

L2 習得による無声摩擦音知覚マップの再構築*

Reconstruction of the Perceptual Map for Fricatives in L2 Acquisition

川崎 貴子¹, 田中 邦佳¹, 竹内 雅樹², マシューズ・ジョン³
Takako Kawasaki, Kuniyoshi Tanaka, Masaki Takeuchi, John Matthews

¹法政大学, ²東京大学大学院, ³中央大学

¹Hosei University, ²The University of Tokyo University, ³Chuo University
kawasaki@hosei.ac.jp

概要

L2 音を新たに習得することは, L1 の音素マップに新たな音を追加することを意味する. 新たな音を追加することは, 知覚マップ上の L1 音のスペースにどのような影響を与えるのだろうか. 本論文では, L2 音声習得が進むことで, 無声摩擦音の音声知覚マップがどのように変化するか, そして既にある L1 音間の距離がどのように変化するかを, ノイズ下での音声混同実験により調査した. 混同傾向の MDS による分析, および選択肢の選択率の比較により, 習得が進むことにより, [i] が後続する環境で, *s, sh* が区別される, *t, s* の距離が近くなるなど, 知覚マップに変化が見られた.

キーワード: 第二言語習得, 音声知覚, 音韻論

1. はじめに

母語 (L1) に加えて, 第二言語 (L2) を習得するに際して L1 の音声文法が L2 に干渉することについては, 古くから多くの研究がなされている (Best, 1995; Flege, 1995; Brown, 1988 他). しかしその一方, L2 を習得することが, 既にある L1 にどのように影響するのかについての研究は比較的新しく (Cook, 2003), ことに音声・音韻面での研究はまだ限られている. 本研究では, L2 習得で新たな音素を習得することにより, 学習者の音素マップにすでに存在する L1 音に影響が見られるのか, L2 音と L1 音との知覚的距離がどう変化するかを探る.

L2 習得が L1 の発話に影響を与えることは, 近年の L2 音声研究によって示されている. たとえば, Flege (1987), Chang (2012) 他は, L2 のイマージョン環境にて学習した場合, L1 発話に L2 の音響的な影響が見られたと報告している. しかし, ほとんどの研究が, 発話に見られる音響的な影響を調査したものであり, 音声知覚における影響に関する研究は少ない (たとえば Tice & Woodley, 2012). また, これまでの研究は, 母音のフォルマント特性 (Flege, 1987; Guion, 2003; Chang, 2012 他), および有声・無声破裂音の VOT 値 (Flege, 1987; Chang, 2012; Harada, 2003) に見られる変化を調査

したものに限られていた. 本研究では, 日本語を母語とする英語 L2 学習者の, 無声摩擦音の知覚を調査対象とした. L2 である英語を学び, 新たな L2 音が追加されることで L1 として存在する音素同士, および L1 音と L2 音の間の知覚的距離がどのように変化するかを調査した.

本研究では, L2 学習者の知覚スペースにおける音素間の距離を計測するため, Miller & Nicely (1955) を元にした知覚混同実験を行った. 参加者には, 英語発話にノイズを合成した音声を呈示し, ターゲットとなる語の摩擦音が, どの音であったかを選択させるものであった.

実験では, 日本語を母語とする英語 L2 学習者を英語圏での留学経験の有無により, (1) 英語圏滞在経験の無い学習者 (SA 群), (2) 1ヶ月以上の英語圏への留学経験者 (NoSA 群) の 2 つに分けた. 多くのインプットを得て, 音声習得が進んでいると考えられる留学経験者群と, 非留学経験者群の間で, 知覚スペースの音素の距離に違いがあるかどうかを分析するため, それぞれの群における子音選択の正答率と, 混同した子音の選択率を比較した. また, 混同先と混同率に基づき, 子音間の知覚的距離を, 多次元尺度構成法により図式化し, 2 つの群の間で知覚マップがどのように変化しているかを分析した.

2. 方法

本研究では, 子音, 特に摩擦音の知覚スペースの変化を調査するため, 英語の 5 つの子音をターゲットとする知覚混同実験を行った. 実験で呈示した知覚のターゲット語は, CV 構造の単語であった. このターゲット語の子音(C)は, 英語の摩擦音 4 音 (*f, θ, s, ʃ*) と破裂音 1 音 (*t*) の合計 5 音で, 母音(V)は, */a, i/* の 2 音であった. 作成したターゲット語は, これらの 5 つの子音と 2 つの母音を組み合わせさせた語, 合計 10 語 (“*fa*”, “*fi*”, “*θa*”, “*θi*”, “*sa*”, “*si*”, “*ʃa*”, “*ʃi*”, “*ta*”, “*ti*”) であった. 2 つの母音環境のうち, [i] の前では, [s], [ʃ] の対立が日本語では

中和される ($s \rightarrow \int / _ i$).¹ よって, [a] の前では [f], [θ] が L1 にはない L2 音であるが, [i] の前では, この 2 音に加え, [j] も L1 音素ではない音であることになる。

実験では, 10 語のターゲット語の前に “Now I say” というキャリア文を付加した文の発話を使用した。実験の音声刺激としてアメリカ英語母語話者が発話した文音声を収録した。収録した音声のサンプリング周波数は 44,100Hz で, 量子化ビット数は 16bit であった。

本実験では, 収録した音声 (バブルノイズ無し: S/N 比 70dB) に加え, 高次の音韻処理を促すため, 音声に異なる 2 つの S/N 比 (0dB, 15dB) でバブルノイズを合成した音声を刺激音声として呈示した。また各刺激音声は平均音圧が 70 dB になるように標準化した。これらの作業には, Praat (Boersma & Weenink, 2019) を使用した。

実験の各試行では, ターゲット語を含む音声を刺激として呈示し, ターゲット語と同一の子音を含む 6 つの異なる実在語を回答の選択肢としてボタンで呈示した。たとえば, 聞き取りのターゲット語が “fa” の場合, “Now I say fa” という音声を呈示し, 画面には, “sand”, “tank”, “thank”, “fan”, “shine”, “child” の 6 語をボタンで呈示した。参加者は, 聴取したターゲット語の子音と同じ子音で始まる語を選択肢から選び回答した。上記の例の場合には, “fan” が正答の選択肢である。

実験には, 練習セッションを設け, このセッションでの正答率が 100% になり次第, 実験セッションに進むように設計した。実験セッションは呈示する音声の S/N 比が異なる 3 つのセッション (1: バブルノイズ無しの S/N 比 70dB, 2: 発話音声とバブルノイズの S/N 比が 15dB, 3: 発話音声とバブルノイズの S/N 比が 0 dB) で構成した。セッションの呈示順は参加者を問わず共通であったが, 各セッションにおける音声刺激の呈示順は参加者毎にランダム化した。

実験の構築及び実施には Inquisit (2018) を使用した。参加者は, PC または iOS 機器にヘッドフォンまたはイヤフォンを接続して使用し, Millisecond 社が用意するサーバーに同社によるアプリケーションを介してアク

セスし実験を実施した。

実験の参加者は, 日本語を母語とする 18 歳から 47 歳までの英語学習者, 合計 42 名であった。各参加者の英語圏での滞在経験の有無をもとに, (1) 英語圏滞在経験の無い学習者 (NoSA 群) 29 名と, (2) 1 ヶ月以上の英語圏への留学経験のある学習者 (SA 群) 13 名の 2 グループに分けた。

3. 結果

全体の正答率は, SA 群の方が NoSA 群より高かった (SA 群: 76.28%, NoSA 群 68.25%)。本実験で呈示した摩擦音ごとの正答率を, 参加者グループ間で比較した。以下の図 1 は後続母音が [a], 図 2 は後続母音が [i] の場合の, 摩擦音ごとの正答率を比較したものである。²

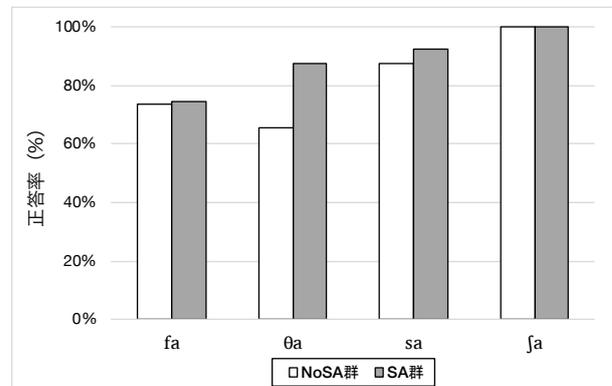


図 1. 後続母音が [a] の子音別正答率

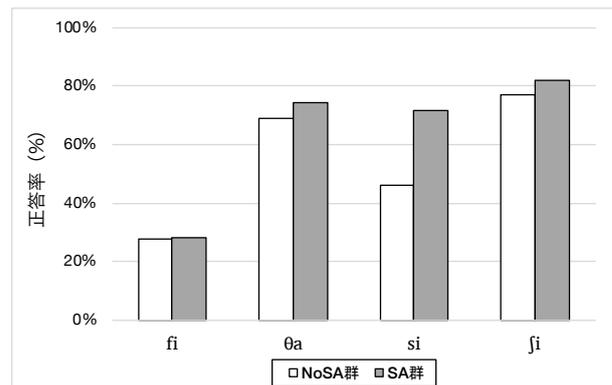


図 2. 後続母音が [i] の子音別正答率

¹ 日本語の /s/ の [i] の前での音は, 厳密には [e] であるが, 本論文では便宜上, 英語の類似音である [i] と同じ子音として議論を進める。

² Cutler et al (2004) にても報告されていたように, 本論文で行った実験の結果においても S/N 比の違いによって, グループ間の正答率の傾向に影響は見られなかった。よって, 本論文では 3 つの S/N 比のセッションでの回答を合わせて正答率を集計した。[i] が後続する

場合には, 前述のように [s] と [j] でも対立が中和されるため, /f, θ, j/ の 3 つの子音のどれもが L1 で音素ではない音であると言える。しかし, 中でも [f] の正答率はどちらの群においても低かった。Cutler et al (2004) によるノイズ下の知覚混同実験では, 英語母語話者を対象とした場合であっても, 語頭の f, θ の正答率は他の子音よりも低いという結果が示された。

図1, 2に見られるように, SA 群の方が NoSA 群よりも正答率が高く, またどちらの群でも, 後続母音が [a] である方が, [i] である場合よりも正答率が高かった. 日本語の /s,ʃ/ は, [a] の前のポジションでは, 対立的である. しかし, [i] の前では [s] と [ʃ] の対立は中和される. よって, [i] が後続する際のこれら2音を呈示した場合の正答率が [a] が後続する場合よりも低くなることは, 予想された結果であった.

[a] を後続母音とする刺激では, 全体的に正答率が高く, どの子音の場合にも2群の正答率間に有意な差は見られなかった.³ 一方, 後続母音が [i] の刺激では, [s] にのみ, 有意差が見られ, SA 群が NoSA 群の正答率よりも有意に高かった ($t(40)=2.429, p=.020, d=0.783$).

次に, Johnson (2011) で示された手法を参考に, 知覚混同率より多次元尺度構成法を用い, それぞれの参加者グループの摩擦音の知覚マップを作成した. 以下の図3, 4はそれぞれ, [a] が後続する場合の NoSA 群, SA 群の知覚マップである. グラフの3次元座標の値は Appendix に付した.

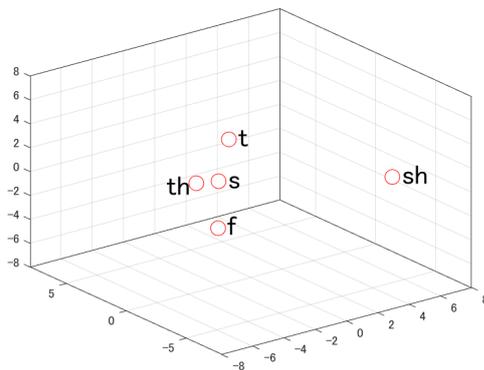


図3. [a]が後続する場合の無声摩擦音の知覚マップ: NoSA 群

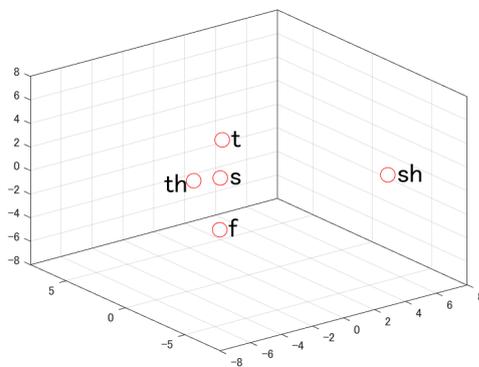


図4. [a]が後続する場合の無声摩擦音の知覚マップ: SA 群

図3, 4から, [a] が後続する場合には, L1 の音素である /t/, /s/, /ʃ/ の位置が離れており, 知覚においてこれらの音が明確に区別されていることが分かる. また, 図3, 4を比較すると SA, NoSA 群の間の知覚マップにはほぼ違いが無いことが分かる.

[a] の場合とは異なり, [i] が後続する環境での知覚では正答率も, 混同傾向もグループ間で異なった. 以下の図5, 6は [i] が後続する場合の, NoSA 群, SA 群の知覚マップである.

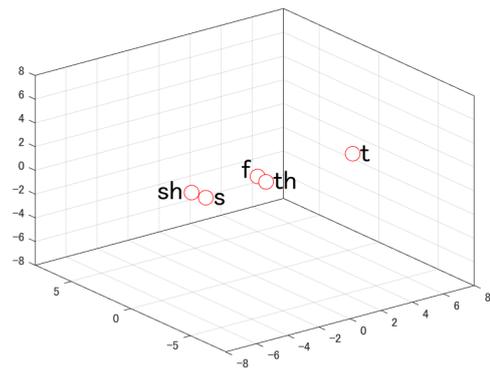


図5. [i]が後続する場合の無声摩擦音の知覚マップ: NoSA 群

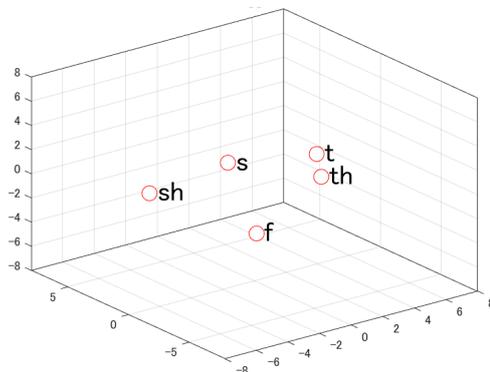


図6. [i]が後続する場合の無声摩擦音の知覚マップ: SA 群

図5より, NoSA 群では [i] が後続する場合, L1 では同環境で異音となる *s* と *f* の距離が近くなっている. また, 音響的に類似しており, Miller & Nicely (1955) により, 英語母語話者でもノイズ下で混同が多いことが報告されている *f* と *θ* も近くに配列されている. また, 図5の NoSA 群の [i] のマップでは *s, f* と *f, θ* そして *t* と, 3つの塊を形成しており, *f, θ* は, 知覚マップ上, 摩擦音の *s, f* と破裂音である *t* との間に位置していることも読み取れる. 一方, 図6の SA 群では, [a] の図

³ [θa] の正答率には2群の間に有意傾向が見られた. ($t(40)=2.185, p=.065, d=0.678$)

同様、5つ音が分かれて配列されている。[a] の場合の図4と比較すると、[i]の前では、*s, f*が*f*と*θ*の間に位置し、*f*と*s, f*との距離が近づいている。NoSA群(図5)とSA群(図6)を比較すると、NoSA群で距離が近い、*s*と*f*が、SA群では離れて位置しており、L1で異音である[si]と[ji]が、SA群では知覚的に区別されてきていることが分かる。また、L1音である[t]と[s]との間の距離は、NoSA群のマップと比べ、SA群では近くなっている。これは*f, θ*の習得により、他の音の区別が進んだことによるのかもしれない。

上記の多次元尺度構成法による知覚マップによる分析は、混同傾向により、音声間の距離を図式化したものであるが、実際にはどの音をどの音に聞き違えたのかの方向性には非対称性が見られる。しかし、Johnson(2011)でも指摘されているように、多次元尺度構成法では混同の「方向性」は考慮されていない。そこで、以下に、それぞれの子音刺激に対する選択肢の選択率をまとめた表を提示した。

表1. [a] が後続する無声摩擦音に対する選択率：NoSA群

	<i>fa</i>	<i>sa</i>	<i>fa</i>	<i>ta</i>	<i>θa</i>
<i>fa</i>	73.56%	4.60%	0.00%	0.00%	21.84%
<i>sa</i>	0.00%	87.36%	0.00%	0.00%	12.64%
<i>fa</i>	0.00%	0.00%	100.00%	0.00%	0.00%
<i>θa</i>	20.69%	12.64%	0.00%	1.15%	65.52%

表2. [a] が後続する無声摩擦音に対する選択率：SA群

	<i>fa</i>	<i>sa</i>	<i>fa</i>	<i>ta</i>	<i>θa</i>
<i>fa</i>	74.36%	2.56%	0.00%	0.00%	23.08%
<i>sa</i>	0.00%	92.31%	0.00%	0.00%	7.69%
<i>fa</i>	0.00%	0.00%	100.00%	0.00%	0.00%
<i>θa</i>	2.56%	7.69%	0.00%	2.56%	87.18%

表3. [i] が後続する無声摩擦音に対する選択率：NoSA群

	<i>fi</i>	<i>si</i>	<i>fi</i>	<i>ti</i>	<i>θi</i>
<i>fi</i>	27.59%	11.49%	12.64%	2.30%	45.98%
<i>si</i>	0.00%	45.98%	36.78%	0.00%	16.09%
<i>ji</i>	0.00%	17.24%	77.01%	0.00%	2.30%
<i>θi</i>	12.64%	5.75%	10.34%	1.15%	68.97%

表4. [i] が後続する無声摩擦音に対する選択率：SA群

	<i>fi</i>	<i>si</i>	<i>fi</i>	<i>ti</i>	<i>θi</i>
<i>fi</i>	28.21%	2.56%	2.56%	0.00%	66.67%
<i>si</i>	0.00%	71.79%	10.26%	0.00%	17.95%
<i>ji</i>	0.00%	17.95%	82.05%	0.00%	0.00%
<i>θi</i>	15.38%	5.13%	0.00%	5.13%	74.36%

表1, 2より、NoSA群とSA群ではともに[fa]を*θa*と混同する回答が20%を超えている一方、[θa]に対して*fa*の選択肢を選んだ割合はNoSA群とSA群では違いが見られる。NoSA群では20.69%と、[fa]と[θa]との混同は対称的であったが、SA群では[θa]に対して*fa*を選んだ割合は低く(2.56%)、混同には非対称性が見られた。表1, 2の[θa]での他の子音との混同率を比較すると、SA群では[θa]を*fa, sa*と混同する割合がNoSA群よりも低下し、正答率が高くなっていることが分かる。

また、子音に[i]が後続する場合の混同率を示した表3, 4から、[a]の場合と同様に[fi]と[θi]との混同が多いことが分かる。NoSA群とSA群のどちらにおいても、[fi]に対する*θi*の選択率は高いが、[θi]に対しての*fi*の選択は、逆のケース([fi]に対する*θi*の選択)ほど多くは無く、混同は非対称的であると言える。つまり、日本語母語話者は[fi]と[θi]を混同して同じ音だと知覚する傾向があるが、どちらの音についても*fi*を避け、*θi*であると知覚するプリファレンス(指向性)が見られる。この指向性により、[f]よりも[θ]の正答率の方が高くなったと考えられる。さらに、表3, 4を比較すると、NoSA群では、[fi]の刺激に対する混同が、*θi*を中心としつつも、*si, fi*にも広がっているのに対し、SA群では*si, fi*との混同が抑えられ、*θi*の選択率が増えている。これはSA群では、[si], [θi], [ji]が知覚上、区別されるようになったことによると考えられる。*fi*の選択を避け、異なる摩擦音を選択するにあたり、[si], [ji], [θi]が知覚的に区別されているため、最も[fi]と音響的に近いとされる*θi*に回答が集中したのであろう。非対称性は[si], [ji]の混同にも見られた。表3のNoSA群の回答では、[si]の正答率よりも[fi]の正答率の方が高かった。また、[si]を*fi*とした回答が36.78%であったのに対し、[ji]を*si*とした回答は17.24%と低くなっている。つまり、NoSA群では、[si], [ji]を混同し、*fi*であると捉える指向性があると言える。一方、図4のSA群では、[si]を*fi*と混同した割合は

10.26%と、*ji* への指向性が抑制されている。このことにより、SA 群の [si] の正答率が高くなったのだと言える。また、図6のマップにおいても、*s*, *f* の距離が離れていることも、SA 群で [si] と [ji] の区別がなされていることを示している。

4. まとめ

本研究では L2 音声習得が進むことにより、L1 音、および L2 音の音声知覚マップがどのように変化するかを調査した。ノイズ下における摩擦音の知覚混同実験により、[i] が後続する環境において、SA 経験者と非 SA 経験者の音声知覚マップには違いが見られた。また、SA 経験者のマップでは、L1 では非対立的である *si* と *ji* の区別が進んでいることが示された。また、摩擦音がどの音と知覚混同されたのかを示す選択率の比較からも、SA 経験者において、*s*, *f*, *θ* の区別が進んでいることが分かった。

References

- Best, C. T. (1995). A direct realist view of cross-language speech perception. *Speech perception and linguistic experience: Issues in cross-language research*. Wnifred Strange ed., Timonium, MD: York Press, 171-204.
- Boersma, P. and Weenink, D. (2019). Praat: doing phonetics by computer [Computer program]. Version 6.0.50, retrieved 24 April 2019 from <http://www.praat.org/>
- Brown, C. A. (1998). The role of the L1 grammar in the L2 acquisition of segmental structure. *Second Language Research*, 14 (2), 136-193.
- Chang, C. B. (2012). Rapid and multifaceted effects of second-language learning on first-language speech production. *Journal of Phonetics*, 40, 249-268.
- Cook, V. J. (ed.) (2003). *Effects of the second language on the first*. Clevedon: Multilingual Matters.
- Cutler, A., Weber, A., R. Smits, and N. Cooper (2004). Patterns of English phoneme confusions by native and non-native listeners. *Journal of Acoustical Society of America* 116, 3668-3678.
- Flege, J. E. (1995). Second language speech learning Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research*, 233-277. Baltimore: York Press.
- Flege, J. E. (1987). The production of “new” and “similar” phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics*, 15, 47-65.
- Guion, S. G. (2003). The vowel systems of Quichua-Spanish bilinguals: Age of acquisition effects on the mutual influence of the first and second languages. *Phonetica*, 60, 98-128.
- Harada, T. (2003) L2 Influence on L1 Speech in the Production of VOT. *Proceedings of the 15th International Congress of Phonetic Sciences at Barcelona*, 1085-1088.
- Johnson, K. (2011). *Acoustic and Auditory Phonetics*, 3rd edition, Chichester : Wiley-Blackwell, 2012, 127-137.
- Miller, G. A., and P. E. Nicely (1955). An analysis of perceptual confusions among some English consonants. *Journal of Acoustical Society of America*, 27, 338-352.
- Strange, W. (2011). Automatic selective perception (ASP) of first and second language speech: A working model. *Journal of Phonetics*, 39, 456-466.
- Tice, M., & Woodley, M. (2012). Pagnettes & bastries: Novice French learners show shifts in native phoneme boundaries. A paper presented at the 86th Annual Meeting of the Linguistic Society of America, Portland, Oregon.

Appendix

A-1. 多次元尺度構成法による [a] が後続する子音の3次元座標

	NoSA			SA		
	x	y	z	x	y	z
fa	-3.377	-1.670	-1.932	-3.318	-1.876	-2.146
sa	-3.322	-1.635	1.967	-3.241	-1.822	2.198
ja	6.138	-3.813	-0.008	6.198	-3.652	-0.013
ta	2.860	5.561	0.000	2.574	5.591	-0.001
θa	-2.298	1.557	-0.026	-2.212	1.759	-0.038

A-2. 多次元尺度構成法による[i]が後続する子音の3次元座標

	NoSA			SA		
	x	y	z	x	y	z
fi	0.471	0.376	0.000	-1.886	-2.641	-2.217
si	-3.517	-0.477	0.000	-2.080	-0.522	2.738
ʃi	-3.638	0.557	0.000	-4.315	3.037	-0.640
ti	6.350	0.054	0.000	5.651	2.221	-0.517
θi	0.334	-0.511	0.000	2.630	-2.094	0.636

*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習得過程における音韻処理変化と知覚スペースの再調整」課題番号: 18K00848) の助成を受けた。