
特集「有声性の対立に関する音声と音韻」
Feature Articles: The Phonetics and Phonology of a Voicing Contrast

有声性の研究はなぜ重要なのか
Why Study Voicing

川原 繁人 (Shigeto KAWAHARA)* · 高田三枝子 (Mieko TAKADA)** ·
松浦 年男 (Toshio MATSUURA)*** · 松井 理直 (Michinao F. MATSUI)****

If the topic of voicing as a distinctive attribute of speech sounds continues to be a subject of lively interest to students of speech communication, it must be because it continues to provoke new questions or to refuse final answers to old ones (Lisker 1986, p. 3).

1. はじめに

教科書的には、[p], [t], [k] vs. [b], [d], [g] を区別するものは「有声性の対立」であり、調音音声学的には「声帯振動の有無」として教えられていることが多い。前者では子音の調音中に声帯が振動せず、後者では声帯が振動するとされる。しかし、この対立を少しでも詳細に観察してみれば、この特徴づけがあまりにも単純化されたものであることがわかる。音響的に「有声性の違い」を観察してみると、様々な言語において、以下のような特徴が観察されることが多い (Lisker 1986 他, 本特集の清水論文も参照)。

- ①有声子音には子音区間中に声帯振動が観察されるが、無声子音には観察されない。
- ②有声子音の方が無声子音よりも子音区間が短い。
- ③有声子音の方が無声子音よりも先行母音が長い。
- ④有声子音の方が周りの母音 (特に後続母音)

の基本周波数 (f_0) が低い¹⁾。

- ⑤有声子音の方が周りの母音の第一フォルマント (F1) が低い。
- ⑥無声子音の後続母音では、F1の開始が高次フォルマントの開始に遅れる (= F1 cut-back)。
- ⑦有声子音の方が無声子音よりも破裂が弱い。

このような、音韻的には一つの対立であっても、音響的には様々な次元において違いが現れることを multiplicity of acoustic cues と呼び、有声性ではこれが顕著に観察される²⁾。この「一つの音韻的な対立が、多数の音響的次元の違いによって表される」という観察は、音声モジュール・音韻モジュール、またこの二つのつながり (= interface) を考える際に多くの問題を提起する。実際に有声性の研究は音韻論・音声学を考える上で欠かせない重要な問題をいくつも提起してきた。例えば次のような具体例がすぐに挙げられる。

- (1) 多次元的な音響的差異を司る音声モジュール

*慶應義塾大学 (Keio University)

**愛知学院大学 (Aichi Gakuin University)

***北星学園大学 (Hokusei Gakuen University)

****大阪保健医療大学 (Osaka Health Science University)

は自動的 (automatic) か制御的 (controlled) か。

- (2) 音声現象と音韻現象の区切りはどこに引かれるべきか。
- (3) 「有声性」と呼ばれる対立はどのような音韻的弁別素性で表されるべきか。
- (4) 音韻パターンは、音声的要請に対してどのくらい自然か。

2. Automatic vs. Controlled Phonetics

まず、音声モジュールは「自動的 (automatic)」か、それとも「制御的 (controlled)」かという問題を考えていく。結論から言えば、現在では後者の説が有力である。しかし、歴史的に振り返ると、かならずしも後者の説が優位というわけではなかった。例えば、生成音韻論の記念碑的研究である *The Sound Pattern of English (SPE)* では、phonetic transcription という表示から実際の調音運動に変換するモジュールが仮定されており、その仕組みは普遍的かつ自動的であるとされていた：

[P]honetic transcriptions omit properties of the signal that are **supplied by universal rules**. These properties include, for example, the different articulatory gestures and various coarticulation effects—the transition between a vowel and an adjacent consonant, the adjustments in the vocal tract shape made in anticipation of subsequent motions, etc (Chomsky and Halle 1968, p. 295, emphasis added).

この考えは現在の生成統語論の主流となっている Minimalist Program にも引き継がれている。

…a condition on phonetic representation is that each symbol be interpreted in terms of articulatory and perceptual

mechanisms in a **language-invariant manner**: a representation that lacks this property is simply not considered a phonetic representation… (Chomsky 1995, p. 151, emphasis added)

Keating (1985, 1988) は、このような考え方を automatic speaking machine と特徴づけている。つまり音韻論が音声モジュールに正しい出力表示を送れば、正しい調音運動が自動的に導かれるはずだという理論である。この理論は確かに簡潔性という点では優れていた。

しかし、この人間の音声モジュールを automatic speaking machine とする考え方に対する反論が、有声性の研究に基づいてなされた。上で述べたように、先行母音の長さは有声子音に先行する場合の方が無声子音に先行する場合に比べて長い。Chen (1970) の通言語的な研究によれば、この現象は多くの言語で観察される。Chen より前の研究でこの傾向を持つとされていたのが英語、ドイツ語、スペイン語、ノルウェー語で、Chen はこれらに加えてフランス語、ロシア語、韓国語においても同様の傾向が観察されることを示した。この結果は「なぜ」有声阻害音の前で母音が長くなるのかという問題は横に置いておくとしても、「長くなる」という観察自体は普遍的な現象であり、automatic speaking machine の存在を示唆しているように思われた。実際に Chen は、有声閉鎖音の前で母音長が長くなる現象に対して、生理学的な要請 (intrinsic factors) を基にした説明を試みている。しかし、後の Keating (1979) の研究では、この先行母音の長さの違いが、言語によって程度が大きく異なり、またチェコ語やポーランド語のように先行母音の長さの違いが観察されない言語があることが指摘された (Keating 1985)。この観察は automatic speaking machine 観では説明がつかない。つまり、有声阻害音の前で母音が伸びるという傾向は普遍的ではあるが、(全く伸ばさないという選択肢を含めて) どのくらい伸ばすかは各言語の音声モジュールに委ねられていると考

えることが必要なのである。これは音声モジュールが話者のコントロール下にあることを物語っている。

このKeating (1985, 1988) の考えは phonetic knowledge (Kingston and Diehl 1994) または phonetic competence と呼べる文法領域の存在を示唆しており、後に大きな研究の呼び水となった。音声モジュールは与えられた入力を調音動作に変換するただの機械ではなく、話者が積極的に制御する文法モジュールの一つとして捉え直されたのである。この結果、抽象的・離散的な音韻カテゴリーを入力として受け取り、連続的な調音動作を導く具体的なモデルが数々提出されることとなった (Fujimura 2000, Goldstein and Browman 1986, Keating 1990, Pierrehumbert and Beckman 1988, 他)。

音声モジュールが自動的か制御的かという議論に関連して、もう少し複雑な問題を考えてみよう。有声障害音は、無声障害音に比べて周辺母音の基本周波数が低い。この現象はどのように説明されるべきであろうか。まず、生理学的な説明が可能である。第5節で詳しく述べるように、有声障害音を発音する際には、口腔内に強い狭窄や閉鎖を伴いつつ、声帯振動を維持させるために、声道空間を膨張させる必要がある。声道空間を膨張させる一つの手立ては、喉頭全体を下げることだが、喉頭が下がると基本周波数が下がることがわかっている (Ewan and Kronen 1974)。また逆に、無声障害音を発音する場合、輪状甲状筋を使い、声帯を伸ばして緊張度を上げることで無声化を促すことがある。そして声帯の緊張度が上がると、基本周波数が上昇する (Yanagihara and von Leden 1966, Löfqvist et al. 1989)。これらの生理学的な要請から無声障害音と有声障害音の基本周波数に差が出ることは間違いない。しかし、同時に、これらの生理学的な要請からだけではすべて説明がつかわけではないことも指摘されている (Kingston and Diehl 1994, pp. 439–440)。つまり、何かしらの理由で、話者が有声障害音の周りの母音の基本周波数を積極的に下げ、無声障害音の周りの母音の基本周波数を積極的に上げていると考える方が

自然だという結論になる (なぜかに関しての一つの仮説は第4節 Auditorist Theory の説明部分で述べる)。

無声子音の周りの母音でF1が高くなる観察に関しても、van Summers (1987) の興味深い観察がある。[bap] と [bab] の調音時の顎の動きを比べると、[bap] の調音時の方が [bab] 調音時の時よりも、[a] の顎の開きが母音開始地点から大きい。顎の開きはF1と正比例の関係にあるので (Stevens 1998)、結果としてF1が無声子音の前の母音で高くなる。しかし、[p] の前で [a] の調音時の顎を大きく開かなければならない生理学的な理由はない。van Summers (1987) の結果は、話者が [p] の前でF1を高めようとして顎の動きを積極的にコントロールしていることを示唆している。

さらに、近年行われている研究によると、個々の発話における f_0 と VOT にはあまり強い相関が見られないという報告が多い (Chodroff and Wilson 2018, Clayards 2018, Dmitrieva et al. 2015, Kirby and Ladd 2016, 本特集の清水論文も参照)⁵⁾。もし生理学的・機械的な要請から、 f_0 の高低が決まるのであれば、VOT と f_0 の間には強い相関が観察されるはずである。しかし現実にはそのような相関が見られないことから、 f_0 は VOT と独立した調音上のターゲットを持っていることが示唆される。つまり、有声性の対立によって、周りの f_0 が上下する現象は、生理学的な要請から自動的に導かれるものではない。

3. 音韻現象と音声現象

音声モジュールが自動的か制御的かという問題は、もう一つ重要な理論的問題と関わる。それは、ある音に関わる現象が、音韻的な現象なのか、それとも音声的な現象なのか、という線引きの問題である。日本語の高母音の無声化はこの問題を考える上で非常に有用な現象であり、実際に数々の議論がなされてきた。音韻論の入門クラスでは、日本語の高母音の無声化は音韻現象であ

り、例えばV [+high]→[-voice] / [-voice] _ [-voice] というようなルールで捉えることができるかと教えられることが多いであろう (Tsuchida 2001, Tsujimura 2007, pp. 23–27)。しかし、Jun and Beckman (1993) や Jun et al. (1998) によって、母音の無声化は、周りの子音の声帯の開きが母音と重なり合うために起こる、いわば音声的な現象である可能性が示唆された。この理論によれば、無声化した母音は音韻モジュールの出力時点で [+voice] であってもよく、音声モジュールの計算の結果、その [+voice] が周りの [-voice] に覆い隠されてしまったということになる。

…we propose that the vowel devoicing processes in Korean and Japanese are much easier to understand if we represent them, not as phonological rules, but as the result of extreme overlap and hiding of the vowel’s glottal gesture by the consonant’s glottal gesture (Jun and Beckman 1993, overhead 15)

この理論に対して、Fujimoto et al. (2002) は、無声化が起こる環境での声帯の開きを PGG (Photoglottogram) を用いて観察した。すると、無声化が起こる […CVC…] という環境で、声帯の開きのピークは一つだけ観察され、その声帯の開きの大きさは、無声子音一つ分のピークの約三倍の大きさであった。

また英語においては、例えば *kiss Ted* というようなフレーズを早い話速で発音すると、[s] と [t] の声帯の開きのピークが重なりあい、その声帯の開きの大きさは [s] と [t] の声帯の開きの大きさの合計とほぼ同じになることが観察されている (Munhall and Löfqvist 1992)。Fujimoto et al. (2002) の結果は、この英語のパターンと決定的に異なる。つまり、日本語の無声化した母音は「積極的に声帯を開く」という指令を受けていると考える方が自然なのである。よって、Fujimoto et al. (2002) の観察は、日本語の無声化母音は

[-voice] という出力を音韻レベルで持っている可能性を示唆している。

さらに、Shaw and Kawahara (2018) は、無声化した [u] の舌の調音を EMA (Electromagnetic articulography) を使って直接観察した。その結果、無声化した母音の舌の調音は「完全になくなっている」か、または「有声母音と全く変わらない」かのどちらかであり、「音声的に弱化している」調音はほとんど見つからなかった。つまり、無声化した母音が消えるか残るかという区別は離散的なものであり、この結果は無声化を司っているモジュールが音韻モジュールであるというさらなる証拠となる。

ここで重要なのは、それぞれの研究の結論よりも、「有声性」という現象に基づいて、音声モジュールと音韻モジュールの区別という理論的に非常に重要な問題が議論され続けているという点である。特に生成文法に基づいた研究では、音韻的な変化と音声的な変化をしっかりと区別せずに、ともすれば自分の理論に都合の良い音の変化パターンはすべて音韻知識に基づくと無批判に仮定する研究もあるが (Kawahara 2015)、日本語の母音の無声化は、そのような研究態度に警鐘を鳴らすものである。

4. 有声性と弁別素性

有声性の研究において、もう一つ重要な理論的問題は、「有声性の音響的な実現は音環境によって大きく異なる」という点である。例えば、有声性をもっとも顕著に定義づけるのは「声帯振動」であるが (Stevens and Blumstein 1981)、英語において語頭では、有声子音であっても声帯振動が実現しない場合が多い。この問題を解決するために提唱された概念が Voice Onset Time (VOT) である (Lisker and Abramson 1964)。VOT は、閉鎖開放から声帯振動が始まるまでの時間として定義される。語頭で声帯振動が観察されない場合も、無声子音の VOT は有声子音の VOT よりも長い。VOT は後に定義の拡張が提案された。それに従

えば、語中では、有声子音では声帯振動が子音中も持続するため、VOTは負の値をとると考えることもできる (Abramson and Whalen 2017, p. 81)。これに従って語中と語頭を一貫して捉えるならば、無声子音の方が有声子音よりもVOTの値が大きいと言える。VOTを使えば、語頭であっても母音間であっても、「無声子音はVOTが値が大きい子音である」と特徴づけられる。

このVOTという概念は提出以来、広く音声学的に有用なものと認められ、結果として英語の「有声性」とは、むしろ「帯気性 (aspiration)」として捉えられるべきものなのではないかという考えが活発に議論されるようになった。英語では、語頭および、強勢直前の語中において、[p], [t], [k] は帯気を伴って発音されるので、「有声性」という概念は不要であると考えられることすらできる。この考えでは、[b], [d], [g] が母音間において声帯振動が観察されるという事実は「受動的声帯振動 (passive voicing)」として説明される。つまり母音発音時は空気が流れ声帯振動が続くため、その母音の間に挟まれた子音は「帯気性」(= 声帯を広げるという指令) を伴っていない限り、自然に声帯が振動してしまうというのである。

ここで [p], [t], [k] vs. [b], [d], [g] の対立をどのような音韻的弁別素性で区別するべきかという問題が生じる。語頭でも語中でも [p], [t], [k] は帯気を伴うので、これらの子音は [-voice] ではなく、「帯気性 ([+ aspirated], [+ spread glottis])」という弁別素性で捉えるべきだと考える理論が1990年代から台頭してきた (Iverson and Salmons 1995, Jessen and Ringen 2002, 他)。Laryngeal Realismと呼ばれるこの理論では、音声的に帯気を伴う音は音韻的にも「帯気性」で表示すべきであり、「有声性」という弁別素性は、どのような環境においても音声的に声帯振動を伴って現れる子音に使われるべきだとする。例えば、ドイツ語やフランス語などの語頭でも声帯振動が観察されるものが具体例になる。つまり、この理論の根底には「音韻的弁別素性と音声的実現が一对一の関係で成り立つべきである」という考え方があ

る。

We take sounds specified as [spread glottis] to be implemented with an active glottal opening gesture...On the acoustic/ auditory level, [spread glottis] in stop production is implemented primarily by aspiration...We take the feature [voice] in stop production to be prototypically implemented with voicing during closure (Jessen and Ringen 2002, pp. 191–192).

この理論の対極に位置するのがSubstance Free Phonologyである (Hale and Reiss 2000, Reiss 2017)。Substance Free Phonologyでは、音韻モジュールに音声的な要素が入り込むことを一切禁止する。

Imagine a theory of phonology that makes no reference to well-formedness, repair, contrast, typology, variation, language change, markedness, ‘child phonology’, faithfulness, constraints, phonotactics, **articulatory or acoustic phonetics, or speech perception**. What remains in such a phonological theory constitutes the components of the Substance Free Phonology (Reiss 2017, p. 425, emphasis added).

つまり、[p], [t], [k] や [b], [d], [g] がどのような音声的な特徴を持つかは、[p], [t], [k] や [b], [d], [g] がどのような音韻素性を持つかということからは独立していると解釈するのである。

この理論の問題は、[p] vs. [b], [t] vs. [d], [k] vs. [g] のペアがすべてある一定の音声的な特徴で特徴づけられるという事実を説明できないことである。例えば、どのペアをとっても後者の方が短いVOTを示し、後続母音の f_0 も低い値を示す。さらに、個人内で [p] vs. [b], [t] vs. [d], [k]

vs. [g] のペアのVOTを比べると、各音の間に強い相関性が浮かび上がってくる (Chodroff and Wilson 2017)。つまり、[p] のVOTが比較的長い話者は、[k] のVOTも長く、[p] のVOTが短い話者は、[k] のVOTも短い。よって、ある個人の [p] の発音の特徴から、その人の [t] や [k] の発音の特徴をある程度推察できるわけである (Theodore et al. 2009)。ただし、同じ話者であっても、[p] と [b] のように有声性が異なる子音間にはVOTの値にそれほど強い相関が観察されなかった。これらの結果は、[p], [t], [k] の間にはある抽象的な関係性が成り立ち、その抽象的な関係によって [b], [d], [g] と区別されていることを示唆している。Substance Free Phonologyでは [p], [t], [k] や [b], [d], [g] がこのような抽象的な関係性を持っているという事実が説明できない (または偶然として扱わなければならない)。

Laryngeal Realismと Substance Free Phonologyの中間に立つ理論が Auditorist Theoryである (Kingston and Diehl 1994, 1995, Kingston et al. 2008)。Auditorist Theoryは、知覚音声学の一般的な理論であると同時に、「有声性」の対立に関しても独自の見解を示している。Auditorist TheoryではSPE的な伝統的な立場に立ち [p], [t], [k] vs. [b], [d], [g] の対立は「有声性」であるとする。Auditorist Theoryの要点は、当該音に現れる表面的な音響情報だけを見ては、「音韻的弁別素性と音声の実現が一对一の関係で成り立つ」というLaryngeal Realismの主張がそもそも成り立たないとする点にある。例えば、音節末の [p], [t], [k] vs. [b], [d], [g] の対立は、声帯振動の有無や帯気性の有無で区別ができない。むしろ、音節末の有声性は「先行する母音の長さ」および「子音の持続時間」(Port and Dalby 1982) や低いF1の値 (van Summers 1987) に顕著に現れる。さらに、音韻的な対立が環境によって変化することは上に見たとおりである⁶⁾。以上から、すべての環境を通して [p], [t], [k] vs. [b], [d], [g] を区別する普遍的な音響的特徴はないと Auditorist Theoryは主張する。

ここで以下のような問題が浮かび上がる。一方で [p], [t], [k] や [b], [d], [g] には抽象的なつながりが存在すると考える方が自然である。しかし他方で、すべての環境を通して、[p], [t], [k] と [b], [d], [g] を区別する音声的な尺度を定めることはできない。このジレンマに対する答えとして、Auditorist Theoryはこの [p], [t], [k] と [b], [d], [g] を区別する音声的な尺度は、表面的な音響情報ではなく、知覚的な情報であるとする。具体的には、有声性は隣接母音の低い基本周波数、低い第一フォルマント、子音の声帯振動などからなる「低い周波数成分」や子音の「短さ」といった知覚的な情報によって一義的に定義できるとする。これらの考え方はKingston et al. (2008) の次の一節にまとめて示されている。

Two IPPs [intermediate perceptual properties] are proposed...The first is called the 'low frequency property', because it represents the integration of acoustic properties that concentrate energy at low frequencies in and near a stop specified for [voice], namely, closure voicing and low F1 and F0 at the edges of flanking vowels. The second is called the 'C/V duration ratio', because it represents the integration of closure and preceding vowel duration — the closure voicing and F1 values at the edges of flanking vowels also contribute to this IPP. This ratio is substantially smaller in stops specified [voice] than those unspecified for this feature (Kingston et al. 2008, p. 29).

この理論によれば、弁別素性は知覚レベルで定義されるべきということになる。また、この理論は有声阻害音の前の母音の長さがなぜ長くなるのかという問題にも一つの答えを与える。有声性は「知覚的な短さ」で定義され、その短さを強調す

るためには前の母音を伸ばすのが効果的だからである (Kluender et al. 1988, Port and Dalby 1982, cf. Fowler 1990, Kingston et al. 2009)。

Auditorist Theory の根本にある思想は、弁別素性は音声的な情報に基づいている必要がある一方で、ある程度の抽象性を持っていなければならないという点にある。[p], [t], [k] と [b], [d], [g] を区別する弁別素性が音声的な情報に基づいているべきである証拠は Substance Free Phonology の問題点に関連して議論した。抽象性に関しては、Kingston et al. (2008) が次のような例を挙げている。英語の過去形を示す *-ed* は、無声音の後では [t]、有声音の後では [d] として発音される。例えば *cause* のように [z] で終わる動詞の後では [d] が選ばれる。ここで興味深いのは、音声的に考えると [b], [d], [g] の有声性と [z] の有声性はかなり異なる形で実現するという点である。[z] の有声性は、後に詳しく述べるように、声帯振動というよりも「長さ」によって表される。それにも関わらず [z] が [d] という音を選ぶ背景には何があるのだろうか。Kingston らは、そこに抽象的な [voice] という弁別素性が関わっているという見解を示している。

最後に、[p], [t], [k] と [b], [d], [g] の対立の背後にある弁別素性の候補は [voice] や [spread glottis] のみではない。無声子音を「より強い調音を伴って発音される」として [fortis] とし、有聲子音を [lenis] とすべきであるという見方もある (Kohler 1984, Trubetzkoy 1939, Zsiga 2018)。実際に、日本語の [t] と [d] の発音を EPG で比べてみると、[t] の方が舌と口蓋の接触面積が広い (Kawahara & Matsui 2017)。また、[p], [t], [k] と [b], [d], [g] の対立がどの言語でも同じ弁別素性で区別されている保証もない (Zsiga 2018)。

本節の議論から浮かび上がってくる重要なメッセージは、どの言語を考えるにあたって、有声性の対立を研究する場合、VOT や声帯振動の長さだけを見ているだけでは不十分だということである。有声性において multiplicity of cues が成り立つ限り、fo や F1、先行母音や子音の時間長な

ど複数の音響特徴を精査しなければ、大事な区別を見逃す可能性がある。近年の VOT に関する概観論文である Abramson and Whalen (2017) においても、VOT が決して喉頭特徴のすべてを捉えることができる万能ツールではないことが強調されている。Kong et al. (2012) は、日本語の音声分析において、VOT という概念は必要であるが十分ではないという見解を述べている。よって、fo や F1、時間長なども分析の対象に入れる必要がある。ここで注意が必要となるのは、fo、F1、時間長は、母音の質や話速に強く影響を受けるといえる点である (Lehiste 1970, Stevens 1998, Whalen and Levitt 1995)。今後の有声性の研究では、これらの影響を考慮に入れつつ、有声性の特徴を様々な音響的次元の観点から分析することが求められる。

5. 空気力学と音韻パターンの音声的自然性

有声阻害音の研究は音韻パターンがいかに音声学的に考えて自然であるかという問題を考える上でも大きく貢献してきた。有声閉鎖音調音時の空気の流れを考えてみよう。声帯を振動させるためには声門を通過する十分な気流が必要であり、そのためには声門下圧が声門上圧よりも十分に高くなければならない。しかし、声道内に閉鎖が生じていると声門上圧は上昇するため、次第に声帯振動を続けるのが困難になる。声帯振動を続けるためには、話者は声道空間を膨張させることによって声門上圧を下げなければならない。その方法は、喉頭を下げる、舌根を前に出す、口蓋帆を上げる、頬を膨らませるなど様々な存在する (Ohala 1983, Ohala and Riordan 1979, Westbury 1979)。

摩擦音の場合、口腔から空気が流れ出るためにこの問題が生じないと思われるかもしれないが、ことはそう簡単ではない。摩擦を作り出すためには狭め位置前後の圧力差が一定以上なければならない。従って、話者はその調音位置で摩擦を作り出せるほど気圧を高めながら、一方で声帯振動が止まらない程度に声門上圧の上昇を抑えなければ

ならない。このコントロールは非常に難しく、英語では有声摩擦音は声帯振動を伴って発音されることは希で、典型的には短い [s] として発音される。また、日本語のように、有声摩擦音が接近音のように (= 摩擦を起こさない状態で) 発音されることも多い。

このように有声阻害音は空気力学的に調音が難しい。そしてこの調音上の難しさが音韻パターンにも反映しているという主張が古くからなされている (Hayes 1999, Hayes and Steriade 2004, Ohala 1983)。この主張は「音声学的に難しい音は音韻的にも避けられる」という原則を用いて、通言語的なパターンを説明しようとする試みの一例として位置づけることができる (Archangeli and Pulleyblank 1994, Stamp 1973, Zipf 1949 など)。

第一に、音韻的に有声阻害音を有している言語はすべて無声阻害音を有している。音韻論の言葉でいうと、有声阻害音は有標 (marked) である。第二に、調音点が後ろになればなるほど、有声閉鎖音の空気力学的な問題は深刻になる (Ohala 1983)。すなわち閉鎖位置が後ろになるほど声門上の空間は小さくなり、それだけ声門上圧も上がりやすい。また頬など膨張しやすい器官をつかって気圧をコントロールすることもできなくなる。こうしたことから、音韻的にも調音点が後ろの有声閉鎖音が避けられやすいことを予測し、この予測が通言語的な一般化と合致しているという主張がある。

この空気力学的な問題は、促音の調音時に特に顕著に表れる。促音は定義上、調音時間が長い子音である。その長い調音時間中ずっと、声帯振動を可能にさせる程度に口腔内気圧を下げておくのにはかなりの負荷がかかる。実際に東京方言の有声阻害促音を観察すると、声帯振動が途中でやむ場合が多い (Kawahara 2006)。ただし、促音中の声帯振動を完全に持続させることは生理学的に不可能なわけではなく、日本の方言のなかにも完全に声帯振動が続く方言も存在するし (松浦 2016, 高田 2017)、アラビア語などでも同じような有声阻害促音が存在する。

おそらく多くの音韻論者が「有声阻害音がある意味で音韻的に避けられる傾向にある」ということには同意している。むしろ議論となっているのは、その音韻的な特徴が音声学的な特徴とどれだけ直接的に関わっているのかという点である。

一方で、話者個々人が音声学的な問題から音韻論的な有標性を導出 (induce) しているという考え方がある (Hayes 1999)。この理論では、音韻モジュールと音声モジュールは不可分な関係にある。

The source of markedness constraints as components of grammar is this [phonetic] knowledge. The effect phonetic knowledge has on the typology of the world's sound systems stems from the fact that certain basic conditions governing speech perception and production are necessarily shared by all languages, experienced by all speakers and implicitly known by all. This shared knowledge leads learners to postulate independently similar constraints. The activity of similar constraints is a source of systematic similarities among grammars and generates a structured phonological typology (Hayes and Steriade 2004, pp. 1–2).

またこの考え方によれば、音韻的な有標性が普遍的 (universal) であったとしても、その普遍性は音声学的な (もっと言えば物理的な) 要請から演繹できるため、生得的 (innate) である必要がない。

他方で、有声阻害音が有標であるように見えるのは、あくまで歴史変化の結果であるという見方もある。例えば、調音点が後ろにある有声閉鎖音はその調音上の問題から無声音として知覚されやすく、結果として無声音として再分析されてしまい、その言語で使われなくなるという説明であ

る。この立場によれば、音韻モジュールは音声モジュールの情報に従う必要はない。

Principled diachronic explanations for sound patterns replace, rather than complement, synchronic explanations, unless independent evidence demonstrates, beyond reasonable doubt, that a separate synchronic account is warranted (Blevins 2004, p. 5).

有声阻害音の調音時における空気力学的問題が、どの程度音韻システムに影響を与えるのか。その影響は通時的なものか (Hayes 1999, Hayes and Steriade 2004), それとも共時的なものか (Brown 2006)。この問題は現在に至るまで活発に議論がなされている。

有声性がこの問題においても理論的に非常に有用な題材となるのは、以下のような理由がある。「音声学的に調音が難しい音は音韻的にも避けられる」という仮説は非常に直感的で魅力的であるが、「音声学的に調音が難しい」という概念を客観的に計量化することは非常に難しい⁷⁾。つまり、この仮説は反証可能性という点で、困難を抱えている。ともすれば、「音韻的に避けられる音Aは音声学的に難しい音で、音声学的に難しいことは音韻的な証拠からもわかる」といった循環論法に陥ってしまう危険すらある。ところが、有声阻害音の空気力学的な問題に関しては、空気力学の問題に還元できるので、厳密な計量化がしやすい。例えば、声帯が振動するためには、声門上圧と下圧に1-2cm H₂Oの差が必要である (Lindqvist 1972)。つまり調音上の難しさを物理的な圧力量に置き換えて考えることができ、従って客観的に計量化して示すことができるのである。

6. 本特集の論文

以上概観してきたように、教科書で「有声性」と呼ばれる対立 (それは実は「有声性」ではない

かもしれない) は、音声理論・音韻理論に対して様々な問題を提起し続けてきた。本特集の論文もそれぞれ独自の観点からこれらの問題に新しい知見をもたらしている。

清水論文では、有声性・無声性の対立のうち特に音声的な側面について、これまでの成果を概観している。著者自身、これまでも日本語を含めた複数の言語における有声性の音声的な側面に関する実態を報告してきたが、本稿ではビルマ語を加えた新しいデータを示しつつ、改めて日本語と有声性に関わる音響的なパラメータとその具体的な表れ方について、特にVOTとfoに重点を置いていくつかの言語と対照しつつ紹介している。

Matsui and Kochetov論文および藤本論文は第5節で概説した有声阻害音調音時の声道気圧のコントロールがどのように行われているかに関して調音音声学的観点から考察を行っている。このうちMatsui and Kochetov論文では、超音波を使いロシア語では舌根を前に出す動作が観察されないことを示している。ロシア語では、palatalizationが音韻対立に大事な役割を担っており、palatalizationの方が舌根の位置に強い影響を与えると考えられている。それに対して藤本論文ではMRIを用いて熊本方言話者は舌根を前に出すことによって、咽頭空間を膨張させ、長い調音時間を通して声帯振動を持続させ得ていることを示している。

高田論文および松浦論文は、上記藤本論文の生理的な観察に対し、日本語の地域変種 (方言) を対象にして、産出された音声の音響音声学的側面を報告したものと言える。高田論文は藤本論文と同じく熊本方言の有声閉鎖促音について、促音中の声帯振動の持続に性別差や年層差が見られることを示しているが、特に若年層の話者でこれが短くなるという結果は、高年層と違って彼らが負荷の高い発音を採用しないと解釈でき興味深い。松浦論文は山形県村山方言において閉鎖区間に対する声帯振動の持続率の分布に話者間の違いが見られること、また、同一話者内でも声帯振動の持続率が幅広く分布することを示している。これらの結果は有声性をどのように実現させるかに

ついて日本語の地域変種に限っても多様なものがあることを示しており、さらなる研究が必要とされていることを示していると言えよう。

同じく Ueta 論文も有声性に関する音響音声学の側面を扱っている。同論文は、モンゴル語のハルハ方言における対立が、第4節で扱った「ある喉頭特徴の対立が有声性の対立なのか帯気性の対立なのか」という問題を、VOTの計測を基に議論している。その結果、文字では無声と有声の対立で書かれているに関わらず、音声学的には帯気性の違いであることが示唆されている。

Kawahara 他論文では、第4節で概観した Kingston らが注目する low frequency component に関して、音象徴の観点から考察している。有声障害音が低い周波数によって特徴づけられるのであれば、有声障害音は大きい振動帯を含意する。そこで、ポケモンの名づけというタスクを用いて、日本人の6歳児がこの音象徴効果を示すことを論じている。

繰り返しになるが、本特集に収録された論文だけでも有声性に関して多様な研究があることが見て取れよう。本特集によって有声性という概念のより深い理解が促されるだけでなく、これをきっかけにして有声性に関するより活発な議論が生まれることを願っている。

〔注〕

- 1) 基本周波数はF0 (エフゼロ) と表記することが多く見られたが、近年 *f*₀ (エフオー) という表記を用いることが推奨されている。榊原ら (2017) による紹介ならびに Titze et al. (2015) を参照のこと。
- 2) 調音点の違いも「破裂の強さ・スペクトルの違い」「後続の母音のF1の高さ」「周辺母音のフォルマント遷移の違い」など多くの音響的な特徴に現れるし、英語の強勢も「母音長」「強さ」「基本周波数」など様々な音響特徴に影響を与える。よって、有声性だけが特別 multiplicity of acoustic cues を体現するわけではない。ただし、有声性の対立において multiplicity of acoustic cues が顕著に観察されることも事実である。
- 3) 同様のSPEに対する批判は、Port et al. 1980も

参照。

- 4) ただし、有声障害音の前で母音が無声障害音の前より短くなる言語は報告されておらず、この点で完全に普遍性が否定されたわけではない。
- 5) VOTの定義に関しては第4節で詳しく述べる。
- 6) さらに言えば、[voice]の対立がどのような環境でどのような音響的特徴に現れるのかは、音声学の原則によって決まる。例えば、語頭において [+voice] であっても声帯振動を伴わないのは、語頭という先行する有声音がない位置で声帯振動を開始することが、母音間で声帯振動を持続させることよりも空気力学的に困難だからであるとされる (Westbury and Keating 1986)。
- 7) Kirchner (1998) において「調音上の難しさの計量化」が提唱されているが、あまりうまくいっているとは言えず、広く受け入れられてもいない。

参考文献

- 榊原健一・竹本浩典・北村達也 (2017) 「Q & A コーナー：先日の国際会議で海外の研究者が基本周波数のことを「エフ・オー」と発音していましたが、そういった言い方もあるのでしょうか」『日本音響学会誌』73(5), 331.
- 高田三枝子 (2017) 「促音閉鎖区間の有声性に関する音声詳細の地域差」『人間文化』32号, 47-61.
- 松浦年男 (2016) 「天草諸方言における有声促音の音韻論的・音声学的記述」『国立国語研究所論集』10号, 159-177.
- Abramson, Arthur S. and Douglas H Whalen (2017) “Voice Onset Time (VOT) at 50: Theoretical and practical issues in measuring voicing distinctions.” *Journal of Phonetics* 63, 75-86.
- Archangeli, Diana and Douglas Pulleyblank (1994) *Grounded phonology*. Cambridge: MIT Press.
- Blevins, Juliette (2004) *Evolutionary phonology: The emergence of sound patterns*. Cambridge: Cambridge University Press.
- Brown, Jason (2006) “An evolutionary perspective on stop inventories.” *Theoretical Linguistics* 32(2), 175-184.
- Chen, Matthew (1970) “Vowel length variation as a function of the voicing of the consonant environment.” *Phonetica* 22, 129-159.
- Chodroff, Eleanor and Colin Wilson (2017) “Structure in talker-specific phonetic realization: Covariation of stop consonant VOT in American English.” *Journal of Phonetics* 61, 30-47.
- Chodroff, Eleanor and Colin Wilson (2018) “Predictability

- of stop consonant phonetics across talkers: Between-category and within-category dependencies among cues for place and voice.” *Linguistics Vanguard*.
- Chomsky, Noam (1995) *The minimalist program*. Cambridge, MA: MIT Press.
- Chomsky, Noam and Morris Halle (1968) *The sound pattern of English*. New York: Harper and Row.
- Clayards, Meghan (2018) “Individual talker and token variability in multiple cues to stop voicing.” *Phonetica* 75(1), 1–28.
- Dmitrieva, Olga, F. Llanos, A. A. Shultz and A. L. Francis (2015) “Phonological status, not voice onset time, determines the acoustic realization of onset f0 as a secondary voicing cue in Spanish and English.” *Journal of Phonetics* 49, 77–95.
- Ewan, William G. and Robert Kronen (1974) “Measuring larynx movement using the thyroumbrometer.” *Journal of Phonetics* 2, 358–362.
- Fowler, Carol (1992) “Vowel duration and closure duration in voiced and unvoiced stops: There are no contrast effects here.” *Journal of Phonetics* 20, 143–165.
- Fujimoto, Masako, Emi Murano, Seiji Niimi and Shigeru Kiritani (2002) “Difference in glottal opening pattern between Tokyo and Osaka dialect speakers: Factors contributing to vowel devoicing.” *Folia Phoniologica et Logopaedica* 54, 133–143.
- Fujimura, Osamu (2000) “The C/D model and prosodic control of articulatory behavior.” *Phonetica* 57, 128–138.
- Goldstein, Louis and Catherine P. Browman (1986) “Representation of voicing contrasts using articulatory gestures.” *Journal of Phonetics* 14, 339–342.
- Hale, Mark and Charles Reiss (2000) ““Substance abuse” and “dysfunctionalism”: Current trends in phonology.” *Linguistic Inquiry* 31, 157–169.
- Hayes, Bruce (1999) “Phonetically-driven phonology: The role of Optimality Theory and inductive grounding.” In Michael Darnell, Edith Moravcsik, Michael Noonan, Frederick Newmeyer and Kathleen Wheatly (eds.) *Functionalism and formalism in linguistics, vol. 1: General papers*, 243–285. Amsterdam: John Benjamins.
- Hayes, Bruce and Donca Steriade (2004) “Introduction: The phonetic bases of phonological markedness.” In Bruce Hayes, Robert Kirchner and Donca Steriade (eds.) *Phonetically based phonology*, 1–33. Cambridge: Cambridge University Press.
- Iverson, Gregory and Joseph Salmons (1995) “Aspiration and laryngeal representation in Germanic.” *Phonology* 12(3), 369–396.
- Jessen, Michael and Catherine Ringen (2002) “Laryngeal features in German.” *Phonology* 19(2), 189–218.
- Jun, Sun-Ah and Mary Beckman (1993) “A gestural-overlap analysis of vowel devoicing in Japanese and Korean.” *Paper presented at the 67th annual meeting of the Linguistic Society of America*, Los Angeles.
- Jun, Sun-Ah, Mary Beckman and H.-J. Lee (1998) “Fiberscopic evidence for the influence on vowel devoicing of the glottal configurations for Korean obstruents.” *UCLA Working Papers in Phonetics* 96, 43–68.
- Kawahara, Shigeto (2006) “A faithfulness ranking projected from a perceptibility scale: The case of [+voice] in Japanese.” *Language* 82(3), 536–574.
- Kawahara, Shigeto (2015) “Can we use rendaku for phonological argumentation?” *Linguistic Vanguard* 1, 3–14.
- Kawahara, Shigeto and Michinao F. Matsui (2017) “Some aspects of Japanese consonant articulation: A preliminary EPG study.” *ICU Working Papers in Linguistics* II, 9–20.
- Keating, Patricia A (1979) *A phonetic study of a voicing contrast in Polish*. Brown University Doctoral dissertation.
- Keating, Patricia A (1985) “Universal phonetics and the organization of grammars.” In Victoria A. Fromkin (ed.) *Phonetic linguistic essays in honor of Peter Ladefoged*, 115–132. Orlando: Academic Press.
- Keating, Patricia A (1988) “The phonology-phonetics interface.” In F. J. Newmeyer (ed.) *Linguistics: The Cambridge survey, vol. 1*, 281–302. Cambridge: Cambridge University Press.
- Keating, Patricia A (1990) “The window model of coarticulation: Articulatory evidence.” In John Kingston and Mary Beckman (eds.) *Papers in laboratory phonology I: Between the grammar and physics of speech*, 451–470. Cambridge: Cambridge University Press.
- Kingston, John and Randy Diehl (1994) “Phonetic knowledge.” *Language* 70, 419–454.
- Kingston, John and Randy Diehl (1995) “Intermediate properties in the perception of distinctive feature values.” In B. Connell and A. Arvaniti (eds.) *Papers in laboratory phonology IV: Phonology and phonetic evidence*, 7–27. Cambridge: Cambridge University Press.
- Kingston, John, Shigeto Kawahara, Della Chambless, Daniel Mash and Eve Brenner-Alsop (2009) “Contextual effects on the perception of duration.” *Journal of Phonetics* 37(3), 297–320.
- Kingston, John, Aditi Lahiri and Randy L. Diehl (2008) *Voice*. Manuscript, University of Massachusetts, Amherst.

- Kirby, James and D Robert Ladd (2016) “Effects of obstruent voicing on vowel f0: Evidence from “true voicing” languages.” *Journal of the Acoustical Society of America* 140(1), 2400–2411.
- Kirchner, Robert (1998) *An effort-based approach to consonant lenition*. University of California, Los Angeles, Doctoral dissertation.
- Kluender, Keith, Randy Diehl and Beverly Wright (1988) “Vowel-length differences before voiced and voiceless consonants: An auditory explanation.” *Journal of Phonetics* 16, 153–169.
- Kohler, K. J. (1984) “Phonetic explanation in phonology: The feature fortis/lenis.” *Phonetica* 41, 150–174.
- Kong, Eun Jong, Mary E. Beckman and Jan Edwards (2012) “Voice onset time is necessary but not always sufficient to describe acquisition of voiced stops: The cases of Greek and Japanese.” *Journal of Phonetics* 40(6), 725–744.
- Lehiste, Ilse (1970) *Suprasegmentals*. Cambridge: MIT Press.
- Lindqvist, J. (1972) “Laryngeal articulation studies in Swedish subjects.” *Royal Institute of Technology Speech Transmission Laboratory, Quarterly Progress and Status Report* 10–27.
- Lisker, Leigh (1986) ““Voicing” in English: A catalog of acoustic features signaling /b/ versus /p/ in trochees.” *Language and Speech* 29, 3–11.
- Lisker, Leigh and Arthur Abramson (1964) “A cross-language study of voicing in initial stops: Acoustical measurements.” *Word* 20, 384–422.
- Löfqvist, Anders, Thomas Baer, Nancy S. McGarr and Robin Seider Story (1989) “The cricothyroid muscle in voicing control.” *Journal of the Acoustical Society of America* 85, 1314–1321.
- Munhall, Kevin. G. and A. Löfqvist (1992) “Gestural aggregation in speech: Laryngeal gestures.” *Journal of Phonetics* 20(1), 111–126.
- Ohala, John J. (1983) “The origin of sound patterns in vocal tract constraints.” In Peter MacNeilage (ed.) *The production of speech*, 189–216. New York: Springer-Verlag.
- Ohala, John J. and Carol J. Riordan (1979) “Passive vocal tract enlargement during voiced stops.” In Jared. J. Wolf and Dennis H. Klatt (eds.) *Speech communication papers*, 89–92. New York: Acoustical Society of America.
- Pierrehumbert, Janet B. and Mary Beckman (1988) *Japanese tone structure*. Cambridge: MIT Press.
- Port, Robert, Salman Al-Ani and Shosaku Maeda (1980) “Temporal compensation and universal phonetics.” *Phonetica* 37, 235–252.
- Port, Robert and Johnathan Dalby (1982) “Consonant/vowel ratio as a cue for voicing in English.” *Perception & Psychophysics* 32, 141–152.
- Reiss, Charles (2017) “Substance free phonology.” In S. J. Hannahs and Anna Bosch (eds.) *Routledge handbook of phonological theory*, Routledge.
- Shaw, Jason and Shigeto Kawahara (2018) “The lingual gesture of devoiced [u] in Japanese.” *Journal of Phonetics* 66, 100–119.
- Stampe, David (1973) *A dissertation on natural phonology*. University of Chicago, Doctoral dissertation. Published by Garland, New York, 1979.
- Stevens, Kenneth (1998) *Acoustic phonetics*. Cambridge, MA: MIT Press.
- Stevens, Kenneth and Sheila Blumstein (1981) “The search for invariant acoustic correlates of phonetic features.” In Peter Eimas and Joanne D. Miller (eds.) *Perspectives on the study of speech*, 1–38. New Jersey: Earlbaum.
- Titze, Ingo, Ronald J Baken, Kenneth Bozeman, Svante Granqvist, Nathalie Henrich, Christian Herbst, David Howard, Eric Hunter, Dean Kaelin, Raymond Kent, Jody Kreiman, Malte Kob, Anders Löfqvist, Scott McCoy, Donald Miller, Hubert No, Ronald Scherer, John Smith, Brad Story and Joe Wolf (2015) “Toward a consensus on symbolic notation of harmonics, resonances, and formants in vocalization.” *Journal of the Acoustical Society of America* 137, 3005. 10.1121/1.4919349
- Theodore, R. M., J. L. Miller and D. DeSteno (2009) “Individual talker differences in voice-onset-time: Contextual influences.” *Journal of the Acoustical Society of America* 125, 3974–3982.
- Trubetzkoy, Nikolai S. (1939) *Grundzüge der phonologie. Göttingen: Vandenhoeck and Ruprecht* [Translated by Christiane A. M. Baltaxe 1969, University of California Press].
- Tsuchida, Ayako (2001) “Japanese vowel devoicing: Cases of consecutive devoicing environments.” *Journal of East Asian Linguistics* 10(3), 225–245.
- Tsujimura, Natsuko (2007) *An introduction to Japanese linguistics, 2nd edition*. Oxford: Blackwell-Wiley.
- van Summers, W. (1987) “Effects of stress and final-consonant voicing on vowel production: Articulatory and acoustic analyses.” *Journal of the Acoustical Society of America* 82, 847–863.
- Westbury, John R. (1979) *Aspects of the temporal control of voicing in consonant clusters in English*. University of Texas, Austin, Doctoral dissertation.
- Westbury, John R. and Patricia Keating (1986) “On the natu-

- ralness of stop consonant voicing.” *Journal of Linguistics* 22, 145–166.
- Whalen, Douglas H and Andrea G. Levitt (1995) “The universality of intrinsic f0 of vowels.” *Journal of Phonetics* 23, 349–366.
- Yanagihara Naoaki and Hans von Leden (1966) “The cricothyroid muscle during phonation: Electromyographic, aerodynamic, and acoustic studies.” *Annals of Otology, Rhinology and Laryngology* 75(4), 987–1006.
- Zipf, George Kingsley (1949) *Human behavior and the principle of least effort*. Cambridge, MA: Addison-Wesley Press.
- Zsiga, Elizabeth C. (2018) “In favor of [fortis]: Evidence from Setswana and Sebirwa.” *Proceedings of Annual Meeting of Phonology 2017*.