

## 日本語有声破裂音における閉鎖調音の弱化

前川 喜久雄\*

### Weakening of Stop Articulation in Japanese Voiced Plosives

Kikuo MAEKAWA\*

**SUMMARY:** Weakening of stop articulation in Japanese voiced plosives was analyzed using the phonetically annotated part of the *Corpus of Spontaneous Japanese* (CSJ). It turned out that the weakening of /b/ and /d/ into [β] and [δ] could be best described as a function of TACA (time allotted for consonant articulation) as was the case in the affricate-fricative variation of Japanese /z/. The location of the voiced plosive phonemes in a linguistic unit showed secondary importance as the factor of the variation, but it was the location in a higher-level unit like accentual phrase or utterance that played a crucial role. The weakening of /g/ into [ɣ] or [ŋ], on the other hand, is somewhat different in that it should be treated differently depending on whether the phoneme was immediately preceded by a moraic nasal /N/. When it was preceded by an /N/, the TACA-RSA (rate of stop articulation) relationship reached a plateau much earlier (at around 70%) than in /b/, /d/, and /g/ not preceded by an /N/ (where the RSA values reaches the level of 90%). It also turned out that the curve of TACA-RSA relationship changed systematically reflecting the complexity of phonological contrast at the point of articulation of the phoneme in question. The more complex the contrast is, the earlier the curve reaches a plateau. Statistical modeling by means of logistic regression analysis revealed it was possible to predict the variation with 68–76% accuracy (closed data) using only the TACA information. The accuracy reached 72–81% when TACA and all other linguistic and extra-linguistic variables were used.

キーワード：閉鎖音，破裂音，弱化，自発音声，コーパス，CSJ

#### 1. 問題の所在

本稿の目的は日本語有声破裂音素 /b/, /d/, /g/ の音声変異の分析である。その分析では単に変異形のリストを作成して生起条件を把握するだけでなく、大量のデータを統計的に解析することによって変異の背後に存在する音声生成のメカニズムに迫りたい。このような目的をもって音声変異を研究するためには多数の話者による大量の発話を組織的に検討する必要があるが、そのような研究のために必要とされる大規模な音声コーパスが整備されはじめたのは比較的近年のことである。

筆者は先に日本語ザ行子音 /z/ における調音様式の変異の要因についてコーパスベースの分析結

果を報告した（前川 2009, Maekawa 2010, 以下では前稿と記す）。/z/ の変異は、多くの教科書がそう記載しているように、これを語中の位置に係る条件変異とみるだけでは十分に説明することができない。むしろ話者が /z/ の調音のために利用できる時間がどれほどあるかという観点からの分析が有効である。前稿では /z/ 自体の持続時間に /z/ の直前に位置する撥音、促音、ポーズの持続時間を足し合わせた測度を定義し TACA (Time Allotted for Consonant Articulation) と呼んだ。種々の解析結果は TACA が /z/ の変異における最大の要因であることを示していた。サンプルに生じている変異の約 70% は TACA によって予測され、これに位置の情報と話者の個人差の情報を追加す

\* 国立国語研究所言語資源研究系 (Department of Corpus Studies, National Institute for Japanese Language and Linguistics)

ると 80% までの予測が可能であった。

ところで、/z/ における有声摩擦音と有声摩擦音との変異は声道に明瞭な閉鎖とその急激な開放が認められるかどうかに係るものであるが、音声学的にこれとよく似た変異は有声破裂音にも認められる。例えば有声両唇破裂音 /b/ に典型的な [b] に加えて [β] で表記すべき有声摩擦音が生じるとは川上 (1977), 天沼・大坪・水谷 (1978), Vance (2008) などに報告されている。有声軟口蓋摩擦音 /g/ においては [g] と [ŋ] が生じるとは周知のとおりだが、さらに [ɣ] が生じるともよく知られている (川上 1977, Vance 2008)。最後に有声歯茎 (歯裏) 破裂音 /d/ にも閉鎖が明瞭でない有声摩擦音に類した音声 (Vance 2008 に従えば [ð]) が観察される。これらはすべて有声破裂音における閉鎖および開放調音が弱化することによって生じる現象である。以下本稿ではこれを閉鎖調音の弱化と呼ぶ。

この変異の要因については、どの文献もあまり言及していないが、川上 (1997) は /b/ について「語頭では確かに破裂音だが語頭以外では摩擦音 [β] である場合がある」(p.32) と述べて語に係る位置の関与を指摘している。一方 Vance (2008) は “This weakening of a voiced stop to a fricative normally happens only if the affected segment is both preceded and followed by a vowel or a semivowel” (p.76) と述べて音声環境の関与を示唆している。

以下本稿では 2 節で利用するデータについて説明し、3 節以下で日本語有声破裂音における閉鎖調音の弱化を前稿と同一の手法で分析する。

## 2. データ

本節では 3 節以下で分析の対象とするデータと分析において利用する変数について説明する。紙幅の制約上できるだけ簡潔な説明を試み、詳細は参考文献に譲ることとする。

### 2.1 『日本語話し言葉コーパス』

分析対象データとして『日本語話し言葉コーパ

ス』(Corpus of Spontaneous Japanese) のうち X-JToBI による音声ラベリングが施された「コア」部分 (以下 CSJ-Core と記す) を利用する。CSJ-Core の概略を表 1 に示す。今回の分析では前報と同一条件とするためにコーパスの大部分を占めるモノローグ (APS と SPS) のみを分析対象とする。

APS (Academic Presentation Speech: 学会講演) は人文、社会、理工学の 3 分野の諸学会で実況録音された研究発表の音声である。話者には大学院生が多く、理工学系では男性、人文系では女性話者が多い。SPS (Simulated Public Speaking: 模擬講演) は、人材派遣会社からの派遣によって年齢と性別をほぼバランスさせた話者群による、少人数の友好的な聴衆を前にした日常的話題についての 10~15 分程度のスピーチである。テーマとしては「私の住む街」「人生で一番嬉しかった／悲しかった出来事」「最近のニュースについてのコメント」などを指定している。読み上げ原稿の準備は禁じられており、メモを参照しながらの朗読が多い。

CSJ-Core の APS ないし SPS の話者は異なりで 123 名である。出生地は東京が 92 名、千葉県が 12 名、神奈川県が 12 名、埼玉県が 3 名であり、他にアメリカ合衆国、京都府、大阪府、北海道生まれであるが東京圏での生活歴が長い話者が各 1 名ずつ含まれている。本稿では CSJ に記録された音声は東京語を基盤とした標準的日本語音声であるとみなすことにする。

CSJ-Core には豊富なアノテーションが施されている。以下ではそのうち本稿に直接関係するものだけを簡単に説明する。アノテーション全般については国立国語研究所 (2006) 参照。

表 1 CSJ-Core における発話と話者の分布

発話タイプ	発話数*	話者数*	短単位語数	時間数†
APS	24/46	23/45	218161	14.2
SPS	54/53	38/37	225572	15.0
対話	9/9	3/3	41964	3.0
再朗読	3/3	3/3	18977	1.4

\* 斜線の左右の数字で女性話者と男性話者を示す。

† 発話間のポーズを除去した実質発話時間

## 2.2 形態論情報

CSJの転記テキストは短単位と長単位と呼ばれる二種類の形態論的単位を用いて二重に形態素解析されている。例えば「国立国語研究所」は長単位としてはひとつの複合語であるが、短単位としては「国立|国語|研究|所」のように4単位に分解される。もう少し複雑な例をあげると「彼||は||京都|駅||まで|は||歩き|つけ||た」(||は長単位境界)のようである。複合名詞「京都駅」、複合動詞「歩きつける」、複合助詞「までは」は短単位としては2単位であるが長単位としては1要素として認定されている。長単位の内部にも構造が認められるが、現在は解析の都合上最大の構造をもって長単位と認定している(短単位、長単位の詳細については小椋 2006 参照)。

日本語にこのようないわゆる膠着語的な性格があることは周知の事実であるが、先行研究では「語」という概念が何をさすのかが必ずしも明らかにされていなかった。3節では短単位、長単位にアクセント句、発話転記単位を加えて各種言語単位の影響を検討する。

さらに語種の影響も検討するが(3.4節)、CSJは語種情報を提供していないので、本研究に關係する短単位の語種を今回新たに判定した。

## 2.3 X-JToBI 情報

CSJ-CoreにはX-JToBI(Maekawa et al. 2002, 五十嵐ほか 2006)による分節音、韻律両面のアノテーションが施されている。以下の分析ではそのうちアクセント句中の位置を問題とするほか、有声破裂音直前の韻律境界の強さ(深さ)を示すBI(break indices)値との関係を分析する。X-JToBIの前身となったJ-ToBI(Venditti 1997)ではBIは基本的に「1, 2, 3」の値をとる。概略、1は語境界、2はアクセント句境界、3はイントネーション句境界以上に該当する強さを表している。発話末端(発話境界)も3に分類されることに注意。

X-JToBIのBIラベルには中間値が許容されている点に特徴がある。2+で始まる一連のラベル

は従来の2と3の中間と判断された韻律境界であり、何故そう判断されたかの理由が2+に続く文字列によって示されている。2+pは直後にポーズが生じているために典型的なアクセント句境界よりも強い(しかし典型的なイントネーション句境界よりは弱い)と判断された場合、2+bはアクセント句末に句末イントネーション(BPM)が生じていることによる中間値、2+bpはBPMとポーズがともに存在していることによる中間値である。同様に1+pはポーズが存在するために通常よりも強いと判断される語境界を表す。X-JToBIにおいて「語」は上述の短単位を指している。

以上の情報の他に、TACAの計算のために分節音ラベルを用いて当該音素および直前要素の持続時間情報を計算する。またポーズの有無、平均発話速度の情報もX-JToBIラベルから導いて分析に利用する。実際の分析にはKikuchi and Maekawa(2007)に報告したCSJのアクセント句単位XML文書を利用した。

## 2.4 破裂音弱化的判定

図1に同じ女性話者が発音した「自分」の例をふたつ示す。音声波形とサウンドスペクトログラムの下にX-JToBIの短単位ラベルと分節音ラベルが表示されている(X-JToBIでは口蓋化した/z/を「zj」で示している)。

X-JToBIの分節音ラベルでは、閉鎖区間とバーストが明瞭に観察される典型的な破裂音は2個のラベルによって表現される。閉鎖区間の終端(すなわちバースト音の直前)にラベル「<cl>」が付与され、バースト音に続く後続母音の始端に破裂音の種別を表すラベル「b, d, g」のいずれかが付与される(図1A参照)。本稿で問題とする閉鎖が弱化した有声破裂音においては、音声波形ないしサウンドスペクトログラムに明瞭な閉鎖区間が認められないか、明瞭なバースト音が観察されないか、あるいはその両方である。このような場合、X-JToBIの一般規約に従って、上述の両ラベルを融合させた「<cl>,b」「<cl>,d」「<cl>,g」が子音区間の終端(後続母音の始端)に付与され

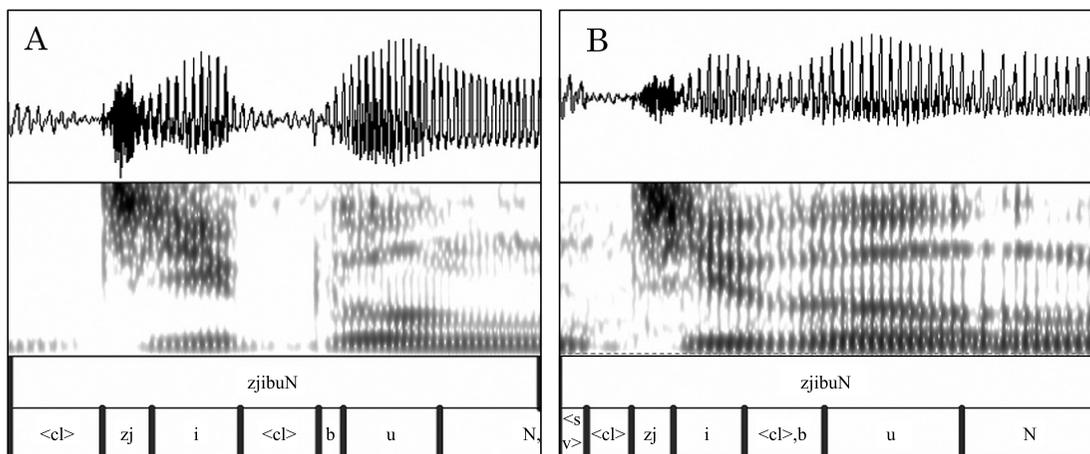


図1 音素 /b/ の変異と X-JToBI ラベル  
閉鎖区間が明瞭でバースト音も明瞭 閉鎖区間が不明瞭、バースト音無し

ている (図 1B 参照)。従って融合ラベルが用いられているかどうかによって、有声破裂音の弱化を検索できる。

ただし CSJ-Core における有声破裂音のラベリング規準にはひとつ問題がある。X-JToBI の分節音ラベリングを解説した藤本ほか (2006) は、撥音 /N/ に有声破裂音が後続する環境において「破裂音のバーストは観察されるが子音の閉鎖区間と /N/ との境界が特定できない場合」には破裂の時刻にラベル「N」を、後続母音の始端に融合ラベル「<cl>,d」等を付与したと述べている。この規準に従ったラベリングでは明瞭なバーストが観察される有声破裂音にも融合ラベルが付与されてしまうので、本研究のデータとしては不都合である。

そこで上記の音声環境におかれた有声破裂音はすべて再ラベリングすることにした。該当するサンプルは /Nb/ が 826 個、/Nd/ が 7733 個、/Ng/ が 3336 個である。ラベリングは筆者が単独で実施した。それ以外のサンプルについては CSJ-Core のラベリング結果をそのまま利用する。

最後にラベリングの精度を簡単に評価しておく。弱化といってもその程度にはかなり大きな幅があり、それがラベルのゆれになってコーパスに反映していることが予想される。そのゆれが大きすぎればデータとしての価値がなくなるので、こ

の問題を予め検証しておく必要がある。

CSJ-Core の APS と SPS には /b/ が 9279 個、/d/ が 34289 個、そして /g/ が 21953 個含まれている。このなかから各音素のサンプルを無作為に約 200 個ずつ抽出して、筆者が弱化の有無について判定をおこなった (ちなみに筆者は X-JToBI の仕様を策定したが、CSJ-Core のラベリング作業には参加していない)。

その結果を CSJ-Core における融合ラベルの有無と比較して Kappa 統計量を計算したところ、/b/ では 0.70、/d/ では 0.62、/g/ では 0.71 の値を得た。Kappa が 0.6 以上であれば、経験的にふたつのラベリング結果は「よく一致している」と判断してよいとされているので、今回利用するデータは信頼に足る精度をもっていると判断する。

なお /g/ の異音には、特に /N/ に後続する環境において [ŋ] が生じている。鼻音の典型的調音では声道に閉鎖もしくは強い狭窄とその開放が生じるが、本稿では鼻音を破裂音の一種とはみなしていないことに注意。

### 3. 分析

CSJ-Core の APS と SPS 全体において、有声破裂音が典型的な破裂音として実現された百分率

(以下これを Rate of Stop Articulation —RSA—と称する) を計算すると、/b/ は 52.9%、/d/ は 76.1%、/g/ は 29.3% である。以下の分析ではこれらの数字が RSA の高低判断のベースラインとなる。

### 3.1 位置との関係

短単位 (SUW), 長単位 (LUW), アクセント句 (AP), 発話転記単位 (IPU) の各単位において、その冒頭 (語頭) とそれ以外 (語中) における RSA の平均値を比較したのが表 2 である。IPU は CSJ の転記ファイルで採用された転記単位で前後を 200ms 以上のポーズで区切られた音声区間 (小磯ほか 2006) である。これは言語学上の単位とは言えないが、ここでは仮に発話に該当する単位として扱うことにした。表 2 から表 8 には RSA の他にサンプル数 (N) と TACA が示されているが本節では RSA にだけ注目する。TACA には

4 節で言及する。表中の R についてもそこで言及する。表 3 中の記号 ~/Ng/ については 4.2 節参照。

表 2 で RSA は、いずれの単位においても冒頭において非冒頭以外よりも顕著に高い値を示している。しかし SUW ないし LUW 冒頭における RSA は、最も低い /g/ においては 30% 強、最も高い /d/ においても 80% 弱であり、これを語中位置に係る条件異音とみる解釈は成立しがたいことがわかる。

表 2 におけるもうひとつの顕著な傾向は、冒頭位置において、SUW<LUW<AP<IPU の順に RSA が上昇していることである。以上ふたつの傾向が /z/ に観察されることは前稿に報告した (ただし前稿では IPU は検討対象としていない)。

### 3.2 先行分節音との関係

表 3 に有声破裂音の直前に位置する分節音によ

表 2 各種言語単位中の位置による RSA (閉鎖率) と TACA の変動

Phoneme	UNIT	INITIAL			NON-INITIAL		
		N	RSA [%]	TACA [ms]	N	RSA [%]	TACA [ms]
/b/	SUW	4341	66.5	64.8	4938	41.0	52.5
	LUW	3286	70.7	65.3	5993	43.1	54.4
	AP	1969	76.3	69.6	7310	46.6	55.2
	IPU	436	92.9	94.8	8843	50.9	56.5
R		0.976			0.959		
/d/	SUW	25911	78.5	62.7	8378	68.6	47.3
	LUW	19498	78.3	58.1	14791	73.1	59.9
	AP	7736	90.3	66.3	26533	71.9	56.7
	IPU	3899	98.2	74.6	30390	73.2	56.9
R		0.948			0.962		
/g/	SUW	14524	33.1	51.3	7429	22.0	52.1
	LUW	9431	33.0	50.6	9431	24.5	52.8
	AP	1957	76.7	66.6	19969	24.7	50.1
	IPU	558	93.4	103.7	21395	27.7	50.2
R		0.908			-0.606		
~/Ng/	SUW	12994	32.6	44.3	5809	20.6	38.5
	LUW	11394	32.8	44.7	7409	22.9	39.2
	AP	1897	77.0	64.8	20056	24.8	40.0
	IPU	522	95.0	92.9	18281	27.0	41.1
R		0.958			0.979		

表3 先行分節音による RSA と TACA の変動

Phoneme	/b/			/d/			/g/		
	N	RSA [%]	TACA [ms]	N	RSA [%]	TACA [ms]	N	RSA [%]	TACA [ms]
/a/	1659	47.4	52.7	4957	74.2	48.9	3067	29.9	44.9
/e/	1099	50.9	53.7	5241	67.8	46.7	1180	33.6	45.2
/i/	1636	45.5	51.7	4211	67.4	43.0	3712	21.2	38.8
/o/	1882	54.8	52.8	5897	77.0	46.8	3523	30.1	42.6
/u/	1109	43.2	48.2	2935	81.9	52.8	4090	30.7	42.2
/H/	1102	57.5	60.3	3545	72.9	46.6	3176	31.1	43.6
/N/	770	84.4	114.3	7448	86.4	100.8	3150	32.0	105.4
/Q/	21	95.2	93.2	32	100.0	93.7	16	93.7	90.3
R	0.917			0.845			0.564 (~/Ng/ 0.997)		

表4 先行短単位境界 BI による RSA と TACA の変動

BI	/b/			/d/			/g/		
	N	RSA [%]	TACA [ms]	N	RSA [%]	TACA [ms]	N	RSA [%]	TACA [ms]
1	2400	48.0	55.0	20467	71.0	57.3	13081	22.1	47.0
1+p	25	72.0	74.3	329	97.3	73.9	214	80.8	116.8
2	1840	51.3	55.3	3876	72.5	52.5	2294	33.5	53.7
2+b	292	56.6	60.9	408	77.2	52.4	449	32.5	57.3
2+bp	138	65.2	70.0	146	89.7	63.6	174	41.4	58.9
2+p	273	66.3	65.5	486	85.8	69.2	399	42.6	62.0
3	2802	54.7	60.0	6728	89.8	66.3	3363	42.2	58.4
F	1490	54.8	60.8	1749	82.8	59.2	1919	39.9	56.2
P	11	45.5	52.5	53	81.1	70.3	32	43.7	45.2
R	0.975			0.759			0.954		

る RSA の変動を示す。/b/, /d/ の場合、直前が撥音 /N/ もしくは促音 /Q/ であると RSA が顕著に上昇している。これと同一の現象が前稿でも観察されている。他方 /g/ の RSA は /Q/ の直後では顕著に上昇しているが、/N/ の直後ではベースラインからほとんど変動していない。この現象は 4.2 節で論じる。

### 3.3 先行 BI およびポーズとの関係

表 4 は、直前の短単位境界に付与された BI 値による RSA の変動を示している。生起頻度が 10 以下のラベルは集計から除外した。ポーズに関係する BI (1+p, 2+bp, 2+p) の直後では RSA が上昇する傾向が認められる。これも前稿と同じである。

ポーズの影響を直接に分析すると表 5 の結果を

得る。いずれの有声破裂音においても、その直前にポーズが存在すると RSA が 90% 以上に達する。

### 3.4 語種との関係

表 6 は、/b/, /d/, /g/ が含まれる短単位の語種による RSA の変動を示している。前稿では語種の影響を分析しなかったが、Hibiya (1988) が東京語のガ行鼻濁音の変異に語種が関与することを指摘しているので今回追加した。語種は、和語、漢語、外来語、混種語の 4 類とした。混種語には和語と漢語の混種語、外来語と漢語の混種語などがあるが、細分すると頻度が低くなりすぎるので区別していない。

表 6 は /b/, /d/, /g/ のいずれにおいても、RSA は混種語において最も高く、和語において最も低い

表5 先行ポーズの有無による RSA と TACA の変動

Phoneme	WITH PRECEDING PAUSE			NO PRECEDING PAUSE		
	N	RSA [%]	TACA [ms]	N	RSA [%]	TACA [ms]
/b/	461	92.4	97.0	8818	50.8	56.2
/d/	3979	97.8	75.3	30311	73.2	56.7
/g/	594	91.8	104.4	21359	27.6	50.1

表6 語種による RSA と TACA の変動

語種	/b/			/d/			/g/		
	N	RSA [%]	TACA [ms]	N	RSA [%]	TACA [ms]	N	RSA [%]	TACA [ms]
混種語	109	54.1	65.3	114	82.5	53.2	167	47.9	64.4
外来語	1366	51.0	60.0	1448	75.5	54.9	791	34.8	67.4
和語	3436	44.6	48.5	29141	62.4	59.0	14675	20.2	46.8
漢語	4355	50.5	65.3	3557	72.0	59.6	6297	39.4	60.3
R			0.918			-0.831			0.798

表7 発話速度による RSA と TACA の変動

Speaking rate	/b/			/d/			/g/		
	N	RSA [%]	TACA [ms]	N	RSA [%]	TACA [ms]	N	RSA [%]	TACA [ms]
=<3	59	69.5	86.1	487	89.3	84.3	163	49.1	71.9
3~6	988	61.3	70.2	3936	86.5	71.0	2343	40.5	65.5
6~9	3814	55.0	60.8	13616	79.1	60.0	9003	32.5	53.1
9~12	3440	51.4	54.6	12709	72.6	55.2	8196	25.5	47.6
12<	978	40.4	47.5	3541	63.8	50.8	2248	17.6	43.9
R			0.972			0.926			0.983

ことを示している。ただし語種のデータには語彙の偏りがあるので、解釈には慎重さが要求される。この問題は4.4節で論じる。

### 3.5 外的要因

言語外的な要因にも簡単に触れておく。RSA を APS と SPS とで比較すると、/b/において51.6% (APS) : 53.9% (SPS), /d/において76.5% : 75.8%, /g/において31.2% : 27.8%であり、一貫した傾向は認められない。

表7は5段階に区分した発話速度とRSAの関係を示している。発話速度はAPごとに計算しており単位はmora/secである。いずれの音素に関しても、発話速度の上昇につれてRSAは例外な

く単調に低下している。これも前稿と同一の結果である。

表8は話者の性別による変動を示している。いずれの音素においても女性の方が男性よりもRSAが高い。前稿でもこれと同じ結果が観察されている。最後に表9は話者の生年代(10年間隔)によるRSAの変動である。「30s」は話者が1930年代生まれであることを意味する。/b/と/g/については話者の年齢が低下するほどRSAが低下する緩やかな傾向が観察される(この傾向は前稿ではより明瞭に観察された)。しかし/g/はこの傾向に従っておらず、RSAはむしろ上昇する傾向をみせている。この問題は4.4節と6節で検討する。

表 8 話者の性別による RSA と TACA の変動

Speaker's sex	/b/			/d/			/g/		
	N	RSA [%]	TACA [ms]	N	RSA [%]	TACA [ms]	N	RSA [%]	TACA [ms]
Female	3826	57.8	60.5	14520	82.8	59.4	9475	35.3	52.9
Male	5435	49.5	56.7	19769	71.1	58.5	12478	24.8	50.5

表 9 話者の生年代による RSA と TACA の変動

Generation	/b/			/d/			/g/		
	N	RSA [%]	TACA [ms]	N	RSA [%]	TACA [ms]	N	RSA [%]	TACA [ms]
30s	350	66.6	70.2	1735	69.8	82.6	922	21.4	61.9
40s	468	53.0	57.4	1876	61.9	83.7	1163	27.4	54.0
50s	1294	61.7	60.7	5814	58.3	75.7	3286	31.9	51.4
60s	3264	51.7	56.8	11787	57.5	76.8	7883	29.3	51.7
70s	3903	49.7	57.7	13077	58.6	73.6	8699	29.5	50.1
R			0.903			0.733			-0.948

#### 4. 議 論

##### 4.1 TACA の導入

ここまで示した分析結果は、3.2 節における /N/ 直後での /g/ の RSA の低さと 3.4 節における /g/ と生年代との相関を例外として、前報における分析結果と一致していた。短単位においても長単位においても語頭位置における /b/, /d/, /g/ の RSA は高々 80% 程度であり、閉鎖の弱化を語中位置に係る条件変異とみることは困難である。

撥音と促音とポーズの直後で RSA が顕著に上昇するという事実を統一的に解釈するためには、子音の調音に利用可能な時間の増大に応じて RSA が上昇すると考えるのがよい。

子音の直前がポーズであればその時間を子音の調音ないし調音の準備に利用できるのは当然であるが、直前が撥音ないし促音の場合にも、子音調音に利用可能な時間は増大する。よく知られているように、撥音ないし促音における声道閉鎖は後続子音の調音位置において後続子音と一体化して実現されるからである。この時間を前稿では TACA と呼んだ。

TACA には幾通りもの定義が可能だが、本稿では (前稿と同じく) 当該子音の直前に撥音、促音、

ポーズがあれば、その持続時間を子音自体の持続時間に加算したものを TACA と呼ぶことにする。撥音、促音、ポーズのうちふたつ以上が子音の直前に生起している場合は、すべての持続時間を加算の対象とする。ただしポーズのなかには持続時間が 10 秒を越えるようなものもあるので、ポーズ長は最長で子音自体の持続時間を越えないという制約を設けることにする。

図 2 にはこのようにして計算された TACA (横軸、単位は ms) と RSA (縦軸、単位は %) の関

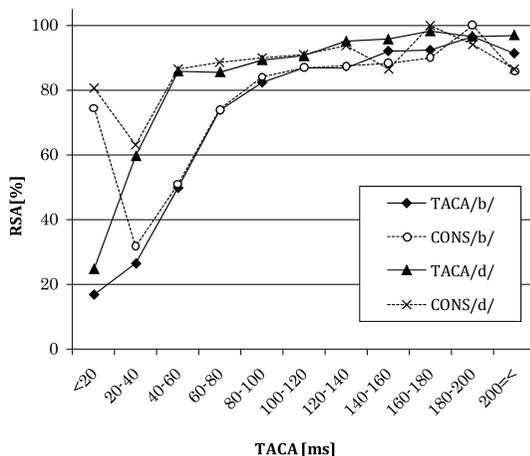


図 2 TACA および子音持続時間長と RSA

係を示した。比較のために TACA の代わりに子音持続時間単体と RSA の関係 (CONS/b/ と CONS/d/) も示してある。/b/, /d/ のいずれにあっても、TACA と RSA の間にはほぼ単調増加とみなせる関係が成立しているのに対し、子音持続時間と RSA の関係では横軸の両端 (殊に左端) において単調性が破綻している。これによって RSA を予測するための変数としては、子音持続時間よりも TACA の方が優れていることがわかる。

ここで有声破裂音の持続時間の測定法を説明しておく。子音 /b/, /d/, /g/ の持続時間とは閉鎖ないし強い狭窄の持続時間と VOT (バースト音から後続母音の開始時刻までの時間長) の合計である。図 1A においては「<cl>」ラベルと「b」ラベルの時間長を合計したものがこれにあたり、図 1B においては融合ラベル「<cl>,b」の時間長がこれにあたる。

しかし、/b/, /d/, /g/ が IPU 頭 (CSJ では多くの場合に 200ms 以上のポーズをとる) に位置しているときには閉鎖の持続時間を正確に知ることができないので、何らかの便法を用いる必要がある。CSJ-Core では典型的な有声破裂音ないし閉鎖が弱化した異音の多くにおいて、声帯振動がバースト音ないし摩擦ノイズに先行して始まることに注目して、IPU 冒頭での声帯振動の開始時点から当該子音に後続する母音の始端までの時間をもって子音の持続時間としている。本稿でもこれに従って子音持続時間を決定している。IPU 内部に 200ms 未満のポーズがあり、その直後に /b/, /d/, /g/ が生じている場合もこれと同じ方法で対処している。ただし稀に IPU 冒頭にバースト音が観察されるものの、それに先行する声帯振動が観察されない発話もあった。その場合は、閉鎖の時間長を知る手段が存在しないので、閉鎖の時間長はゼロとして処理している。

#### 4.2 /g/ の特異性

図 2 には示さなかった /g/ のデータを図 3 に示す。破線と黒丸の組合せが /g/ における子音持続時間 (CONS/g/) と RSA の関係、実線と矩形の

組合せが TACA と RSA の関係である。グラフの左端で CONS/g/ が単調性に乱れを見せているのに対し、TACA はそうっていない点は /b/, /d/ と同じである。一方、TACA の RSA が早く頭打ちになり、最終的に 70% にも達しない点は /b/, /d/ との顕著な相違点である。

その原因として考えられるのは /g/ の異音には軟口蓋鼻音 [ŋ] が含まれることである。CSJ-Core の分節音ラベルでは [ŋ] と [ɣ] を区別していないが、実際にラベリングを行うと撥音直後の環境では /g/ は多くの場合に [ŋ] として実現されていることがわかる (以下この環境を /Ng/ 環境と呼び、直前に撥音が存在しない環境を ~Ng/ 環境と呼ぶことにする)。/Ng/ 環境における TACA は相対的に大きな値をとるが、2.4 節に述べたように今回の分析では [ŋ] は破裂音に認定されないでグラフが頭打ちになるのだと考えられる。

この仮説の正しさの傍証として、図 3 には ~Ng/ 環境のグラフも示してある (実線と三角の組合せ)。そこには /b/, /d/ と同様の単調増加関係が認められることから仮説の正しさが示唆される。最後に /Ng/ 環境のグラフ (細点線と×の組合せ) は、他のいずれのグラフよりも下に位置しており RSA は最高でも 60% に達しない。ただしこの環境においても TACA と RSA の間には緩やかな単調増加関係が認められることは注目に値する。

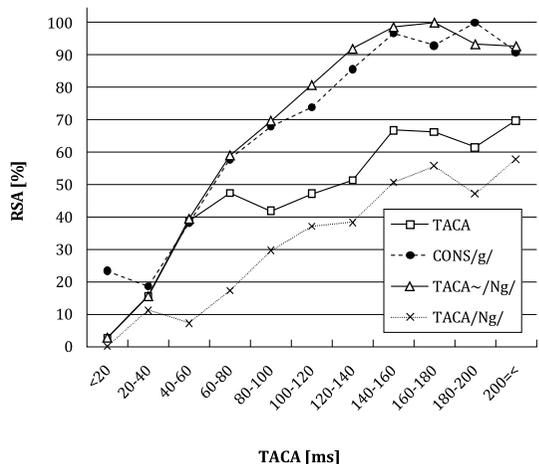


図 3 /g/ における TACA-RSA 関係

### 4.3 TACA による分析

有声破裂音の弱化要因としての TACA の有効性を示すために 3 節で検討した諸要因を TACA との関係という観点から再分析する。TACA の測定値は表 2 から表 9 に掲載済である。

位置の影響を分析した表 2 において、TACA と RSA は共変関係にあり、いずれの音素に対しても高い正の相関が認められる (表中の R がピアソン積率相関係数)。また単位冒頭位置だけでなく非冒頭位置でも高い相関が認められる。唯一の例外は /g/ における非冒頭位置であるが、この例外は非冒頭位置では冒頭位置よりも /g/ に [ŋ] が生じやすいことに起因していると考えられる。実際、撥音に後続する /g/ を除外して再計算すると高い正の相関が得られる (表最下列の ~Ng/ 参照)。

直前の分節音の影響を分析した表 3 においても /b/, /d/ に関しては TACA と RSA の間に高い相関が認められる。/g/ の場合データ全体を分析すると相関が低いが、ここでも /Ng/ 環境のサンプルを除外すると高い相関が得られる (相関係数の列中の ~Ng/ 参照)。

直前短単位境界における BI の影響を検討した表 4 では /b/, /g/ については非常に高い相関が、/d/ についても高い相関が認められる。

表 5 においては、直前にポーズが存在する場合に TACA は顕著に大きな値をとる。これは TACA の定義上、当然である。

表 6 の場合、/b/ と /g/ においては TACA と RSA の間に高い相関が認められるが、/d/ では逆に高い負の相関が生じている。この問題は次節で検討する。

表 7 において発話測度と TACA が相関するのは TACA の定義から当然予測される結果であり、RSA との相関も音素によらず高い。ただしこのことから TACA と発話速度を同一視してはならない。図 4 は AP ごとに計測した発話速度を 6mora/sec 以下、6~9mora/sec、9mora/sec 以上に区分化した上で、TACA と RSA の関係を検討した結果である。どの群においても TACA は広い範囲に分布しており、RSA との間にはほぼ単調

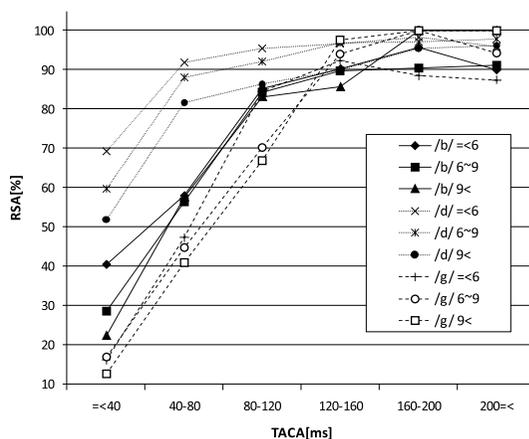


図 4 発話測度と TACA

増加の関係が成立している。TACA は直前の撥音・促音・ポーズによっても変動するため、どの発話速度群においても幅広く分布しているのである。図 4 では RSA の単調増加のあり方が音素ごとにかなり異なっていることが注目されるが、この現象については 6 節で論じる。

以上のように TACA と RSA との間には広い範囲に高い相関が認められる。言語単位中の位置、先行分節音の種別、先行韻律境界の種別など表面的には大きく異なる言語的要因が TACA を利用することによって統一的に解釈できることは、TACA が有声破裂音弱化的の本質的要因であることを強く示唆している。

### 4.4 例外の検討

しかし説明を要する例外が三つある。表 6 の /d/ における負の相関、表 8 における性別の影響、表 9 の /g/ における負の相関である。最初に表 6 (語種の影響) の問題を検討する。

ここでは /d/ の RSA と TACA に負の相関が生じていることが問題なのであるが、/d/ においては (後述する表 8 ほどではないにせよ) TACA の値には語種による顕著な差は生じていない。

また CSJ-Core の語彙には、語種の観点から分類するかぎり、相当著しい偏りが存在している。例えば /d/ は和語に 29141 回生じているが、その

45%は助動詞の「です」(7896回)と「だ」(5231回)の2語によって占められている。また /g/ は和語に14675回生じているが、その66%にあたる9726回が助詞「が」のデータである。

さらに /d/ には以下のような特殊事情もある。和語の高頻度語である「です」は準体助詞ないし否定の助動詞起源の撥音の直後に生じやすい(「僕んです」「歩くんです」「ありませんでした」等)。実際「です」の用例の47%がこの環境に生じており、TACAの増大を引き起こしているにもかかわらず、この環境におけるRSAは上昇せず、低い値にとどまっている(撥音直後の「です」のRSAとTACAは25.5%と92.1ms)。

その原因としては、「です」が高頻度の形式語であるゆえに、語の個別的特徴として弱化しやすいのではないかと考えられる。東京語のくだけた発話スタイルでは、上述の環境におかれた「です」の「で」全体が弱化して「僕んす」「歩くんす」「ありませんした」のような発音がおこなわれることはよく知られているだろう。

以上の理由により、表6におけるRSAの差をそのまま語種の影響と解釈することには慎重であるべきだと考える。/d/における負の相関も /b/、/g/における正の相関も、語彙の偏りと高頻度語の個別的特徴に影響された偶然の結果である可能性が否定できない。

次に表8(性別の影響)の問題を検討する。ここではTACAとRSAは一応期待される共変関係にあるのだが、そこに生じているTACAの差はRSAの差の原因とするには僅少すぎる点が問題となる。正しくは、TACAがRSAの差を生み出しているのではなく、所与のTACAの水準において、女性音声では男性音声よりも高いRSA値が実現される(つまり女性の方が男性よりも調音が丁寧である)と考えるべきであろう。前稿の結論もこれと同一であった。

最後に表9(生年代の影響)の問題を検討する。ここでは /b/、/d/、/g/ のすべてにおいてTACAの平均値は若い世代ほど減少する傾向をみせている。そこから予想されるのは若い世代ほどRSAが低下

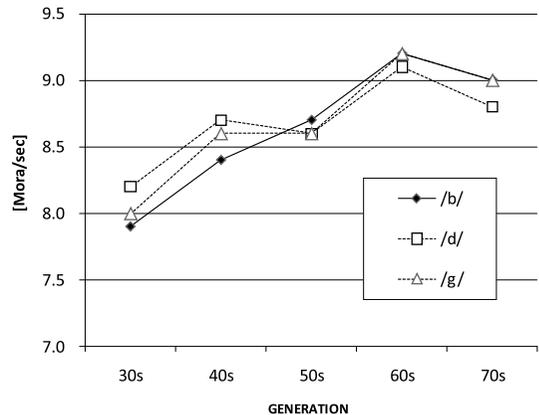


図5 話者の生年代による平均発話速度の変化

するという傾向である。実際、/b/と/d/にはその傾向が観察され、TACA-RSA間に高い正の相関が生じているのだが、/g/だけは反対に若い世代のRSAが上昇している。前稿において分析した /z/ の破擦率も若い世代ほど下降していたことを考えあわせると、一般的なのは /b/、/d/ のパターンであり、特殊なのは /g/ のパターンであると言える。

ここで図5はCSJ-Core全体を使って生年代ごとの平均発話速度を計算した結果である。この図からは若い世代ほど発話速度が速くなる傾向が読みとれる。TACAの定義上、平均発話速度の変化はそのままTACAに反映されるので、表9において若い世代ほどTACAが低下しているのはその結果と考えられる。そして本稿で想定しているTACAとRSAの因果関係からすれば、この傾向は若い世代においてRSAの低下を引き起こすと予想される。

実際 /b/、/d/ には予想どおりの変化が生じている。しかし何らかの理由で /g/ においてはRSAが低下せず、逆に上昇している。その原因は残念ながら不明であるが(ただし6節の議論参照)、/g/の生成には発話速度に帰着させることのできない世代差が存在していることは確実である。

なお、上述の /g/ の特異性は先行する撥音の影響によるものではない。~/Ng/ 環境におけるRSAとTACAの相関係数は-0.830であり、表9と同様、負の相関を示すからである。

## 5. 統計的予測

前節までの検討で TACA が有声破裂音弱化の有力な要因であることが明らかになった。本節では、統計的にはどの程度まで有声破裂音の弱化を予測できるかを検討する。予測のための統計手法としてロジスティック回帰分析を用いる。従属変数は有声破裂音における弱化の有無という 2 値変数である。独立変数としては 3 節でとりあげた各種変数を利用する。分析には R 言語 (Version 2.10.1) の glm 関数を用い (Baayen 2008, 青木 2009), 3 種の有声破裂音それぞれを独立に分析した。

表 10 に分析結果をまとめた。表の各行が独立変数に対応している。Talktype は APS か SPS かの別, spksex は話者の性別, generation は世代の別, precphoneme は直前分節音素, precsuwbi は直前短単位境界 BI, pause は直前ポーズの有無, suwinitial から ipuinitial までは, 短単位, 長単位, アクセント句, 発話転記単位の冒頭に位置しているかどうかの別である。表末の wclassl から wclasss までは語種に関する変数である。

ここで TACA 以外のすべての変数はカテゴリ変数なので, 独立変数にカテゴリ変数を含む回帰分析 (一般化線形モデル) の定石として, 各変数に含まれる各水準が 2 値のダミー変数として展開されている。例えば spksexM は話者の性別が M であることを示す 2 値変数であり, pause1 はポーズの存在を示す 2 値変数である。一般に N 個の水準をもつ変数は N-1 個のダミー変数に展開されるので, 性別やポーズのような 2 値変数では 1 個のダミー変数が, generation のように 5 水準の変数では 4 個のダミー変数が用いられている。語種は 4 水準なので, 外来語 (loan) が wclassl, 和語 (native) が wclassn, 漢語 (Sino-Japanese) が wclasss の 3 変数に展開されている。

表の列のうち, estimate はモデルによって推定された偏回帰係数である。z value は偏回帰係数を標準誤差で割って標準化した値で, その確率が  $\Pr(>/z)$  である。偏回帰係数はその独立変数の従

属変数 (閉鎖の弱化) への影響の強さを表しているが, TACA とダミー変数とでは単位が異なるので直接比較できない。そこで  $\Pr(>/z)$  が 0.05 未満の変数について, 標準化偏回帰係数 (表中の SPRC。当該偏回帰係数と当該独立変数の標準偏差の積を従属変数の標準偏差で除した値) を計算した。SPRC を用いれば各独立変数の従属変数への影響の強さを直接に比較することができる。いずれの音素についても TACA の SPRC の絶対値が最大となっているのは, 前節までの分析結果と合致し, 前稿とも一致する結果である。

表 11 にはロジスティック回帰式による予測精度 (closed data での正答率) を示した。表 10 に示した全独立変数を用いた分析にくわえて, 語種を除く言語的変数 (suwinitial, luwinitial, apinital, ipuinitial) と TACA を独立変数に用いた分析結果を示し, さらに TACA だけを独立変数とした単回帰分析の結果も示した。

言語的変数+TACA による予測精度 (表の 2 行目) は全変数よりも 1~2% 低下するだけである。TACA 単独での単回帰による予測精度 (3 行目) はさらに最大で 5% 程度低下するが, それでも 69~77% の精度が保たれており, TACA の重要性を示している。

## 6. 音韻体系との関係

本節では上に報告した分析結果について音韻体系の観点から検討を加える。日本語 (東京語) の有声破裂音と調音位置を共有する有声摩擦音と鼻音は図 6 の音素記号のように分布しており, /b/, /d/, /g/ の 3 音素はすべて異なる対立関係におかれている。対立が最も複雑なのは歯裏 (歯茎) 音であり, 破裂音, 摩擦音, 鼻音のすべてが対立している。両唇音には破裂音と鼻音の対立があるが, 軟口蓋音には破裂音しか存在しないので, 三者の調音様式間に積極的な対立は存在しない。

このような場合, 各音素の異音の実現範囲も音韻対立の複雑さを反映して変化すると考えられる。図 6 中の楕円は音韻対立が単純な場合には異

表 10 ロジスティック回帰分析の結果 \*

	/b/				/d/				/g/			
	Estimate	z value	Pr(>  z )	SPRC	Estimate	z value	Pr(>  z )	SPRC	Estimate	z value	Pr(>  z )	SPRC
(Intercept)	9.53	0.05	0.961		-0.02	-0.03	0.978		-2.38	-2.60	0.009	
talktypeS	0.21	3.92	0.000	0.21	-0.15	-4.61	0.000	-0.16	-0.09	-2.57	0.010	-0.10
spksexM	-0.27	-5.39	0.000	-0.27	-0.71	-23.36	0.000	-0.82	-0.64	-18.08	0.000	-0.70
generation40s	-0.42	-2.53	0.011	-0.18	0.17	1.77	0.077		0.63	5.04	0.000	0.31
generation50s	-0.09	-0.63	0.527		-0.29	-3.66	0.000	-0.25	0.99	9.19	0.000	0.77
generation60s	-0.42	-3.06	0.002	-0.40	-0.20	-2.64	0.008	-0.22	0.79	7.57	0.000	0.83
generation70s	-0.61	-4.46	0.000	-0.60	-0.38	-5.13	0.000	-0.44	0.95	9.18	0.000	1.02
precphonemea	-11.36	-0.06	0.954		13.43	0.03	0.974		-0.20	-0.19	0.847	
precphonemee	-11.40	-0.06	0.954		13.34	0.03	0.974		-0.33	-0.32	0.749	
precphonemeH	-10.99	-0.06	0.956		13.48	0.03	0.974		-0.35	-0.33	0.739	
precphonemei	-11.20	-0.06	0.955		13.39	0.03	0.974		-0.49	-0.47	0.640	
precphonemeN	-10.96	-0.06	0.956		12.22	0.03	0.976		-1.79	-1.71	0.087	
precphonemeo	-11.11	-0.06	0.955		13.79	0.03	0.973		-0.38	-0.36	0.717	
precphonemeQ	-8.85	-0.05	0.964		25.46	0.06	0.953		2.17	1.45	0.146	
precphonemeu	-11.26	-0.06	0.954		13.60	0.03	0.974		0.12	0.11	0.912	
precswbi1	0.11	0.09	0.926		-13.34	-0.03	0.974		-0.46	-0.50	0.619	
precswbi1+	-11.03	-0.08	0.937		-13.67	-0.03	0.973		-10.18	-0.08	0.937	
precswbi1+p	1.10	0.86	0.390		-10.82	-0.03	0.979		0.80	0.85	0.396	
precswbi2	0.30	0.26	0.796		-13.58	-0.03	0.974		0.09	0.10	0.922	
precswbi2+b	0.43	0.37	0.714		-13.39	-0.03	0.974		0.02	0.03	0.979	
precswbi2+bp	0.67	0.57	0.572		-12.75	-0.03	0.975		0.45	0.48	0.632	
precswbi2+p	0.71	0.61	0.545		-13.25	-0.03	0.974		0.41	0.44	0.658	
precswbi3	0.47	0.40	0.690		-13.15	-0.03	0.974		0.34	0.37	0.709	
precswbiF	0.51	0.44	0.661		-12.86	-0.03	0.975		0.53	0.57	0.566	
precswbiP	-0.14	-0.10	0.918		-13.34	-0.03	0.974		0.53	0.53	0.595	
precswbiPB	—	—	—		-1.24	0.00	0.998		1.42	1.13	0.257	
precswbiW	-0.82	-0.44	0.663		-13.34	-0.03	0.974		—	—	—	
pause1	-0.74	-1.30	0.195		-1.98	-4.98	0.000	-1.49	-1.39	-3.16	0.002	-0.50
suwinitial1	0.56	5.07	0.000	0.56	-0.08	-1.47	0.142		0.80	8.97	0.000	0.83
luwinitial1	0.51	5.23	0.000	0.49	0.14	2.93	0.003	0.16	0.31	4.64	0.000	0.34
apinitial1	0.20	1.76	0.079		0.62	8.32	0.000	0.61	0.83	7.91	0.000	0.52
ipuinitial1	2.18	3.64	0.000	0.93	4.13	9.93	0.000	3.07	2.61	5.54	0.000	0.90
taca	0.04	22.58	0.000	2.30	0.05	48.19	0.000	4.60	0.03	31.62	0.000	2.26
wclass1	-0.51	-2.42	0.015	-0.36	-0.27	-0.85	0.397		-0.02	-0.10	0.920	
wclassn	-0.47	-2.31	0.021	-0.45	-0.83	-2.64	0.008	-0.69	-1.02	-5.39	0.000	-1.05
wclasss	-0.62	-3.05	0.002	-0.62	-0.28	-0.89	0.375		-0.26	-1.38	0.168	

\* 表中の—は変数が生じなかったことを示す。

表 11 有声破裂音弱化の予測精度

	/b/	/d/	/g/
All variables (Table 10)	72.4 %	81.0%	78.5%
Linguistic variables and TACA	71.5%	79.2%	76.3%
TACA exclusively	68.8%	76.7%	71.3%

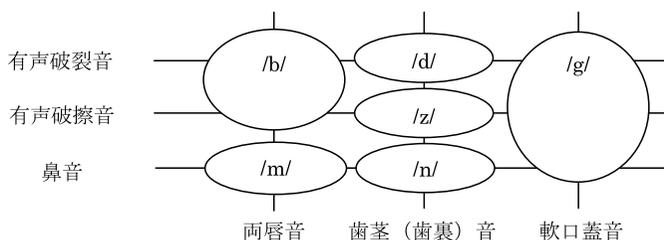


図6 日本語有声破裂音・有声破擦音・鼻音の分布

音の実現範囲が変化する様を概念的に表したものである。

このような関係が TACA と RSA の関係に実際反映されていることは大変興味深い。図7は図2, 3中の TACA 関係のグラフに前稿に報告した /z/ のデータを重ねて表示したものである。RSA が 90% 以上に達するグラフの右端を除けば、TACA の広い範囲において、所与の RSA に対応する TACA の値について /d/ < /b/ < /g/ の関係が成立していることがわかる。

例えばグラフから RSA が 80% に達する TACA の値を読みとると、/d/ では 50ms 前後、/b/ では 90ms 前後、/g/ (~Ng/) では 110ms 前後である。つまり、調音様式の対立関係が複雑な破裂音にあっては、対立関係が単純な破裂音よりも相対的に短い TACA で破裂をともなう閉鎖調音が実行可能なのである。この事実は、調音運動は音韻、言語環境のみならず発話状況などにも応じて、情報伝

達に最低限必要な対立を実現するように適応制御されているとする Lindblom (1989) の H&H 理論などと符合する。

同じく興味深いのは /z/ と /Ng/ の比較である。/z/ における声道の閉鎖（つまり破擦調音）は音韻の対立には関与していない。そのため図7における /z/ のグラフは /b/, /d/ よりも緩やかに上昇する。しかし /z/ のグラフは TACA とともにほぼ単調に増加し、最終的な破擦音の比率は 95% 以上に達する。前稿ではこの事実をもって /z/ は理想状態においては破擦音となるべく調音されていると推論した。

これに対して /g/ のうち /Ng/ 環境のサンプルはグラフの横軸の終端に達しても高々 70% の水準に達するに過ぎない。グラフの概形から判断しても TACA 値 160ms 前後で既に高原状態が生じており、グラフの右端を越えて TACA 値を増大させても RSA が 100% に達することはないと思われる。この点で /Ng/ 環境における /g/ は /b/, /d/, ~Ng/ はもとより /z/ と異なるふるまいを示している。今回のデータに記録された音声に関する限り、/Ng/ 環境における /g/ が理想状態においては有声破裂音となるべく調音されているとみなせる証拠は見つからなかった。一方 ~Ng/ 環境の /g/ は図3のデータから有声破裂音となるべく調音されているものとみなすことができる。

ひとつの可能性として指摘するならば、現代東京語の /Ng/ 環境では、鼻音の直後という自然な同化の環境において、かつて東京語においてひとつの音素として存在していたガ行鼻濁音 /ŋ/ の調音運動が化石的に保存されているのかもしれない

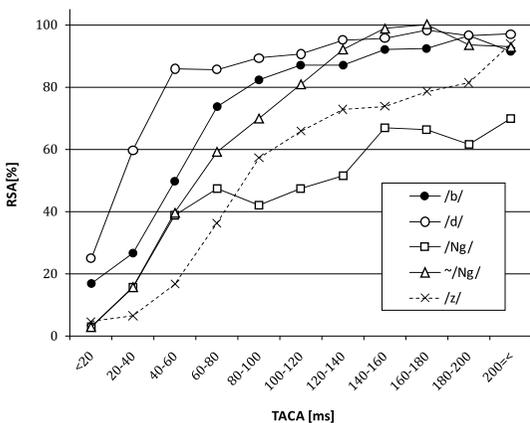


図7 /b/, /d/, /g/ および /z/ における TACA-RSA 関係

ない。そうであるならば、/g/において/b/, /d/とは反対に若い世代ほどRSAが上昇していたこと(4.4節)は、化石化した/g/の調音が最終的に消滅する方向への変化であるとみなすことができるだろう。

## 7. 結論

現代日本語(標準語)の有声破裂音/b/, /d/, /g/における閉鎖調音弱化的本質的要因はTACAである。TACAと閉鎖調音の実現率の関係は音素によって異なるが、そこには各音素の調音位置における調音様式に係る音韻対立の複雑さが反映されている。

従来、変異の要因として指摘されることのあった語ないし言語単位中の位置は副次的要因として限定的に機能しているが、その効果が顕著に生じるのはアクセント句や発話といった上位単位の冒頭であり、これを「語」の効果と解釈することには無理がある。

今回の分析結果の大部分は先に報告した/z/における調音様式の変異の分析結果と一致するが、/g/については一部で/b/, /d/, /z/と異なるふるまいが観察された。その大部分は撥音直後の/g/が[n]として実現されやすいことに起因していたが、/g/では若い世代の話者ほど破裂率が上昇する傾向については撥音との関係が認められなかった。その説明は今後の課題である。

## 謝辞

草稿に対してコメントをいただいた広島大学の五十嵐陽介氏に感謝します。また本稿における語種の影響の分析は匿名査読者の指摘に基づいて実施したものです。記して感謝します。

## 参考文献

青木繁伸(2009)『Rによる統計解析』オーム社。  
天沼寧・大坪一夫・水谷修(1978)『日本語音声学』くろしお出版。

- 五十嵐陽介・菊地英明・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』347-453, 国立国語研究所。  
小椋秀樹(2006)「形態論情報」『日本語話し言葉コーパスの構築法』134-186, 国立国語研究所。  
川上葵(1997)『日本語音声概説』桜楓社。  
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』23-132, 国立国語研究所。  
国立国語研究所(2006)『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)。  
藤本雅子・菊地英明・前川喜久雄(2006)「分節音情報」『日本語話し言葉コーパスの構築法』323-346, 国立国語研究所。  
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, 111-133。  
前川喜久雄(2009)「日本語ザ行音の調音様式の変異について」『第23回日本音声学全国大会予稿集』169-174。  
Baayen, R. H. (2008) *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge: Cambridge University Press.  
Hibiya, Junko (1988) "A Quantitative Study of Tokyo Japanese." Ph.D. diss. University of Pennsylvania.  
Kikuchi, Hideaki and Kikuo Maekawa (2007) "Construction of XML documents for the study of prosody using the Corpus of Spontaneous Japanese." *Proc. Oriental-COCOSDA 2007*, 38-42.  
Lindblom, Björn (1998) "Explaining phonetic variation: A sketch of H&H theory." In W. J. Hardcastle and A. Marchal (eds.) *Speech production and speech modelling*. Dordrecht: Kluwer Academic Publishers.  
Maekawa, Kikuo (2010) "Coarticulatory reinterpretation of allophonic variation: Corpus-based analysis of /z/ in spontaneous Japanese." *Journal of Phonetics*, 38: 3, 360-374.  
Maekawa, Kikuo, Hideaki Kikuchi, Yosuke Igarashi and Jennifer Venditti (2002) "X-JToBI: An extended J\_ToBI for spontaneous speech." *Proc. ICSLP 2002*, Denver, Colorado, 1545-1548.  
Vance, Timothy (2008) *The sounds of Japanese*. Cambridge: Cambridge University Press.  
Venditti, Jennifer (1997) "Japanese ToBI Labelling Guidelines." In K. Ainsworth-Darnell and M. D'Imperio (eds.) *Papers from the Linguistics Laboratory. Ohio State University Working Papers in Linguistics*, 50 (pp.127-162).

(Received Jul. 4, 2010, Accepted Aug. 5, 2010)