.32, p < 0.001$
/d/	13.1 (3.0)	
/k/	50.1 (12.1)	$t(18) = 8.29, p < 0.001$
/g/	24.1 (4.4)	

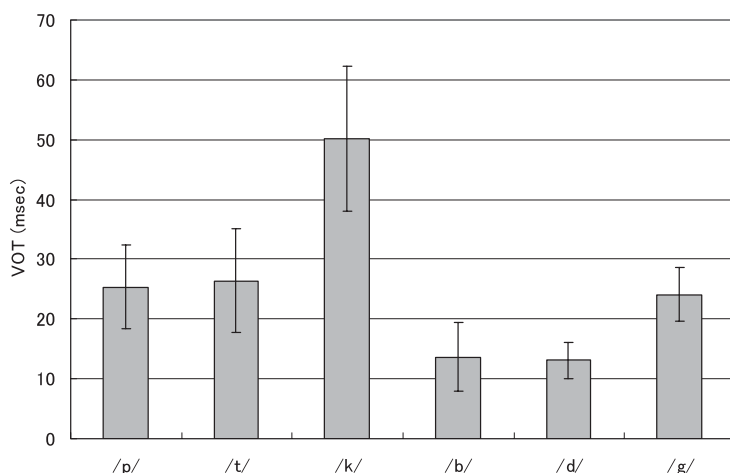


図1 +VOT話者グループのVOT (平均値±SD)

/t/-/d/が同時にこの数値内に納まることは弁別
の面でも不都合を生じることが予想される。し
かし、このデータでは、無声/有声ペア間の差
は統計的に有意であり、短い領域内に有声/無
声の両範疇が共存していることを意味する。こ
のような範疇間の差が非常に小さい共存関係
は、これまでのVOT研究ではあまり報告され
ておらず、稀有な現象であると考えられる。
似た問題として、韓国語における濃音と平音
の対立がVOTの研究が始まった当初から議論
されているが、Shimizu (1996) の研究を見
ると、2つの音のVOT値はオーバーラップして
おり、今回のデータのように尺度内での整合性
(数値上の棲み分け) は取れていない。Lisker

& Abramson (1964) のVOT尺度を使った範
疇化の適切性が問われ始めているが (Cho &
Ladefoged, 1999), この日本語有声破裂音の+
VOT化の現象も興味深い問題を投げかけてい
る。

3.2 考察

本稿の考察課題として最初に述べた3つの事
項について順に検討を行う。

まず、+VOT化の問題であるが、この地方
の大学生の破裂音においても確実にこの変化が
進行していることが確認された。分析の対象
とした123個の有声破裂音のうち2/3が+VOT
で生成されたものであった。Homma (1980)

は、ひとりの被験者が歯音を+VOTで生成したという報告をしているが、日本語の例外的な調音形態と見なしている。その後は、VOTの研究論文の中にも+VOT化への言及がないことから、進行しつつも気づかれていなかった現象のひとつと言える。高田(2004)は回帰式による統計予測から、1933年から約60年の間に進んだ言語変化と捉えており、また同時に+VOT化の割合が100%になるのは1988年から89年に生まれた世代であると推測している。本研究の被験者は1987年-88年に生まれた世代であり、名古屋を中心とする東海地方にもこの言語変化は進行していたものと推測される。¹⁾

次に、+VOT化によってもたらされる問題を聴覚知覚と生成の面から考察したい。以下、説明の便宜上、個々の破裂音に言及する際にはVOTが(－)の値となる有声破裂音をD、VOTが(＋)で短い無声無気破裂音をT、VOTが長い無声有気破裂音をT^hとする。

まず聴覚知覚に関わる問題であるが、Keating et al. (1983)は、言語が本来もつ内在的特性として、対立要素が存在する場合、言語はその違いを強める方向へ変化すると主張しており、それを「極性化の原理(polarization principle)」と呼んでいる。知覚の面でより距離の離れた属性であればあるほど、その違いが際立つことを意味しており、従来D:Tの対立性を解消する+VOT化はこの原理への違反となる。+VOTによって聴覚知覚にどのような影響が出るかは、高田(2004)によって東京方言話者を使って調べられており、+VOTが出現する年齢層と+VOTの聴き取りには相関性がなく、世代にかかわらず声帯振動を伴わない+VOTの/d/も適切に聞き分けがなされていると分かっている。このことは、極性の原

理に違反する言語変化にかかわらず、有声/無声の聞き分けには影響が出ていないことを意味しており、聴き取りにおけるVOTの関与性の低さを示唆する結果となっている。²⁾

有声/無声の聴覚上の境界を画定する研究も様々な言語で行われており、範疇が隣接するD:TとT:T^hにおける識別度の比較がなされている。例えば、Beach et al. (2001)やLasky et al. (1975)は、話者の母語に関係なく、T:T^hの境界で識別度が最高になることを報告している。また、Williams (1980)は、D:Tの対立をもつスペイン語母語者にあってもこの反応は変わらず、母語の識別に重要なD:Tの境界よりもT:T^h境界での識別が高かったことを指摘している。T:T^h境界での聴き取りがなぜ優位なのか、その理由はVOTが長いことによって有気(aspiration)を含めフォルマント遷移や後続母音のF₀開始周波数など多くの音響情報が知覚しやすいことに求められる。D:Tの境界域はこのように聴覚知覚にとって不利な場所であり、以上からも、VOT=0に近い領域に2つの範疇が共存することは好ましくない範疇化と言える。+VOT化は、従って、有声音と無声音の聴覚的識別がしにくい方向への変化と言える。

次に+VOT化を生成の面から考察する。+VOTはDからTへの変化であり、その調音タイミングもまったく異なるものとなる。TはこれまでDやT^hと比べて、諸言語への生起率も高いことなどから、無標として捉えられることが多かった。しかしながら、Vaux & Samuels (2005)はこれらの従来までの考え方に対して、多くの反例を使って、無標性をもつのはT^hあり、逆に最も有標性の高いのがTであるという論を展開している。清水(2006)などの研究では、T^hの生成には破裂と声門開口のタイミ

ングが最も重要で、VOTの変動の幅が少なく、TにおいてはVOTの変動幅が逆に大きくなるとされているが、Vaux & Samuels (2005)の主張はこれとは逆で、微細な調音タイミングが要求されるのはTであり、T^hの方には必要がないとしている。+VOT化との関連でこれらの主張を勘案すれば、従来のD:Tの対立が保持されている話者では、TのVOTは拡がりの自由度をもつと考えることができるが、+VOT化を起こした話者のVOTは少なくともマイナス側への変動が制限される訳で、何らかのタイミング調整が必要になってくるのではないかと予想される。Vaux & Samuels(2005)は、D:Tのシステムをもつ言語がD:T^hやT:T^hに変わった多くのケースをあげ、この事実がT^hが最も無標であり、Tが有標であることの根拠のひとつであるとしている。

以上、+VOT化を起こした話者のデータに基づき、2つの範疇がVOTの尺度上でひとつの領域で共存しているのは、聴覚的にも調音タイミングの観点からも不安定であることを中心に考察を行った。

最後に、日本語の無声破裂音に関して Riney et al. (2007) によってなされた主張について簡単に考察を行いたい。特にここで問題とするのは、日本語の無声破裂音におけるVOT値

がLisker & Abramson (1964) の区分を適用すると、無気破裂音の占める領域と有気破裂音の占める領域のちょうど中間に位置することを指摘した点である。このVOTの中位的特性(intermediacy)の指摘は、有気性の有無に関して曖昧であるという日本語無声破裂音の一面をVOTの尺度上で捉えたひとつの見解であるが、私見では中位的特性の主張は強すぎるように思われる。彼らの論拠は、日本語/p, t, k/がVOTの時間尺度でshort-lag (0–25 ms) とlong-lag (60–100 ms) の中間領域に値を取ることに置かれているが、これはLisker & Abramson (1964) で計測された7言語に基づく区分であり、その後のより多くの言語を使ったVOT研究(Shimizu, 1996; Cho & Ladefoged, 1999)に従えば、無気破裂音と有気破裂音が占める尺度上での範囲は異なっている。

表4は、Cho & Ladefoged (1999) と Shimizu (1996) による範疇化をVOTの分布幅で示したものである。Cho & Ladefoged (1999) のものは、図の中に示された分布幅を読み取ったもので、数字上の誤差が存在する可能性がある。特にこの表の中では、両方の研究において、short-lagの数値がLisker & Abramson (1964) より大きく伸びており、Shimizu (1996) は

表4 VOT値の分布幅

Source	Category	VOT Range (ms)
Cho & Ladefoged (1999)	/k/	20 ~ 40
	/k ^h /(slightly aspirated)	40 ~ 60
	/k ^h /(aspirated)	65 ~ 100
	/k ^h /(highly aspirated)	120 ~ 160
Shimizu (1996), 清水 (1998)	/p/, /t/, /k/	5 ~ 45
	/p ^h /, /t ^h /, /k ^h /	70 ~ 110

short-lagの変動幅が言語によって大きく違うことを指摘している。Cho & Ladefoged(1999)の研究は一方、long-lagの言語による変動幅の大きさを有気(aspiration)の3段階によって更に細区分しており、連続的なものであることを強調している。これによって、Lisker & Abramson(1964)の分布幅の中では存在したshort-lagとlong-lagの間の空白地帯(25–60 ms)はなくなってしまっている。Shimizu(1996)がnon-aspiratedとaspiratedの間の曖昧な(fuzzy)な境界と捉えた範囲であり、線引きによって一応の名称(slightly aspirated)は与えられている。

図2は、Cho & Ladefoged(1999)で取り上げられた無声無気破裂音のVOTデータのうち、特に日本語と同じ歯音をもつ言語を選び出して、その数値をグラフ化したものである。一番右側に本稿で論じた名古屋方言のVOTも合わせて載せた。ここで特に注目するのは軟口蓋音の数値であり、他の両唇音/歯音はほぼ25 msの範囲に入っておりunaspiratedへの分類で何ら問題はない。しかしながら、軟口蓋音/k/は別で5つの言語で40 msを越えており、unaspirated/aspiratedの境界に入っている。こ

のあたりが、日本語にaspirationがあるかどうかという問題の根源になっていると思われるが、軟口蓋音は調音の特性上、空気力学或いは舌運動の遅速性という問題によってVOTが長く伸びることは宿命となっている。グラフを見ても、概ね/k/のVOTは、/p/と/t/の平均を倍にした長さに相当しているように見える。この点に関して、Cho & Ladefoged(1999)は、彼らの18言語の分析を通して、/t/と/k/の長さ差は平均で18.9 ms、/t^h/と/k^h/の場合は16.7 msであったことを報告している。つまり調音点の違いによって、18 ms前後のVOTの変動があることを指摘している。有気性の問題は、この軟口蓋音のこの特性をも考慮に入れて論じるべきであると考ええる。

Riney et al.(2007)による日本語VOTはshort-lagとlong-lagの中間に位置するという中位特性の主張に対して簡単に考察を行った。日本語はaspirationをもたない言語であるという定説に疑問を投げかけている訳だが、現段階では、日本語はunaspiratedな言語であるというこれまでの定説を否定するまでには至っていないと考える。

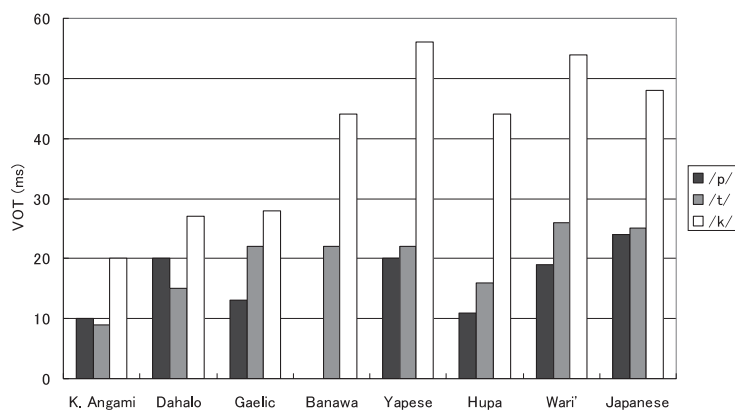


図2 8言語のVOT (Cho & Ladefoged, 1999のデータに基づく)

4. おわりに

名古屋方言の発音実相を報告し、それに基づいていくつかの点を考察した。今回のデータは非常に限られたものであり、名古屋方言の実相のある一面を描いたものにすぎない。+VOT化は現在も進行している言語変化であると考えられ、より若い世代から年齢層の高い世代まで音声データ収集の規模を広げなければならないと考えている。またVOT一般についても日本語のデータは他の言語比べて少なく、今後は方言差や男女差、また加齢によるVOTの変化なども視野に入れなければならないだろう。次稿では、現在進めている名古屋方言を話す子供たちのVOTの実相について報告する。

* 本稿は、名古屋学院大学個人研究奨励金(2005年度)の補助を受けて行った研究成果の一部である。

注

- 1) 本研究で破裂音/b, d, g/を一貫して+VOTで生成した学生の出身地は、東海地方全般に広がっており、また同時に-VOTによる発話者もこの地方にいることから、言語生育歴を含め様々な要因が絡んでいるものと思われる。
- 2) 極性の原理に適った言語変化としては、現代ベルシャ語のケースがあり、当初D:Tの対立であったものがD:Thへ変化を起こしたことが報告されている(Pisowicz, 1987)。

参考文献

- Beach, E., Burnham, D. & Kitamura, C. (2001). Bilingualism and the relationship between perception and production: Greek/English bilinguals and Thai bilabial stops. *International Journal of Bilingualism* 5, 221-235.
- Cho, T-C. & Ladefoged, P. (1999). Variation and universals in VOT: evidence from 18 languages. *Journal of Phonetics* 27, 207-229.
- Fischer-Jørgensen, E. (1954). Acoustic analysis of stop consonants. *Miscellanea Phonetica* 2, 42-59.
- Homma, Y. (1980). Voice Onset Time in Japanese stops. 『音声学会会報』 163, 7-9.
- Homma, Y. (1981). Durational relationship between Japanese stops and vowels. *Journal of Phonetics* 9, 273-281.
- Isshiki, N. & Ringel, R. (1964). Airflow during the production of selected consonants. *Journal of Speech and Hearing Research* 7, 233-244.
- Keating, P. A., Linker, W., & Huffman, M. (1983). Patterns in allophone distribution for voiced and voiceless stops. *Journal of Phonetics* 11, 277-290.
- Lasky, R. E., Syrdal-Lasky, A. & Klein, R. F. (1975). VOT discrimination by four to six and a half month old infants from Spanish environments. *Journal of Experimental Child Psychology* 20, 215-225.
- Lisker, L. & Abramson, A. S. (1964). A cross-language study of voicing in initial stops: acoustical measurements. *Word*, 20, 384-422.
- Lisker, L. & Abramson, A. S. (1970). Some effects of context on voice onset time in English Stops," in *Proceedings of the 6th International Congress of Phonetic Sciences*, 563-567.
- Peterson, G. E. & Lehiste, I. (1960). Duration of syllable nuclei in English. *Journal of the Acoustical Society of America* 32, 693-703.
- Pisowicz, A. (1987). "Development tendencies in the voiced-voiceless opposition of stop consonants (on the basis of data from Persian and Armenian)," in *Problems of Asian and African languages*, Pilaszewicz, S. & Tulisow, J. (eds.) Polish Academy of Sciences, Warsaw,

- 233-238.
- Riney, T. J., Takagi, N., Ota, K. & Uchida, Y. (2007). The intermediate degree of VOT in Japanese initial voiceless stops. *Journal of Phonetics* 35, 439-443.
- Shimizu, K. (1996). *A cross-language study of voicing contrasts of stop consonants in Asian languages*. Seibido Publishing Co., Tokyo.
- 清水克正 (1998). 「閉鎖子音の音声的特徴—有声性・無声性を中心として」『名古屋学院大学外国語教育紀要』28, 1-9.
- 清水克正 (2006). The Representation of Laryngeal Features of Obsruents. 『名古屋学院大学論集』17 : 2, 55-62.
- 城哲哉 (2006). 「日本語無声破裂音の有声化現象—予備研究」『名古屋学院大学論集』43 : 1, 43-56.
- 高田三枝子 (2004). 「日本語の語頭の有声歯茎破裂音/d/における+VOT化と世代差」『音声研究』8 : 3, 57-66.
- Tsujimura, N. (1996). *An introduction to Japanese linguistics*. Blackwell, Cambridge, MA.
- Vaux, B. & Samuels, B. (2005). Laryngeal markedness and aspiration. *Phonology* 22, 395-436.
- Williams, L. (1980). “Phonetic variation as a function of second-language learning,” in *Child Phonology*, 2 vols., Yeni-Komshian, G. H., Kavanagh, J. F. & Ferguson, C. (eds.) Academic Press, New York, 185-215.