

L2 英語摩擦音の知覚における 高周波数帯域情報の利用

川崎 貴子・田中 邦佳

要 旨

母語に無い音素を習得するには、適切な音響手がかりに基づき音韻カテゴリーを構築する必要がある。ターゲット言語の音韻カテゴリーがまだ築けていない段階のL2学習者は、どのような手がかりに基づいてL2音声の知覚を行っているのだろうか。本論文では、ターゲット言語の音韻カテゴリー形成が出来ていないL2学習者は、様々な音響手がかりに注意を払っている、とする川崎他(2011), Matthews & Kawasaki (2012)の説を検証した。

本研究では7,000 Hz以上の高周波数帯域の摩擦雑音情報が、日本語母語話者の英語の摩擦音知覚に利用されているかどうかを調査した。その結果、日本語話者は英語の母語話者が弁別に使用していない高周波数帯域の情報を弁別に使用していることが分かった。この結果は、川崎他(2011), Matthews & Kawasaki (2012)の説を支持するものである。

キーワード : L2 Phonology, phonetics, fricative

1. はじめに

第二言語音韻習得において、母語にない音の知覚は困難であることが多い。母語に存在しない音素対立を知覚する際、L2学習者は、本来弁別に利用すべき音響的特徴に基づいて判断出来ないことがある。分節音の習得では、手がかりとすべき音響特徴と音素対立を適切に関連づけ、音響知覚スペースに新たなカテゴリー境界を構築する必要があるのである。

Iverson et al (2003)では、/l/ vs. /ɹ/を音素として持つ英語を母語とする話者と、/l/ vs. /ɹ/の対立を持たない日本語を母語とするL2英語学習者が、どのように/l/ vs. /ɹ/を弁別するかを調査した。その結果、母語に対立が存在する英語母語話者の/l/ vs. /ɹ/弁別では、第3フォ

ルマント(F3)の音響スペースにカテゴリー境界が見られた。一方、母語に/l, ɹ/が存在しない日本語母語話者の知覚には、しかるべきF3帯域にカテゴリー境界が見られず、日本語母語話者は/l/ vs. /ɹ/の知覚にF3という音響情報を手がかりとして利用できていないことが分かった(図1, 図2参照)。

では、L2音知覚のためのカテゴリー境界が形成できていない段階の学習者は、どのような音響手がかりに基づいてL2知覚を行っているのだろうか。川崎他(2011), Matthews & Kawasaki (2012)では、日本語母語話者に英語の摩擦音の知覚実験を行った。そこで[s] vs. [ʃ]のように、母語の同カテゴリー(/s/)にまとめて知覚される音の弁別は困難である事を確認した。さらに弁別成績の悪かった音素のペアにおいて、L2学習者は対象音に隣接する母音からの移行という音響

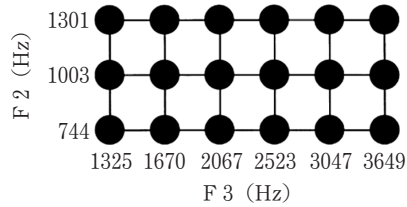


図1 Iverson & Kuhl (1996), Iverson et al (2003) にて使用された /la/~/ɭa/ 刺激。F 2, F 3 の値が異なる 18 の刺激が使用された。図は Iverson et al (2003: B 50) による。

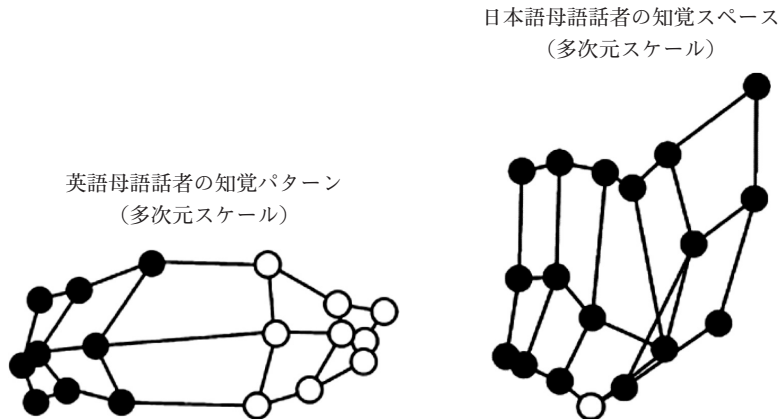


図2 英語母語話者、日本語母語話者の /l/, /ɭ/ 判断結果, およびグッドネス判断の結果を多次元スケールで表したもの。黒丸は /ɭ/, 白丸は /l/ として知覚されたことを示す。(Iverson et al 2003: B 53)

情報に注意を払い、手がかりとしていたことが分かった。このことから、母語話者是对立の知覚に必要な手がかりのみに注意を払うのに対し、L2音韻カテゴリーが構築されていないL2学習者は、注意を払うべき手がかりが分からず、様々な音響的特徴に注意を払っているのではないかと指摘した。

本論文では、川崎他 (2011), Matthews & Kawasaki (2012) により提唱された上記の説を検証するため、追実験を行った。本研究では、日本語母語話者の英語摩擦音の弁別成績が、高周波数帯域情報の有無による影響をうけるのかどうかを調査した。日本語を母語とする英語のL2学習者にとって、[s] と [θ] の区別が困難であるということはよく知られている (Weinberger 1990; Lombardi 2000; Brannen 2002; Hancin-Bhatt 1994, 他)。摩擦音の音響特性を比較すると、[s] では 4,000 Hz より高い部分に強エネルギー帯域

(high intensity) があるのに対し、[θ] は摩擦雑音のエネルギーが弱く、[s] のような高エネルギー帯が見られない。日本語を母語とする英語L2学習者はこのような摩擦エネルギーの違いを知覚の手がかりとして利用することができず、[s] vs. [θ] の弁別が出来ないのだと考えられる。もし日本語話者がこのような摩擦雑音の音響的な差を知覚の手がかりとして利用していないのであれば、高周波数帯域の音響情報が存在する音声での弁別と、音声をローパスフィルターに通すことにより高周波数帯域の情報を除いた音声での弁別の正答率に差は無いはずである。一方、川崎他 (2011) Matthews & Kawasaki (2012) にて提唱されたように、もし音韻カテゴリー生成が出来ていない学習者が様々な音響情報を、L2知覚に利用しているのであれば、高周波数帯域の情報が削除された場合とそうでない場合では、弁別成績の差が見られる可能性がある。

2. Tabain (1998)

Tabain (1998) は、英語母語話者による英語摩擦音の知覚を調査した。英語の非歯擦 (non-sibilants) [f] vs. [θ] の摩擦部分の音響特徴を比較すると、10,000 Hz 以下の帯域ではほとんど差が見られない。そこで Tabain は、英語母語話者は 10,000 Hz 以上の帯域の摩擦音の特徴を [f] と [θ] の弁別に利用しているのかどうかを確認するため、9,166 Hz (Bark スケールで 1~22) までの周波数帯域情報のみの音声と、15,415 Hz (Bark 24) までの周波数帯域の情報を持つ摩擦音で同定実験を行い、高周波数帯域情報の有無により、摩擦音の同定結果に差が出るかどうか調べた。その結果、非実在語においては 10,000 Hz 以上の情報の有無が [f], [θ] の知覚に影響を与えたものの、実在語では高周波数帯域情報の影響は見られなかった。

Tabain (1988) の実験における無意味語と実在語での結果の差は、無意味語・実在語の音声処理プロセスの間に違いがあった可能性を示唆している。実在語の処理では、音の同定に必要な音響情報に集中して注意を払い、処理が行われるが、無意味語の音声知覚処理においては、実在語の処理の場合に比べ、同定に必要な音響情報のみならず、その他の音響情報に注意を払うのではないか。もしそうであれば、Tabain に見られた母語話者の無意味語の知覚プロセスは、川崎他 (2011)、Matthews & Kawasaki (2012) に見られた L2 学習者の知覚プロセスに類するものであると言える。

本研究では、Tabain (1998) が実験で比較したように、異なる周波数帯域情報を持つ刺激 3 パターンを使用し、摩擦音の知覚弁別実験を行った。これら三種類の刺激パターンの比較により、高周波数帯域情報が、L2 英語学習者に手がかりとして利用されているのかどうかを調べた。もし日本語話者が摩擦音の知覚に高周波数帯域の情報を利用しているのであれば、刺激のパターン間で弁別成績に差がでることが考えられる。そしてそのよう

な結果は、音韻カテゴリー形成前の L2 学習者が様々な音響情報に注意を払っているとすると、川崎他 (2011)、Matthews & Kawasaki (2012) の説を支持するといえる。

3. 方 法

実験の刺激に使用するため、16 語の英語の無意味語 (疑似語) を作成した。用いた単語は全て /CVC/ 構造を持つ 1 音節語であった。単語内の各セグメントを C_1VC_2 と表記すると、 C_1 は、[f], [v], [θ], [ð], [s], [z], [t], [d] の 8 つの子音のいずれか、V は母音 [a] または [o], C_2 は全ての単語に共通の [f] であった。

作成した単語 16 語を 1 人のアメリカ英語母語話者の男性が読み上げ、Rode RD-NT 1-A マイクロフォンに M-Audio Mobile Pre USB プレアンプを接続した MacBook Pro を使用して音声の収録を行った。収録音声のサンプリング周波数は、44.1 kHz、量子化ビットレートは、16 bit であった。音声の処理は、Praat 5.3.04 (Boersma & Weenink, 2012) を使用して行った。

本実験に先立ち、収録した音声データ *saf*, *θaf*, *faf* の先頭の [s], [θ], [f] の安定部 10 ミリ秒を抽出した。図 3 は、その抽出部分のスペクトログラムである。スペクトログラムの y 軸は周波数 (Hz) で、内部の表示が濃くなればなるほどその周波数帯がより共鳴し、エネルギーが強いことを示す。歯擦音の [s] と非歯擦音の [θ], [f] を比較すると、[s] の方が摩擦雑音のエネルギーが強く、特に 4,000 Hz から 9,000 Hz の帯域で強くなっていることが分かる。一方、非歯擦音の 2 音を比較すると、7,000 Hz 以下の帯域ではほとんど差が見られないが、[f] では 7,000 Hz から 10,000 Hz の間の帯域でエネルギーがやや強くなっている。

上に述べた様に、本実験で使用した音声では、[s] と非歯擦音の差は 7,000 Hz 以下の部分で十分に顕著である。一方、[θ] と [f] の音響的違いは 7,000 Hz から 10,000 Hz の間の摩擦雑音強度

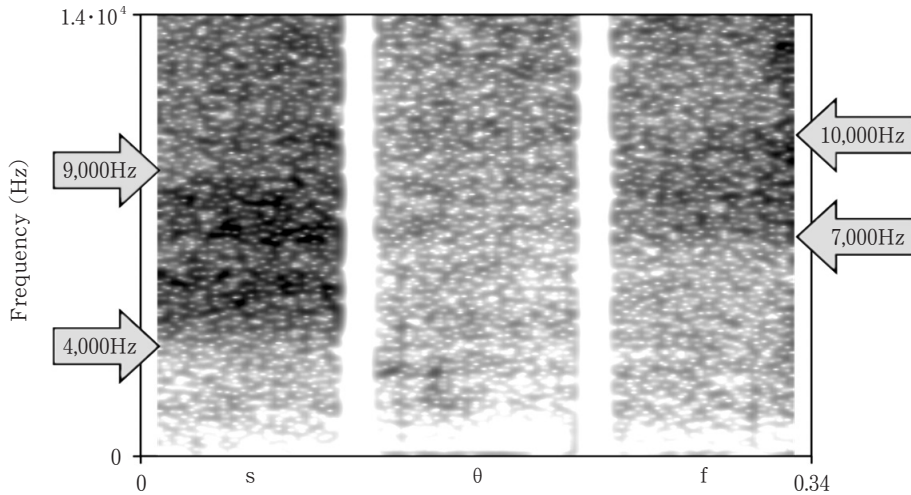


図3 英語摩擦音のスペクトログラム比較

に見られることがわかった。したがって、異なる周波数帯域での摩擦雑音情報利用の有無を調査するため、本実験では、0～7,000 Hz、0～10,000 Hz、0～22,050 Hzの3種類の周波数帯域情報を持つ刺激を作成した。

収録された音声を各単語毎の音声ファイルに分割した。その後、周波数帯域情報が異なる3つのパターンの音声を作成するため、各音声を10,000 Hz、7,000 Hzをカットオフ周波数とするローパスフィルターに通す作業を行った。また、全ての音声の音量の正規化を行った。

本実験は、AX法を使用して行った。録音した16語の C_1VC_2 の単語からs-θ、f-θ、t-θ、s-tの4つのタイプの組み合わせを作成した。各タイプ内の単語対は、A-X間で VC_2 が共通し C_1 のみが異なるように作られた。例えば、s-θタイプの試行には、saf-θaf、sof-θof、zaf-θaf、zof-θofの4つの単語対が含まれた。順序効果を考慮し、AとXの提示順を交換した単語対も作成した。用意された単語対の合計数は、4(タイプ)×4(単語対)×2(提示順)の32対であった。これらの単語対を異なる3種類のカットオフ周波数で提示した。したがって、試行の合計数は、32(単語対)×3(カットオフ周波数)の96試行であった。実験試行に36試行のフィラー試行を加え、合計132試行からなる試行リストが作成された。

実験の作成および実施には、SuperLab 4.08を使用した。A-Xの刺激間隔(ISI)は250ミリ秒であった。実験参加者毎に試行の提示順がランダム化されるように設計した。

実験の参加者は、都内の私立大学および大学院に在籍する日本語母語話者12名であった。

各参加者は、MacBook Airに接続されたUltra-sone HFI-580ヘッドフォンから聴こえてくる2つの音声は「同じ」か「違う」かを判断し、コンピューターのキーボードのキーを押して回答した。各参加者は、3試行の練習試行を行った後、132試行の実験試行を行った。実験の所要時間は約10分であった。

4. 結 果

実験で得られた弁別課題の正答率を4種類の試行のタイプ(s-θ、f-θ、t-θ、s-t)、3種類のローパスフィルターのカットオフ周波数(処理無し、10,000 Hz、7,000 Hz)別に分類して集計し、分析を行った。図4に、試行のタイプ、カットオフ周波数別の正答率の平均値を示す。

図4から、いずれの試行タイプにおいてもカットオフ周波数が低くなるほど正答率が低くなる傾向が見られる。一方、s-tタイプのように、正答率の幅が93.75%～98.96%と、その差が僅かなタ

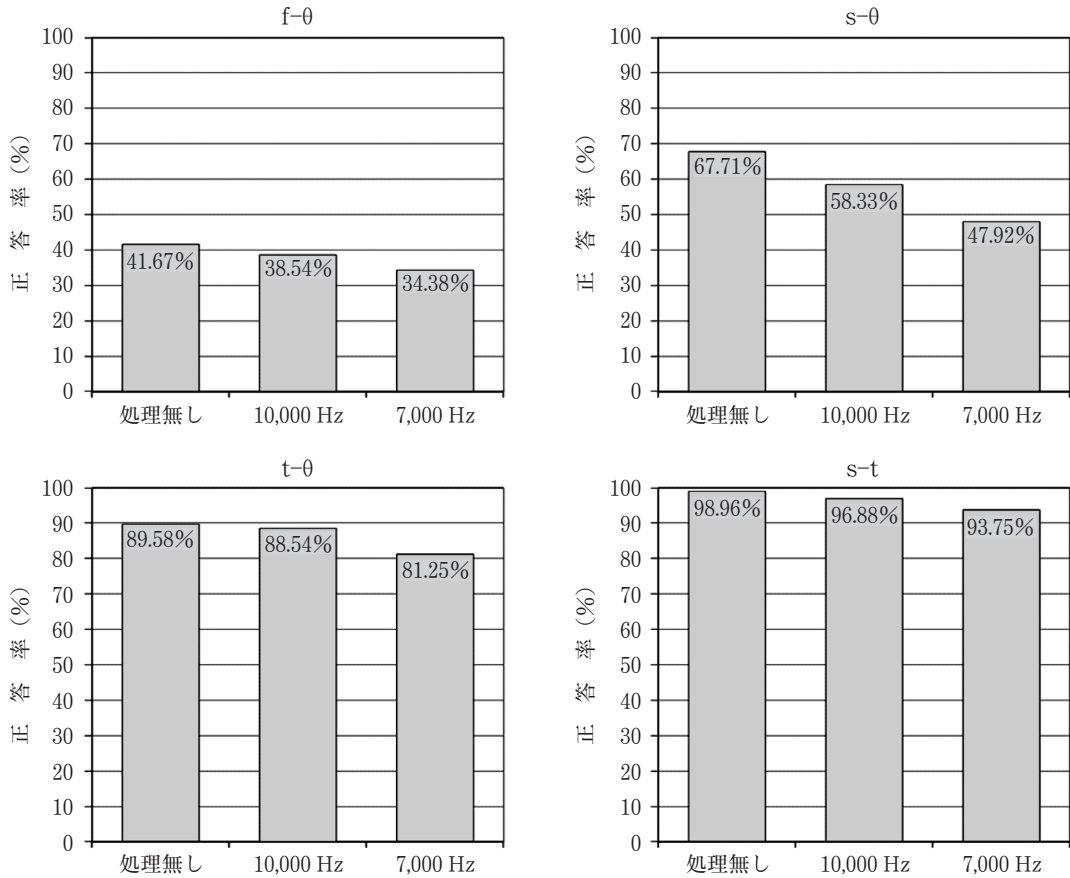


図4 試行のタイプ・カットオフ周波数別正答率の平均値

表1 カットオフ周波数別正答率の平均値

試行のタイプ	処理無し		10,000 Hz		7,000 Hz		Sig.
	Mean	S.D.	Mean	S.D.	Mean	S.D.	
f-θ	41.67%	17.94	38.54%	25.82	34.38%	23.31	n.s.
s-θ	67.71%	21.62	58.33%	26.83	47.92%	27.09	$p < .01$
t-θ	89.58%	12.87	88.54%	11.25	81.25%	14.60	n.s.
s-t	98.96%	3.61	96.88%	7.77	93.75%	11.31	n.s.

タイプがあることもわかる。これらの正答率の統計的な差異を検定するため、試行のタイプ別にローパスフィルターのカットオフ周波数を要因とした一要因の分散分析を行った。

正答率の平均値、標準偏差 (S.D.), 統計結果をまとめたものが表1である。分散分析の結果、s-θタイプではカットオフ周波数間に有意な差は示さ

れず [F(2, 22)=1.13, n.s.], t-θタイプ [F(2, 22)=1.92, n.s.], s-tタイプ [F(2, 22)=2.35, n.s.] においても有意な差はみられなかった。f-θタイプでは、カットオフ周波数要因の主効果が有意であったため [F(2, 22)=10.17, $p < .01$], Bonferroni法を用いて多重比較を行った。その結果、処理無し試行とカットオフ周波数 7,000 Hz 試行の正答率

の間に有意な差がみられた ($p < .05$)。

5. まとめ

本実験では、s-θペアの弁別のみにはあるが、カットオフ周波数による有意差が見られた。この結果は、日本語母語話者が高周波数領域の摩擦雑音の特徴に注意を払っていることを示している。しかし、むしろ7,000 Hzより高い帯域の情報が決定的であると思われたf-θの弁別グループでは、カットオフ周波数による有意差は見られなかった。これは、f-θの両方の音が日本語に存在しないことに起因するかもしれない。どちらの子音も音響特徴の情報との結びつきが弱く、どの音響情報を弁別に利用すべきなのか、試行錯誤している段階なのかもしれない。また、[f, θ]の非歯擦音の摩擦雑音は、歯擦音の摩擦雑音に比べ、強度が弱いという特徴がある。そのため、7,000 Hz以上の帯域にある音響的な差も、[s]の摩擦雑音と比べ知覚されづらいと言える。このように、f-θの知覚ではカットオフ周波数間で弁別成績の差は表れなかったが、同じく弁別成績が低いs-θの対立では有意差が見られた。これは、日本語母語話者が高周波数帯域の摩擦雑音を手がかりとして利用していることを示唆し、川崎他(2011)、Matthews & Kawasaki(2012)によるL2学習者の音響情報利用を支持するものである。

*本研究の実施にあたっては日本学術振興会科学研究費補助金(基盤研究C)(「L2音韻習得における二重モデルの構築」課題番号:23520709)の助成を受けた。

References

- Boersma, P. and Weenink, D. (2012). Praat: doing phonetics by computer [Computer program]. Version 5.3.04, retrieved 12 January 2012 from <http://www.praat.org/>
- Hancin-Bhatt, B. (1994). Segment transfer: a consequence of a dynamic system. *Second Language Research* 10(3): 241-269.
- Iverson, P., Kuhl, P., Akahane-Yamada, R., Diesch, E., Tohkura, Y., Kettermann, A. and Siebert, C. (2003). A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition*, 87, B 47-B 57.
- 川崎貴子, マシューズジョン, 田中邦佳, 大館善文 (2011). 「音韻知覚における音響の手がかりの影響: 日本語母語話者による歯間摩擦音知覚」口頭発表, 日本第二言語習得学会夏季セミナー, 関西セミナーハウス.
- Lombardi, L. (2000). Second language data and constraints on manner: explaining substitutions for the English interdentals, ROA_418_09100.
- Matthews, J. and T. Kawasaki (2012). Mapping acoustic cues in L2 speech perception and the acquisition of new segmental categories. Presented at the 12th annual conference of the Japan Second Language Association. Hosei University, Tokyo Japan.
- Tabain, M. (1998). Non-Sibilant Fricatives in English: Spectral Information above 10 kHz. *Phonetica* 55, 107-130.
- Weinberger, S. H. (1990). Minimal segments in second language phonology. In Leather, Jonathan and Alan James eds., *New sounds 90: Proceedings of the 1990 Amsterdam symposium on the acquisition of second language speech*. University of Amsterdam.
- Werker, J. F., J. S. Logan (1985). Cross-language evidence for three factors in speech perception. *Perception and Psychophysics* 37(1), 35-44.

Use of High-frequency Acoustic Cues in the Perception of English Fricatives by Japanese Native Speakers

KAWASAKI Takako and TANAKA Kuniyoshi

Abstract

In second language (L2) acquisition of phonology, learners have to create new phonological categories based on the appropriate acoustic cues. If learners have not yet established the categories required for target language perception of L2 sounds, how can they perceive the target sounds? What kind of cues do they rely on in perceiving new sounds? The present paper examined the hypothesis by Kawasaki *et al.* (2011), and Matthews & Kawasaki (2012) that L2 learners who have not established the L2 target phonological categories pay attention to various acoustic cues.

In this study, we investigated whether or not the acoustic information of frication noise above 7,000 Hz is used by Japanese learners of English. Our results show that Japanese native speakers use information from this high-frequency range that native English speakers do not.