

音響的特徴に基づく中国語声調のカテゴリー知覚
—母語話者と日本人学習者を対象に—

朱虹（中南財經政法大学・上智大学理工学部） 吉本啓（東北大学国際文化研究科）

荒井隆行（上智大学理工学部）

zhuhong200812@yahoo.co.jp, kei@compling.jp, arai@sophia.ac.jp

1. はじめに

音声に関するカテゴリー知覚は、Lieberman et al. (1957) によっではじめて報告された。最初は主に分節音素である母音と子音を中心とするカテゴリー知覚の研究であったが、90年代の後半から、超分節音素 (suprasegmental phoneme) を対象とするカテゴリー知覚の研究 (Abramson 1979; Francis et al. 2003) が増えてきた。その中、中国語声調 (Tone) (声調の種類が方言により異なるが、本稿では標準語における四つの声調を中心に扱う) に対するカテゴリー知覚の検討は Wang (1976) をはじめ、多くの研究者 (Hallé et al. 2004; Xu et al. 2006; 王他 2010; 榮 他 2013) に注目された。しかしながら、四つの声調における音響的な特徴は異なり、音声刺激の作成に使われるパラメータも異なるため、声調に対する知覚はカテゴリー的か否かはまた議論されている。

声調に関する音響的な研究において、莊 他 (1975) は基本周波数 (以下、fo) の変化範囲、振幅、持続時間から分析し、第1声 (以下、T1) は「平らなパターン」、第2声 (以下、T2) は第3声 (以下、T3) と同じ「凹型パターン」、第4声 (以下、T4) は「放物線パターン」のようにまとめた。さらに、声調の習得研究では、中国語母語話者のみならず、T1とT4、T2とT3は弁別しにくいペアとして多数の研究によって指摘され、混同しやすいペアになる理由は、T1とT4のfoが始点において類似していること、ならびにT2とT3のfo曲線が類似していることが報告されている (Huang 2001; So 2005; Lee, 2008)。

また、声調のカテゴリー知覚に関する習得研究において、主に中国語母語話者および欧米人母語話者を対象にして、声調言語と非声調言語の言語経験に着目し、検討されてきた。特に日本人中国語学習者を対象とする研究は非常に少ないが、その中の一つに張 (2010) がある。張 (2010) はT1からT2までの刺激連続体を使い、それぞれ同定実験と弁別実験を行った結果、日本人学習者は母語の経験の影響により音声の高さに敏感で、習熟度が高いほど声調に対するカテゴリー知覚の程度が高くなると述べた一方で、張 (2010) はT2を「上昇調」として扱い、「凹型パターン」に相応しくないと考える。従って、本研究は4つの声調の音響特徴に基づき、カテゴリー知覚の実験的手続きに従い、日本人中国語学習者および中国語母語話者における声調の習得しにくいT1とT4、T2とT3に対する認知モードを究明することを研究目的とし実験を行う。

2. T1 と T4 に関するカテゴリー知覚

2.1. 終点の fo に基づく音声刺激

T1 の fo は高く、平らなパターンに対して、T4 は始めの fo は高いが、放物線的パターンで急激に下降するという特徴を持っている (莊 1975)。また、中国語声調の記述に広く用いられる「五度制」によって、低い方から 1~5 の数字を振り、T1 と T4 をそれぞれ [55]、[51] のように記述する。すなわち、T1 と T4 の始点は大体同じピッチレベルであるが、終点の fo ははるかに異なることがわかった。

そこで、T1 と T4 に関して、終点の fo をパラメータとして単母音 /u/ に関する音声刺激連続体を作成する。T1 から T4 までの連続体は図 1 に示すように、持続時間を T1 と T4 の平均値の 335 ms、始点の fo を平均値の 273 Hz にした。音声合成のパラメータとなる終点の fo については、T1 の終点の 266 Hz から T4 の終点の 166 Hz まで 10 Hz の間隔で計 11 個の刺激を Praat によって作成した。

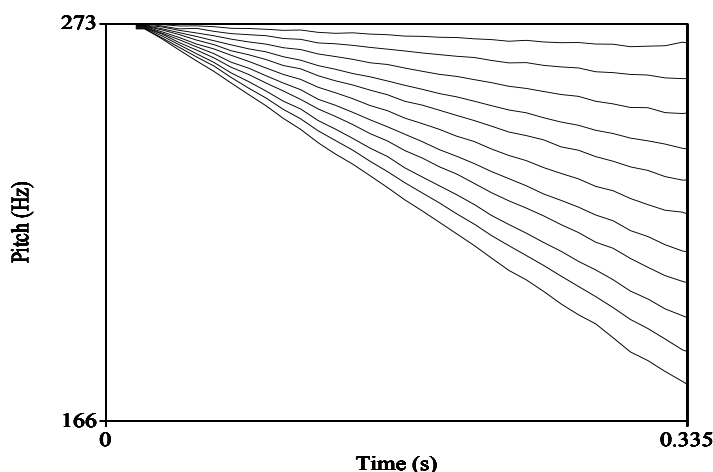


図 1: T1 から T4 までの刺激連続体

2.2. 被験者と実験手順

被験者は日本人中国語上級者 20 名、日本人中国語初級者 20 名、北京出身の中国語母語話者 20 名の三つのグループに分けられた。全員とも聴覚に障害がないことを確認した。上述した被験者にそれぞれ同定実験と弁別実験を行った。同定課題では、11 個の音声刺激が 1 つずつランダムで 10 回提示され、被験者にパソコン画面上に選択肢のボタンを提示し判断させ、その刺激が何かを答えてもらった。なお、刺激 40 個おきに休憩をとってもらった。一方、弁別課題では使われる刺激音は同定実験に使用された 11 個の刺激音から 2 ステップ離れたステップ番号「1, 3」、「2, 4」、「3, 5」、「4, 6」、「5, 7」、「6, 8」、「7, 9」、「8, 10」、「9, 11」の 9 組の刺激ペアを用い、連続して提示され、被験者にこの 2 つの刺激が同じかを答えさせた。合計は 90 組 (9 組の刺激ペア×10 回繰り返し) となった。

2.3. 実験結果

同定実験ならびに弁別実験では、刺激の間に有意差が見られた。日本人学習者も中国語母語話者も同定曲線 (図 2) では、連続している刺激の間に同定率が急激に変化し、急峻な下降が見られた。この急な下降点は弁別曲線 (図 3) におけるピークと一致したので、T1 と T4 に対する知覚はカテゴリー知覚だと判断した。日本人学習者は習熟度が高くなると、知覚範疇化の程度も高くなり、カテゴリー知覚の境界は始点が 273 Hz、終点が 216 Hz~226 Hz のところである。一方、母語話者は範疇内の刺激に対する反応は敏感ではなく、弁別率が低い、異なる範疇間の刺激に対する反応は敏感で、弁別率が日本人学習者より高かった。知覚範疇化の程度は日本人学習者より高い、カテゴリー境界は始点が 273 Hz、終点が 226 Hz のところである。

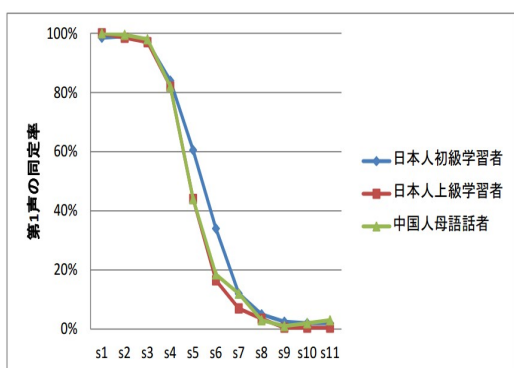


図 2: T1 と T4 に関する同定実験の結果

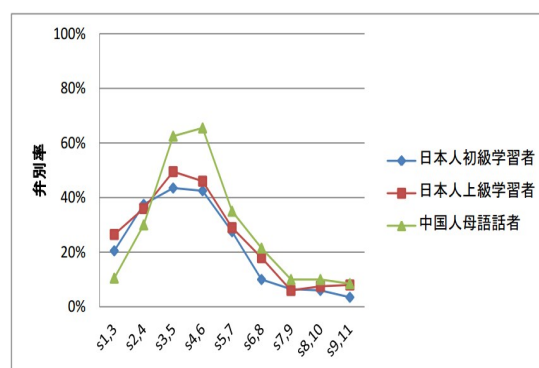


図 3: T1 と T4 に関する弁別実験の結果

3. T2 と T3 に関するカテゴリー知覚

3.1. Δf_0 と TTP に基づく音声刺激

T2 と T3 の弁別に関する重要な聴覚的キューとして、転換点 (ピッチ曲線が下降から上昇へと変化する点) までの時間 (Time of Turning Point, 以下、TTP) および Δf_0 (音節開始点から転換点までの f_0 の差) という 2 つが指摘されている (Shen et al. 1993; Moor & Jongman 1997)。朱 他 (2012) はこの 2 つのキューが日本人学習者および母語話者の T2 と T3 の弁別に重要な役割を果たしているを実証した。従って、本実験は TTP、 Δf_0 をパラメータにして、T2 と T3 のような「凹型パターン」の認知モードを検証する。

単母音 /u/ に関して、 Δf_0 は、10 Hz の間隔で 30 Hz から 80 Hz まで計 6 個、TTP は、20 ms の間隔で 50 ms から 190 ms まで計 8 個、合計 48 個の刺激を作成した。40 個の刺激に基づき、「TTP を変化させる連続体」6 組 (図 4)、「 Δf_0 を変化させる連続体」8 組 (図 5) 及び「 Δf_0 と TTP を同時に変化させる連続体」18 組 (図 6) という 3 種類の連続体 (合計 32 組) を作成した。

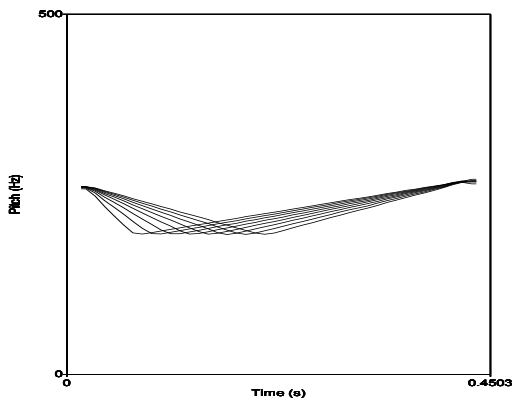


図 4: TTP を変化させる連続体の一例

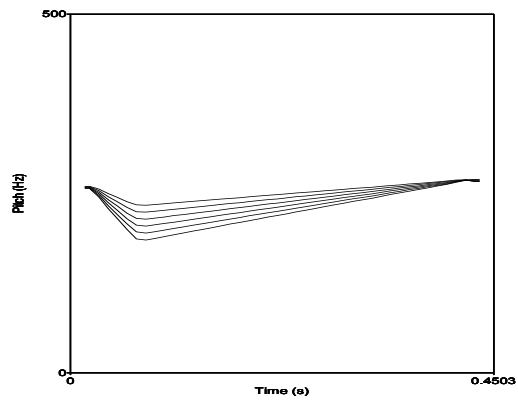


図 5: Δfo を変化させる連続体の一例

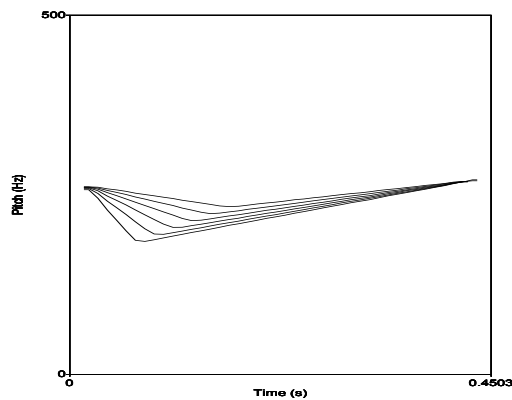


図 6: Δfo と TTP を同時に変化させる連続体の一例

3.2. 被験者と実験手順

被験者は 2.2 と同じ母語話者を含め三つのグループがあり、合計 60 名である。カテゴリー一知覚の実験的な手続きに従い、それぞれ同定実験と弁別実験を行った。同定実験では、刺激は 48 個あり、3 回繰り返して 144 個 (48 個×3 回) になる。弁別実験では、32 組の連続体から 2 ステップ離れた 116 組の刺激ペアを抽出し、2 回繰り返して合計 232 組 (116 組×2 回) となった。両実験もパソコンでランダムな順序で呈示し回答を得た。

3.3. 実験結果

3.3.1 TTP を変化させる連続体の結果

日本人初級学習者も上級学習者も TTP を変化させる連続体において、TTP が 50 ms~190 ms の間に変化してもカテゴリー知覚は発見されなかった。それに対して、母語話者は Δfo が 50 Hz またそれ以上になる連続体に、カテゴリー知覚の傾向が生じ、知覚のカテゴリー境界も明らかになる。凹型パターンの声調において fo の初めの下降は母語話者にとって、重要な役割を果たしていることがわかった。

3.3.2 Δf_0 を変化させる連続体の結果

Δf_0 を変化させる連続体の中で、カテゴリ知覚が発見されたのは、日本人初級学習者については TTP=90 ms であったが、上級学習者の場合は TTP=130 ms であった。それに対して、母語話者の場合は TTP が 170 ms である連続体にカテゴリ知覚が発見された。要するに、習熟度が高くなるにつれて、カテゴリ知覚が発見された連続体における転換点が遅くなる。また、日本人学習者は、 Δf_0 を変化させる 8 組の連続体に対しほぼカテゴリ知覚の傾向が見られた。そのため、日本人学習者にとって、 Δf_0 が T2 と T3 に関するカテゴリ知覚の判断では重要であることはわかった。

3.3.3 Δf_0 と TTP を同時に変化させる連続体の結果

Δf_0 と TTP を同時に変化させる連続体は 18 組あり、連続体の刺激数によって、「刺激数が 3 つある」、「刺激数が 4 つある」、「刺激数が 5 つある」、「刺激数が 6 つある」と 4 種類に分けられる。それぞれの判断基準に基づき、日本人初級学習者は 2 組の連続体にカテゴリ知覚が発見され、上級学習者と母語話者はそれぞれ 4 組、5 組の連続体からカテゴリ知覚が見られた。3.3.1、3.3.2 を合わせ、32 組の連続体に関するカテゴリ知覚の結果は表 1 のようになる。

表 1: カテゴリ知覚が発見された連続体

	TTP	Δf_0	Δf_0 と TTP	合計
初級者	なし	1 組	2 組	3 組
上級者	なし	1 組	4 組	5 組
母語話者	2 組	1 組	5 組	8 組

つまり、日本人学習者は T2 と T3 に対してカテゴリ知覚の傾向が見られ、習熟度が高くなると、知覚範疇化の程度も高くなる。また、カテゴリ境界は Δf_0 が 60 Hz、70 Hz のところに集中していることが観察された。一方、母語話者についても、T2 と T3 に対する知覚はカテゴリ知覚の傾向が見られ、カテゴリ境界を全体的に見れば、TTP が早く来る場合は Δf_0 が高くなるが、TTP が遅く来る場合は Δf_0 が低くなる。特に、 Δf_0 は 30 Hz、40 Hz である場合カテゴリ境界が見いだされないため、ピッチ曲線で最初の下降の幅が小さい場合 (50 Hz 以下) には T2 と T3 を区別できないことがわかった。

4. 結論

本稿は日本人学習者が声調知覚の習得において混同しやすい T1 と T4 および T2 と T3 に対して、これらの区別に重要な音響パラメータを用い、カテゴリ知覚について検討した。その結果、カテゴリ知覚またはその傾向が見られ、使われた音響パラメータが日本人学習者および母語話者の声調習得に重要な役割を果たしていることがわかった。

参考文献

- 栄蓉,石峰 (2013) 「音高和时长对普通话阴平和上声的听感影响」『語言科学』 2, 17-26.
- 王韞佳,覃夕航 (2012) 「再論普通話陽平和上声的感知」 第10届中国語言学學術會議
- 許勢常安 (1994) 「専大の第十一回中国語聴解力コンテストについて」『専修商学論集』 58, 319-357.
- 朱虹,張鵬,張立波,北原良夫,吉本啓 (2012) 「日本人学習者による中国語声調の習得の研究—転換点と ΔF_0 の視点から—」『国際文化研究』 18, 75-81
- 莊秋広 (1975) 「標準中国語の単音節語の四声の音響的特徴」『日本音響学会誌』 31, 369-380.
- 張林軍 (2010) 「日本留学生漢語声調的カテゴリー化知覚」『語言教学与研究』 13, 9-15.
- Abramson, A.S. (1979) “The noncategorical perception of tone categories in Thai”. *Frontiers of speech communication Research*, 127-134.
- Francis, A.L. (2003) “On the (non) categorical perception of lexical tones”. *Perception & Psychophysics*, 65, 1029-1044.
- Hallé, P., Chang Y.H. & Best, C.T. (2004) “Identification and Discrimination of Mandarin Chinese Tones vs. French listeners”. *Journal of Phonetics*, 32, 395-421.
- Huang, T. (2001) “The interplay of perception and phonology in Tone3 sandhi in Chinese Putonghua”. In E, Hume and K.Johnson (eds.) *Studies on the Interplay of Speech Perception and Phonology, Ohio State University Working Papers in Linguistics*, 55, 23-42.
- Lee, C.-Y., Tao, L., & Bond, Z.S. (2008) “Identification of acoustically modified Mandarin tones by native listeners”. *Journal of Phonetics*, 36, 537-563.
- Lieberman, A.M., Harris, K.S., Hoffman, H.S. & Griffith, B.C. (1957) “The Discrimination of speech sounds within and across phonemic boundaries”, *Journal of Experimental Psychology*, 54, 358-368.
- Moore, C.B., & Jongman, A. (1997) Speaker normalization in the perception of Mandarin Chinese. *J.Acoust.Soc.Am.* 102.1864-1876
- Shen, X.-S., Lin, M., & Yan, J. (1993) F_0 turning point as an F_0 cue to tonal contrast: A case study of Mandarin tones 2 and 3. *J.Acoust.Soc.Am.* 93, 2241-2243
- So, C.K. (2005) “The influence of L1 prosodic background on The learning of Mandarin tones: patterns of tonal confusion by Cantonese and Japanese native listeners”. *Proceedings of the 2005 annual conference of the Canadian Linguistic Association*, 1-7.
- Wang.W.S.-Y. (1976) “Language Change”. *Annals of the New York Academy of Sciences*, 28, 61-72.
- Xu, Y. S., Gandour, J. & Francis, A. L. (2006) “Effects of Language Experience and Stimulus Complexity on the Categorical Perception of Pitch Direction”, *J. Acoust.Soc.Am.* 120, 1063-1074.

中国共通語話者による第三言語の語頭破裂音の知覚習得について

劉佳琦(復旦大学外国言語文学学院)・陳忠敏(復旦大学中文系)
 曾婷(復旦大学外国言語文学学院)・鹿秀川(復旦大学外国言語文学学院)
 jiaqiliu@fudan.edu.cn

1. はじめに

中国における大学外国語教育は空前の多言語環境に直面している。英語以外に、大学在籍期間内に第三言語(L3)を学ぶ機会を望む学生が増えている。ただ、L3の習得過程において、母語(L1)と第二言語(L2)に影響されることがあると言われている(Williams and Hammarberg 1998, Cenoz, J. et al. 2001)。このような影響が最も顕著に現れるのは音声である(Onishi 2016)。教育現場の経験からわかるように、外国語学習初期には、有声・無声破裂音の指導と習得は教師にも学習者にも困難である。今まで、Second Language Acquisition(SLA)分野において、主に母語背景の異なる英語学習者を対象とした研究が盛んに行われてきた(Flege 1992)。しかし、管見の及ぶ限り、Third Language Acquisition(TLA)分野においては、有声・無声破裂音に関する習得研究は数が少なく、多くの習得表層現象がまだ解明されていない(Trembley, 2007; Llana et al., 2010; Wunder, 2010; Wrembel, 2014)。

本研究は中国の大学で最も学習人数の多いL3日本語、L3スペイン語学習者を対象に、中国共通語話者におけるL3破裂音知覚の差異を比較し、言語間の相違性と類似性がL3破裂音知覚に与える影響を明らかにする。したがって、本研究の成果は音声習得理論問題の究明に貢献し、多言語的背景におけるL3の音声教育改善の裏付けとなる。

2. L1、L2、L3の破裂音体系

破裂音の有声性は多くの音響パラメーターと関係している。その中、破裂音の音韻対立を判断するには、閉鎖区間の声帯振動(prevoicing)の有無と開放後の帯気(aspiration)という音響的特徴を観察するのは一般的である。音響的パラメーターVoice Onset Time (Lisker and Abramson 1964)は開放から声帯振動開始までの時間長で、prevoicingおよびaspirationという重要な音響的特徴を同時に捉えることができる。そのため、ここではL1、L2とL3破裂音のVoiced Onset Time(VOT)パターンを観察・比較する。

本研究の協力者は中国共通語(Mandarin Chinese)話者である。吳宗濟(1988)は中国共通語の有気・無気破裂音の音韻対立を生理的・音響的角度から分析を行った。注目すべきなのは、中国共通語話者がL1の有気・無気対立を判断するキューとなっているのは帯気性の強さである。中国共通語の語頭有気・無気破裂音のVOT参照値は表1のとおりである(表1は鮑懷翹, 林茂燦(2014)を参照して、作成した)。

L2英語の語頭破裂音は有声・無声対立を成すと言われているが、実際Lisker and Abramson(1964)の音響的実験から、英語の語頭有声破裂音は声帯振動(prevoicing)を伴わないことがわかった。Klatt (1975)とDocherty (1992)の研究でも、英語母語話者の語頭有声破裂音はプラスVOT値で、閉鎖区間に声帯は振動しないことが明らかになった(表2)。

表 1: 中国共通語有気・無気破裂音の VOT 参照値

Voicing	Place of Articulation	Aspiration	VOT
voiceless	Labial [p ^h]	aspirated	92.5 ms
voiceless	Alveolar [t ^h]	aspirated	102 ms
voiceless	Velar [k ^h]	aspirated	96.5 ms
voiceless	Labial [p]	unaspirated	7.5 ms
voiceless	Alveolar [t]	unaspirated	6 ms
voiceless	Velar [k]	unaspirated	14.5 ms

表 2: 英語の語頭破裂音の VOT 参照値

	Klatt (1975)	Docherty (1992)
[p ^h]	47 ms	42 ms
[t ^h]	65 ms	64 ms
[k ^h]	70 ms	63 ms
[p]	11 ms	15 ms
[t]	17 ms	21 ms
[k]	27 ms	27 ms

L3 日本語と L3 スペイン語は L1、L2 と異なる破裂音の音韻対立を持っている。(1)日本語の語頭破裂音は無声有気音[p^h、t^h、k^h]と有声音[b、d、g]である (The International Phonetic Association 1999:117); (2)スペイン語の無声破裂音の場合は、開放後すぐに母音が続き、帯気性が弱い。有聲破裂音の場合は、閉鎖区間における声帯振動(prevoicing)特徴が顕著に見られる(Abramson and Lisker 1973)。

4つの言語の中で、意味弁別機能を持つ有聲・無声あるいは有気・無気破裂音の対立を表3のように整理した。学習者のL1とL2の語頭破裂音は帯気性(aspiration)によって音韻対立を区別しているが、L3日本語とスペイン語は閉鎖区間の声帯振動(prevoicing)の有無によって有聲・無声の音韻対立を成していることがわかった。

表 3: L1, L2, L3 の語頭破裂音の音韻対立

Language	Voiceless		Voiced
	Aspirated	Unaspirated	
L1	Mandarin Chinese	p ^h , t ^h , k ^h	p, t, k
L2	English	p ^h , t ^h , k ^h	p, t, k
L3	Japanese	p ^h , t ^h , k ^h	b, d, g
	Spanish		p, t, k

3. 知覚実験

3.1. 実験協力者

知覚実験の協力者は39名である。二つのグループに分けられている。(1)L1中国共通語, L2英語, L3日本語(N=20); (2)L1中国共通語, L2英語, L3スペイン語(N=19)。学習歴は2ヶ月で、初級学習者にあたる。発音と表記は習得済みである。

3.2. 実験語

知覚実験の音声刺激は日本語とスペイン語各24語(刺激語)+12語(非刺激語)である。調

音点は両唇 pb、歯茎 td、軟口蓋 kg となっている。音声刺激は語頭に位置し、後続母音は[a]に統一されている。音声刺激にはアクセントが置かれていない。

3.3. 実験手順

知覚実験の音声刺激は日本語母語話者 2 名(東京都出身)、スペイン語母語話者 2 名(Segovia、Málaga 出身)に提供してもらった。合計 48 個の音声刺激をそれぞれの言語のキャリアセンテンス、日本語「これは_____」、スペイン語「Es_____」に入れた。録音は防音室あるいは静かな空間で行われ、録音機材は TASCAM DR-44WL linear PCM recorder(44.1-kHz sampling rate, 16-bit quantization level), AKG C544L head micromic。本研究は Praat 6.0 を使って知覚実験の音声刺激を作成した。知覚実験は Praat 6.0 の「ExperimentMFC 6」スクリプトを使って実施した。実験機材はノートパソコン、SONY MDR-ZX110NC headphone である。

データ処理の際にはまず SPPAS(Ver.1.8.6)を使って自動的にアノテーションとラベリング作業を行い、そのあと手作業で確認した。それから、Daniel Hirst(Ver.2010/10/30)が作成した「analyse_tier.praat」スクリプトで各ティアの音響データを収集した。本研究は R(Ver.3.4.0)を使って、データ整理、統計分析、図表作成をした。

4. 実験結果と考察

4.1. L2 語頭破裂音のVOT

学習者の L2(英語)破裂音の習得現状を把握するため、生成実験を行った。その手順は本文 3.3 と一致している。[p^h, t^h, k^h]と[p, t, k]が語頭に位置する単音節語を実験語とした。Independent-Samples t-test の結果、L3 日本語学習者と L3 スペイン語学習者の間に L2 英語破裂音 VOT 値の有意差が存在しない。本知覚実験を行う際に、両グループの L2 英語破裂音の習得状況が同じであることがいえる。中国人学習者の英語破裂音 VOT 値分布は、有気破裂音[p^h, t^h, k^h]は 65ms から 95ms、無気破裂音[p, t, k]は 10ms から 25ms となっている。学習者が生成される L2 語頭破裂音の VOT 値はプラスで、有気か無気かは VOT の数値によって分けられている。学習者は 30~35msVOT を境に L2 有気・無気破裂音を分けている。

学習者の L2 英語は英語母語話者の語頭無気破裂音の音声特徴とほぼ同様であり、またそれと学習者 L1 の無気破裂音の特徴も同様であることがわかった。一方、学習者の L2 無声有気破裂音の場合は、学習者 L1 と L2、英語母語話者の VOT 平均値が少し差異が存在しているが、VOT 分布区間が部分的に重なっていることがわかった。

4.2. L3 語頭破裂音の知覚実験の結果と考察

4.2.1. L3 語頭無声破裂音

L3 有声・無声破裂音の知覚実験の結果は図 1 のようである。Independent-Samples t-test の結果、L3 学習者の正聴率が L3 母語話者より有意に低いことがわかった。Independent-Samples t-test の結果、(1) 日本語無声破裂音 $t(146)=7.038, p=6.995e-11$ 、日本語有声破裂音 $t(479)=3.79$,

p=.00016、(2) スペイン語無声破裂音 $t(111)=7.64, p=8.099e-12$ 、スペイン語有声破裂音 $t(80)=6.17, p=2.636e-08$ 。

知覚実験の結果から、両L3グループの間に正聴率の有意差が存在していることがわかった。L3 スペイン語より、L3 日本語の無声破裂音の正聴率の方が有意に高いことがわかった。

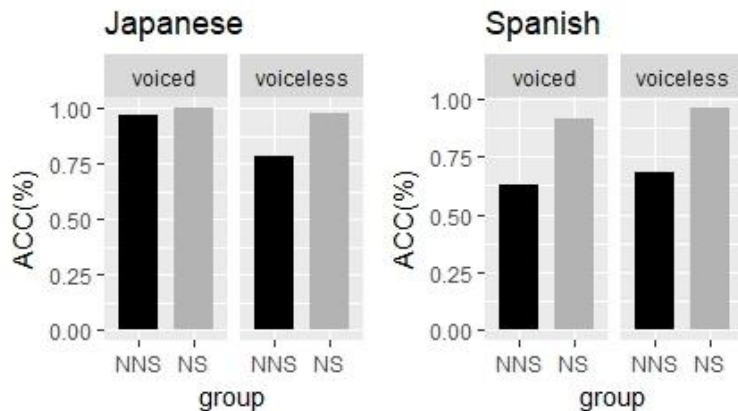


図 1: L3 有声・無声破裂音知覚実験の平均正聴率

Independent-Samples t-test の結果：

$t(907.36)=3.51, p<.001$ 。知覚実験に使用される母語話者の無声破裂音刺激の VOT 分布区間は図 2 のようである。図 2 からわかるように、両 L3 の VOT 値とその分布区間が異なっている。日本語の無声破裂音刺激の VOT 平均値(30ms)が大きく、Median : 35.08ms、Standard Deviation (SD) は 18.48 で、VOT 分布区間も大きい。それと比べると、スペイン語音声刺激の VOT Median : 15ms、Standard Deviation (SD) は 9.60 で、VOT 分布区間が 30ms を下回る。30msVOT は普遍的な意義を持つ音響パラメーターである。一連の通言語的研究から、世界の言語の多くは 30msVOT を境に有気音と無気音を区別していることがわかっている Keith(2003 : 101-102)。

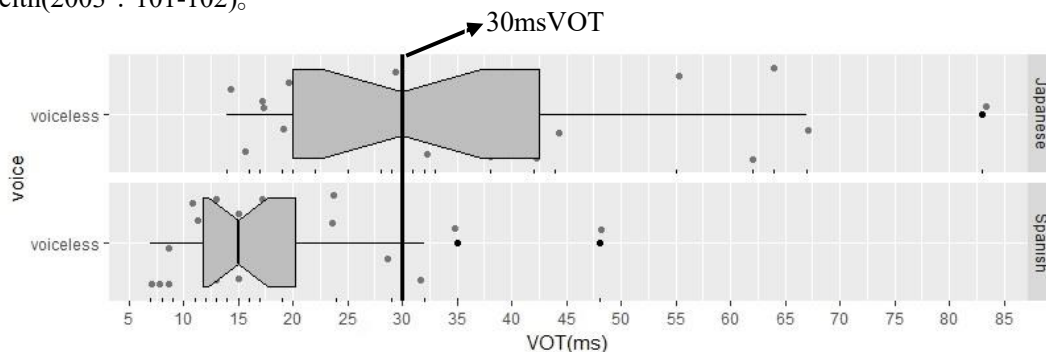


図 2: 日本語とスペイン語の無声破裂音 VOT 値の分布区間

また、両 L3 グループの無声破裂音の正聴率と実験刺激の VOT 値との相関性を調べた。Pearson Correlation の結果：日本語 $r=.41$ 、スペイン語 $r=.20$ 。L3 日本語学習者の正聴率と実験刺激の音響特徴 VOT 値の間に顕著な相関性が存在している。VOT 値が 0ms に近づけば近づくほど、正聴率も次第に下がることがわかった。VOT 数値は L3 日本語学習者が有声性を判断する音響的キューであることが言える。一方、L3 スペイン語学習者の場合は、正聴率と実験刺激の音響特徴 VOT 値の間に有意な相関関係が見られなかった。したがって、VOT 数値は L3 スペイン語学習者が有声・無声を判断する音響的キューではないことがいえる。

Bohn and Flege (1993)の研究では、スペイン語話者の英語習得早期段階においては、語頭

破裂音の正聴率はその学習者の L1 と L2 破裂音素の対応性に関係していることが報告されている。学習者の L1 無声有気破裂音の VOT 値区間は 92.5ms から 102ms(表 1)で、L2 無声有気破裂音の VOT 値区間は 65ms から 95ms(本稿 4.1)、両方とも 30ms を遥かに上回る。L3 日本語無声破裂音の VOT 値とその区間が大きく、学習者の L1 と L2 の無声有気破裂音 VOT 区間に近い。そのため、学習者は L3 日本語の無声破裂音の帯気性特徴により気づきやすく、有声・無声破裂音を区別して知覚し、正聴率が高い。

その一方、L3 日本語と比べると、L3 スペイン語の無声破裂音の正聴率が低い。結果分析からわかるように、学習者の L1 無声無気破裂音の VOT 区間は 6ms から 14.5ms(表 1)、L2 無声無気破裂音の VOT 区間は 10ms から 25ms(本稿 4.1)である。L3 スペイン語の無声破裂音 VOT 区間は、学習者の L1 と L2 の無声無気破裂音の VOT 区間と似ていることがわかる。本研究の結果からは、L1 と L2 の無声無気破裂音 VOT 区間が L3 無声破裂音区間と類似している場合、知覚混同が生じやすく、習得が遅れることが明らかになった。

4.2.2. L3 語頭有声破裂音

統計分析の結果、(1)L3 日本語有声破裂音の正聴率は無声破裂音のより高い、(2)L3 スペイン語の有声・無声破裂音の間に正聴率の差異が存在しない。知覚実験に使用されている母語話者音声刺激のデータを音響的に分析したところ、両 L3 母語話者の有声破裂音 VOT 値はともにマイナスであることがわかった。その平均値は、日本語 -74.42ms、スペイン語 -88.25ms。しかし、学習者の L1 と L2 語頭無気破裂音の VOT 値はプラスであり、L3 有破裂音の音響特徴とは大きく異なっている。それにもかかわらず、中国人学習者が L3 有声破裂音を知覚する際に、無声破裂音より高いあるいは同等な正聴率を見せている。したがって、中国人学習者にとって、L3 破裂音体系における有声破裂音の知覚は無声破裂音より有利である。早期 SLA 研究理論(Lado 1957, Eckman 1977, 1991)に対して、「equivalence classification」(Flege 1987)と SLM(Flege 1995)は、母語と類似度の高い目標言語音声項目の習得が難しいのは、学習者はそれらの音声項目を母語と同じ類に分類したからであると仮説を立ててきた。逆に新しいまたは類似度の低い目標言語音声項目の習得が容易なのは、学習者が母語と目標言語の相違性を早期に発見・区別できるからなのである。

次に、学習者の L3 有声破裂音の正聴率と知覚実験の音声刺激 VOT 値との相関関係を調べた。統計分析の結果、日本語 $r=.02$ 、スペイン語 $r=.01$ 。結果から、学習者の L3 有声破裂音の正聴率と実験刺激の VOT 値の間に相関性が存在しないことがわかった。VOT の数値は中国人学習者の L3 有声破裂音の有効な知覚キューではない。VOT 値が連続変量であることが知られている。また、人類は範疇知覚メカニズムの制約(Liberman, Harris and Hoffman 1957)によって、繊細な音響的特徴の弁別よりも意味弁別機能を働く音素の弁別能力の方が遥かに高い。仮に学習者が L3 と L1, L2 の相違性を察知し、三者の区別ができるとすれば、L3 音声知覚基準の確立には有利であろう。逆に、学習者が類似度の高い音声項目の知覚弁別ができなければ、言語間の相違性についての更なる検証・再現する操作も怠け、知覚混同

に陥ってしまう(Kingston 2003)。本研究では TLA における L3 有声・無声破裂音の知覚実験を行い、多言語学習者は母語と既習言語と差異の大きい L3 有声破裂音の prevoicing 特徴を知覚し、新たな L3 音声体系を築き上げていることがわかった。この発見は Flege's (1995) Speech Learning Model (SLM)とも一致する部分がある。この習得モデルでは、人間は生涯を通して、言語音声体系の適応性を保持し、常に修正・確立を繰り返していると指摘している。このような SLA 理論仮説は L2 習得研究分野のみならず、中国人学習者の L3 音声習得にも同じく適用することが本研究の結果から明白である。

5. まとめ

中国は前代未聞の複雑な多言語教育環境にある。本研究は L1 中国共通語、L2 英語、L3 日本語あるいはスペイン語の多言語学習者を対象に、L3 語頭破裂音の知覚実験を実施し、L3 の知覚習得状況および L1 と L2 の破裂音体系が L3 音声知覚に与える影響を分析した。まず、L3 無声破裂音の知覚は実験刺激の VOT 値とその区間に関係していて、L3 有声破裂音の知覚は prevoicing 特徴の有無に関係していることが明らかになった。また、L3 無声破裂音の知覚混同は L3 と L1、L2 の音素体系の類似性に起因していることがわかった。一方、L3 と L1、L2 有声破裂音の顕著な相違性が L3 有声破裂音の知覚に有利に働くことがわかった。研究結果から証明されたように、L3 破裂音の知覚習得困難を引き起こしたのは、L3 と既知言語との相違性ではなく、類似性が主な影響要因なのである。本研究の結果は言語間の相違性と類似性をめぐる言語習得理論問題の有力な実証データを提供するとともに、複雑な多言語音声教育改善の科学的な裏付けとなるだろう。

謝辞 この実験に協力してくれた学習者に感謝する。本研究は中国国家社会科学基金[代表者：劉佳琦、課題番号 18BYY227]による助成を受けている。

参考文献

- Flege J. E. (1995) "Second language speech learning: Theory, findings and problems." In W. Strange (Eds.), *Speech perception and linguistic experience: Issues in cross-language research*. (pp. 233-277). Baltimore: York Press.
- Lisker, L. and Abramson, A. S. (1964) "A cross-language study of voicing in initial stops: acoustical measurements." *Word* 20. (pp. 384-422).
- Liu, Jiaqi, Zeng, Ting and Lu, Xiuchuan (to appear) "Challenges in Multi-language Pronunciation Teaching: A Cross-linguistic Study of Chinese Students' Perception of Voiced and Voiceless Stops", *Circle of applied linguistics for communication*.
- Williams, S. and Hammarberg, B. (1998) "Language switches in L3 production: Implications for a polyglot speaking model." *Applied Linguistics* 19. (pp. 295-333).

ベトナム人日本語学習者のヤ行・ザ行・ジャ行音の知覚

佐藤 桐子 (首都大学東京大学院人文科学研究科)
sato-kiriko@ed.tmu.ac.jp

1. はじめに

近年日本国内外でベトナム人日本語学習者 (以下, ベトナム人学習者) が急増している。これまでのベトナム人学習者の日本語音声の知覚と産出に関する研究で法貴 (2000) や重川・中村 (2005) 等いくつかの研究結果においてベトナム人学習者がヤ行・ザ行・ジャ行の混同をすることが報告されている。日本語語彙の中でザ行・ジャ行音を含む語彙が多いため, これらの音の誤用はコミュニケーションに支障をきたす可能性が高いと考えられる。そこで本発表ではベトナム人学習者が日本語のヤ行・ザ行・ジャ行音をどのように知覚しているかを明らかにするためベトナム語北部方言話者 (以下, 北部方言話者) に対して子音の語内位置とアクセント型を統制した知覚実験を実施した。

2. 先行研究

2.1. ベトナム人学習者の日本語音声に関する研究

法貴 (2000) は日本国内の大学で日本語を学ぶベトナム人学習者 7 名に対して, 学習開始直後に聴取・発話・リピートの 3 つのテストを行った結果, ジャ行音とザ行音を混同すること, 聴取テストでシャ行音・チャ行音・ピヤ行音の誤答率が高く, 発話テストではピヤ行音およびシ・ジが難しかったと述べている。重川・中村 (2005) はベトナム人研修生 6 名に対して約 3 か月間の日本語発音指導を行い, ベトナム人学習者の発音の問題点として「つ」が「ちゅ」, 「ざ」が「じゃ」, 「ぞ」が「じょ」になること, 「や・ゆ・よ」が「じゃ・じゅ・じょ」になること等を報告している。法貴 (2000)・重川・中村 (2005) とともに被験者の人数が少なく被験者の母方言も考慮されていない点から, より詳細な研究が必要であることがわかる。ファム (2006) はベトナムの大学で北部方言話者と南部方言話者のベトナム人学習者 264 名に知覚実験を行った結果, いずれの母音でも調音点が隣接している音声の聞き分けが難しいこと, ザ行が語頭にある場合ジャ行との聞き間違いが多くなること, 母音[u] が後続する場合比較的聞き分けやすいこと, アクセント型の影響はないことを明らかにしている。しかし, 被験者の日本語レベル・日本語学習歴が不明であること, 一部の考察では行ごとの分析しか行っていないこと等から十分に研究が行われたとは言い難い。

2.2. ベトナム語の方言区分

ベトナム語の方言について富田 (1988) は北部方言はタイン・ホア省あたりまで, 南部方言はクアン・ナム省とダナン省以南, そして中部方言は北部方言と南部方言の中間地域であるとしている。なお, 北部方言はベトナムの首都ハノイの周辺地域で話されていることからベトナムの共通語とされている。本発表ではベトナムの方言区分についてその研究の詳細さから富田 (1988) の区分を援用して被験者の母方言を分類することとした。

3. 知覚実験

3.1. 調査語（刺激音声）

本実験に用いた調査語はヤ・ザ・ジャ行子音に母音ア・ウ・オをつけた9語にマを組み合わせた18の無意味語（ヤマ・マヤ・ユマ・マユ・ヨマ・マヨ・ザマ・マザ・ズマ・マズ・ゾマ・マゾ・ジャマ・マジヤ・ジュマ・マジユ・ジョマ・マジョ）である。この18語を頭高型（以下、HL）と平板型（以下、LH）の2種のアクセントで日本語母語話者（声優・男性）が読み上げたものをSONY製PCMレコーダー（PCM-10, サンプル周波数44.1kHz, 量子化16bit）およびマイクロフォン（ECM-PCV40）で録音し刺激音声とした。録音した刺激音声は発表者と他1名の日本語母語話者が日本語として自然であることを確認した。

3.2. 被験者と調査手順

実験は東京都内の日本語教育機関で2018年5月から6月にかけて実施した。被験者は初級クラスに在籍する北部方言話者19名である。被験者の年齢は平均21.62歳（ $SD=3.975$ ）、日本語学習歴は平均9.95か月（ $SD=1.870$ ）、滞日期間は平均3.47か月（ $SD=1.744$ ）である。

調査語18語とダミー6語を合わせた23語をランダムに並べ替え1セットとしたものを3試行実施し、被験者に強制3択式で答えさせた。各セットともHLとLH別々に実施したため実験は全36セット行った。解答の選択肢は、例えば調査語「ヤマ」に対して「ザマ・ヤマ・ジャマ」という形式に設定した。

3.3. 結果と考察

被験者の回答をMS Excelのシートの入力し、正答・誤答をCOUNTIF関数を用いて集計した。また、統計分析の際はIBM SPSS Statistics 25を用いた。各調査語のデータ数は被験者19名×3試行で57である。はじめに、各調査語の正答率を図1に示す。

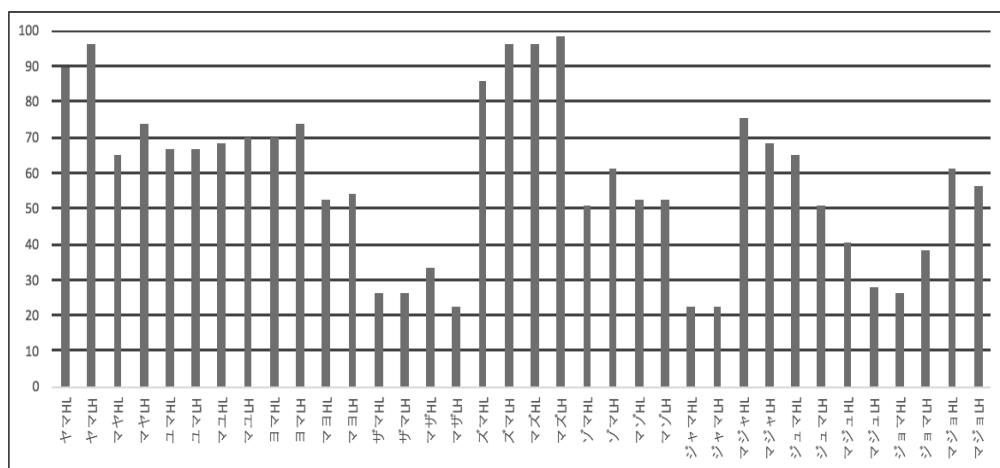


図1: 調査語ごとの正答率

最も正答率が高かったのはマズ LH (98.2%)、次いでマズ HL・ズマ LH・ヤマ LH (96.5%)

であり、ズを含む語の正答率が比較的高くなっている。一方、正答率が低かったのはマザ LH・ジャマ HL・ジャマ LH (22.8%)、次いでザマ HL・ザマ LH・ジョマ HL (26.3%) で、ザとジャを含む語の正答率が比較的低くなっている。

はじめに、子音の語内位置(語頭・語中)によって正答に差があるかを対応のある t 検定で分析したところ、ヤマ HL とマヤ HL, ヤマ LH とマヤ LH, ズマ HL とマズ HL, ジャマ HL とマジャ HL, ジャマ LH とマジャ LH, ジュマ HL とマジュ HL, およびジョマ HL とマジョ HL との間に有意差がみられた。有意差がみられた調査語の t 検定の結果を表 1 に示す。

表 1: 語内位置(語頭・語中)についての t 検定の結果一覧

調査語	M	SD	t	df
ヤマ HL-マヤ HL	.737	1.147	2.800*	18
ヤマ LH-マヤ LH	.684	1.108	2.691*	18
ズマ HL-マズ HL	-.316	.478	-2.882*	18
ジャマ HL-マジャ HL	-1.579	1.305	-5.276**	18
ジャマ LH-マジャ LH	-1.368	1.461	-4.083**	18
ジュマ HL-マジュ HL	.737	1.046	3.071**	18
ジョマ HL-マジョ HL	-1.053	1.353	-3.391**	18

**: $p<0.01$, * $p<0.05$

t 検定の結果から、ヤマ HL とマヤ HL, ヤマ LH とマヤ LH, ジュマ HL とマジュ HL では語頭の調査語のほうが有意に正答が多かったことから、ヤについては語頭のほうが知覚しやすい傾向があると考えられる。またジュについては語頭のほうがやや知覚しやすく、語中の場合やや知覚しにくいと考えられる。一方、ズマ HL とマズ HL, ジャマ HL とマジャ HL, ジャマ LH とマジャ LH, ジョマ HL とマジョ HL では語中のほうが有意に正答が多くなっている。このため語中のズは知覚しやすく語頭のズは知覚が難しいと考えられる。なお、ズマ LH の正答率は 96.5%, マズ LH は 98.2%とズは全体的に知覚が容易であるものの、語頭の場合のみ若干正答率が下がる傾向があるようである。次に、語頭のジャはかなり知覚しにくい語中の場合は比較的知覚が容易であることがわかった。ジョについても語頭のほうが知覚しにくく語中はやや知覚が容易であるという傾向があるようである。

次に、調査語ごとにアクセント型で正答に差があるかを対応のある t 検定を用いて分析した結果ジョマ HL (正答率 26.3%) とジョマ LH (正答率 38.6%) でジョマ LH のほうが有意に正答数が多いことがわかった ($t=-.002, df=18, p<.05$)。このことから、ジョに関しては LH のほうが知覚が容易であると考えられる。なお、他の調査語ではアクセント型による差は見られなかった。

ヤ・ザ・ジャの各行で母音によって正答に差があるかについてヤ・ザ・ジャの行ごとに語内位置（語頭・語中）とアクセント型（HL・LH）別で正答数を従属変数，母音ア・ウ・オを因子として分散分析を行った．その結果，ヤ行語頭 HL ($F(2,54)=3.675, p<.05$)，ヤ行語頭 LH ($F(2,54)=4.249, p<.05$)，ザ行語 HL ($F(2,54)=25.515, p<.01$)，ザ行語頭 LH ($F(2,54)=25.000, p<.01$)，ザ行語中 HL ($F(2,54)=22.801, p<.01$)，ザ行語中 LH ($F(2,54)=26.828, p<.01$)，ジャ行語頭 HL ($F(2,54)=11.400, p<.01$)，ジャ行語頭 LH ($F(2,54)=3.187, p<.05$)，ジャ行語中 HL ($F(2,54)=5.202, p<.01$) およびジャ行語中 LH ($F(2,54)=6.024, p<.01$) に有意差がみられたため，Tukey HSD 法で多重比較を実施した．多重比較の結果有意差があった組み合わせを表 2 に示す．また，各調査語の正用の平均値を語内位置・アクセント型ごとにまとめたものを図 2 に示す．

表 2: 多重比較の結果一覧

	I(母音)	J(母音)	I-J	SD		I(母音)	J(母音)	I-J	SD	
ヤ行語頭 HL	あ	う	.684*	.272	ザ行語中 LH	あ	う	-2.263**	.311	
	う	あ	-.684*	.272			お	-.895*	.311	
ヤ行語頭 LH	あ	う	.895*	.321		う	あ	2.263**	.311	
	う	あ	-.895*	.321			お	1.368**	.311	
ザ行語頭 HL	あ	う	-1.789**	.252		お	あ	.895*	.311	
		お	-.737*	.252			う	-1.368**	.311	
	う	あ	1.789**	.252		ジャ行語頭 HL	あ	う	-1.263**	.294
		お	1.053**	.252			う	あ	1.263**	.294
	お	あ	.737*	.252	お			1.158**	.294	
		う	-1.043**	.252	お		う	-1.158**	.294	
ザ行語頭 LH	あ	う	-2.105**	.298	ジャ行語頭 LH	あ	う	-.842*	.334	
		お	-1.053**	.298		う	あ	.842*	.334	
	う	あ	2.105**	.298	ジャ行語中 HL	あ	う	1.053**	.329	
		お	1.053**	.298		う	あ	-1.053**	.329	
	お	あ	1.053**	.298	ジャ行語中 LH	あ	う	1.211**	.358	
		う	-1.053**	.298		う	あ	-1.211**	.358	
ザ行語中 HL	あ	う	-1.895**	.288						
	う	あ	1.895**	.288						
		お	1.316**	.288						
	お	う	-1.316**	.288						

**: $p<0.01$, *: $p<0.05$

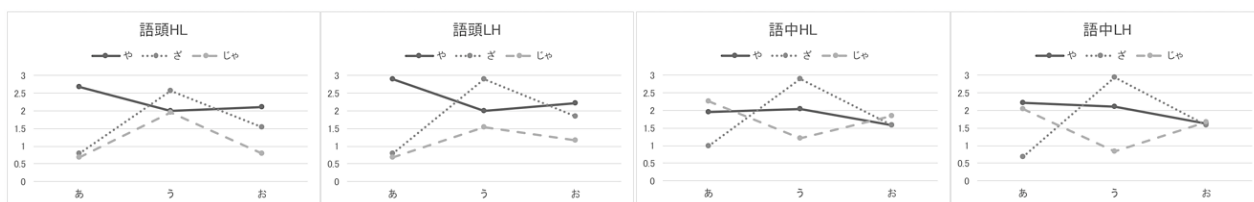


図 2: 語内位置・アクセントごとの平均値の比較

表 3 および図 2 からヤ行において語頭の場合アクセントを問わず母音アとウではウが後続する場合正しく知覚されやすいことがわかった。ザ行においては、まず語内位置・アクセント型を問わず母音ウが後続する場合他の母音より知覚しやすいことがわかった。次に、語頭 HL・語頭 LH・語中 LH で後続する母音がアとオの場合オのほうが知覚されやすいこと、ウとオの場合ウのほうが知覚されやすいことがわかった。ジャ行の場合、語頭 HL ではア・オが後続する場合よりウが後続する場合のほうが知覚されやすいこと、語頭 LH ではア後続する場合よりウが後続する場合のほうが知覚されやすいこと、そして語中 HL と LH ではウが後続する場合よりアが後続する場合のほうが知覚されやすいことがわかった。

最後に正答率が 30%以下だった調査語 8 つについて誤用の内訳を分析した (表 3)。

表 3: 誤用パターンの分類

調査語	正用数(%)	誤用数(%)	調査語	正用数(%)	誤用数(%)
ジャマ HL	ジャ 13 (22.8)	→ザ 16 (28.1) →ヤ 28 (49.1)	ジャマ LH	ジャ 13 (22.8)	→ザ 11 (19.3) →ヤ 33 (57.9)
マザ LH	ザ 13 (22.8)	→ヤ 2 (3.51) →ジャ 42 (73.7)	ザマ HL	ザ 15 (26.3)	→ヤ 28 (49.1) →ジャ 14 (24.6)
ザマ LH	ザ 15 (26.3)	→ヤ 20 (35.1) →ジャ 22 (38.6)	ジョマ HL	ジョ 15 (26.3)	→ゾ 25 (43.9) →ヨ 17 (29.8)
マジュ LH	ジュ 16 (28.1)	→ズ 32 (56.1) →ユ 9 (15.8)			

表 3 からジャマ HL とジャマ LH ではザと聞き間違えるよりもヤと聞き間違える誤用のほうが多いことがわかった。一方マジュ LH とジョマ HL ではユ/ヨよりもズ/ゾと聞き間違える誤用のほうが多くなっている。マザ LH とザマ LH ではジャと聞き間違える誤用のほうがヤと聞き間違えるよりも多いが、マザ LH では 73.7%がジャと聞き間違えているのに対してザマ LH ではヤ・ジャの違いはほとんどない。一方ザマ HL ではジャと聞き間違える誤用よりもヤと聞き間違える誤用のほうが多くなっている。

この結果から、ジャ行では母音アが後続する場合ヤ行との聞き間違えが、ウ・オが後続する場合ザ行との聞き間違えが起りやすいと推測できる。

4. 結論と今後の課題

ベトナム人学習者のヤ行・ザ行・ジャ行の知覚について明らかにするため知覚実験を行った。結果についてはじめに、子音の語内位置について対応のある *t* 検定で分析したところヤとジュの場合語頭のほうが知覚しやすく、ズとジャ・ジョは語中の方が知覚しやすいことがわかった。

次に、アクセント型が知覚に影響するかについて対応のある *t* 検定を行った結果、ジョのみ LH のほうが HL より知覚がしやすいことがわかった。

後続母音の影響を分析したところ、ヤ行が語頭の場合アクセントを問わず母音ウが後続するとき正しく知覚されやすいことがわかった。ザ行においては、まず語内位置・アクセント型を問わず母音ウが後続する場合他の母音より知覚しやすいことがわかった。また語頭 HL・語頭 LH・語中 LH で後続する母音アとオの場合オのほうが知覚されやすいこと、ウとオの場合ウのほうが知覚されやすいことがわかった。ジャ行の場合、語頭 HL ではウが後続する場合知覚されやすいこと、語頭 LH ではア後続する場合よりウが後続する場合のほうが知覚されやすいこと、そして語中 HL と LH ではウが後続する場合よりアが後続する場合のほうが知覚されやすいことがわかった。

最後に正答率が 30% だった調査語についてその誤用の内訳を分析したところ、ジャマ HL とジャマ LH ではヤと聞き間違える誤用が多いこと、マジュ LH とジョマ HL ではズ/ゾと聞き間違える誤用が多いこと、マザ LH とザマ LH ではジャと聞き間違える誤用が多いこと、そしてザマ HL ではヤと聞き間違える誤用が多いことがわかった。よってジャ行では母音アが後続する場合ヤ行との聞き間違えが、ウ・オが後続する場合ザ行との聞き間違えが起りやすいことが推測できる。

今後の課題としてまず特殊拍の有無や先行母音等他の音環境での実験が必要であることと、中部方言話者・南部方言話者に対する実験を行い北部方言話者と比較する必要があることがあげられる。

参考文献

- ファム, T.H. (2006) 「ベトナム語母語話者による日本語のザ行音・ジャ行音・ヤ行音の聞き分け」『日本言語文化研究会論集』 2, 83-108.
- 法貴則子 (2000) 「ベトナム人、カンボジア人、ラオス人学習者の音声上の問題点」『東京外国語大学留学生日本語教育センター論集』 26, 183-197.
- 重川明美・中村倫子 (2005) 「ベトナム人学習者への発音指導の一例」『日本語教育方法研究会紙』 12:2, 22-23.
- 富田健次 (1988) 「ヴェトナム語」『言語学大辞典第 1 巻世界言語編 (上)』 759-787, 三省堂.

中国人日本語学習者による促音の知覚

—中国語普通話話者を対象に—

任 宏昊・劉 奕櫟 (早稲田大学大学院) ・近藤 真理子 (早稲田大学)
 {haoer21262,liuyili}@{fujii,toki}.waseda.jp, mkondo@waseda.jp

1. はじめに

日本語の音声習得において、特殊拍の一つである促音/Q/は、中国語（普通話）の音韻体系には存在しておらず、中国人日本語学習者が促音を習得する際にとりわけ問題となっている。学習者を対象とする多くの先行研究においては、語中子音の閉鎖・摩擦持続時間が促音・非促音を聞き分ける主な手がかりであると言われている。しかし、促音部の持続時間は発話速度や個人の発話習慣の違いによって左右されるため、学習者による促音判断に揺れが生じ、不安定な傾向を示す原因となる。

また、促音部の持続時間は促音判断における唯一の手がかりではないという指摘もある（福井 1978）。アクセント型や先行・後続母音の時間長といった促音の存在する環境を考慮しながらその知覚を考察する研究は見られるが、促音を形成する語中子音による知覚への影響に関する討論は管見の限り十分なされていない。

そこで本稿は、中国人日本語学習者（普通話話者）の促音の知覚実態（正聴率・範疇知覚など）を明らかにし、それに影響する絶対的な持続時間以外の手がかりを探り出していくことを目的とする。

2. 先行研究

2.1. 促音の知覚における時間的な手がかり

従来、日本語の促音の知覚において、時間長は主な手がかりとされてきた。時間的な要素としては、促音部の持続時間（closure duration）とその隣接母音の長さがよく取り上げられ、促音知覚にどのような影響を与えているかしばしば論述されてきた。子音持続時間の促音の知覚に対する役割について、Fujisaki et al. (1973) は/ise/-/isse/、/ita/-/itta/という促音・非促音対立を含む語を基にし、子音持続時間を伸縮させた加工音声で知覚実験を行ったところ、絶対的な持続時間長が促音の知覚に示差的な役割を果たしていることを示した。

また、隣接する母音の長さが促音の知覚に関わっているとする先行研究も複数見られる。渡部・平藤（1985）は、異なる速度で発話された/apa/ /ata/ /aka/について語中無声区間の時間長を伸縮させることによって、知覚同定実験を行った。その結果、被験者による促音の判断境界値が先行母音の長さとは一次関数的に関連しており、促音に先行する母音長が促音の判断閾値に影響を与えているとした。一方、大深・森・桐谷（2005）は、実在語の/uta(Q)tane/（転寝/歌ったね）を刺激とし、促音部に隣接する母音の長さを変化させ、「次は、____、らしいです。」というキャリア文内で提示することで、21名の日本語母語話者による促音の

判断境界値を求めた。同研究は「先行母音が長く、後続母音が短い (LS) パターンが最も促音に聞こえやすく、後続母音が長いパターン (LL と SL) が促音に聞こえにくい」という傾向から、先行母音と後続母音の長さが共に促音の知覚に有意な影響を及ぼすとの結果を得た。

促音知覚に関わる非時間的な手がかりについても、いくつかの研究が行われている。大深 (2003) は、アクセントパターンの異なる促音・非促音語「カタ」「カッタ」(頭高型:肩/勝った、平板型:型/買った) における /t/ の持続時間を変化させ、それらの語をキャリア文に埋め込んで被験者に提示し、知覚実験を行った。その結果、促音知覚判断の境界値は、アクセントパターンが平板型である場合、少し大きくなる傾向にあった。すなわち、平板型の刺激を促音語として知覚するには、より長い閉鎖持続時間が必要とされる。

一方、柳澤・荒川 (2015) は、合成音声を用いて、「フォルマント遷移の有無」と「インテンシティの減衰」による促音知覚への影響について調査した。その結果、フォルマント遷移の有無が促音知覚に明確な影響を及ぼすことが明らかになったが、インテンシティの減衰による促音知覚への影響は、フォルマント遷移がない場合にしか見られなかった。

以上の研究結果を踏まえて、本稿では、相対的持続である「単語時間長 (EWL) に対する促音部持続時間長 (CDL) の比率」(以下 CDL/EWL と略す) を主な変数とし、促音知覚判断の手がかりとしての可能性を検証する。更に、先行・後続母音の聞こえ度、語中子音の調音方法と学習者日本語レベルといった要素が促音の知覚に影響を与えるかどうかを考察する。

3. 実験方法

3.1. 刺激

本研究は先行・後続母音の聞こえ度、語中子音の調音方法の違いによる中国人学習者の促音の知覚の差異を捉えるため、刺激語における先行・後続母音を聞こえ度の異なる /a/ /e/ /i/ とし、語中子音を破裂音の /k/ /p/ と摩擦音の /s/ とする。本実験で使う刺激語を表 1 に挙げる。

刺激語の一部はアクセントにより有意味語 (実在語) となるが、意味的要素に干渉されないように、被験者には全ての刺激を無意味語として扱ってもらい、刺激語の意味だけで促音の有無を判断しないように注意を呼びかけておいた。

表 1: 本研究で使用了刺激リスト

	両唇音		歯茎音	
	非促音	促音	非促音	促音
破裂音	<i>apa</i>	<i>aQpa</i>	<i>aka</i>	<i>aQka</i>
	<i>epe</i>	<i>eQpe</i>	<i>eke</i>	<i>eQke</i>
	<i>ipi</i>	<i>iQpi</i>	<i>iki</i>	<i>iQki</i>
摩擦音			<i>asa</i>	<i>aQsa</i>
			<i>ese</i>	<i>eQse</i>
			<i>isi</i>	<i>iQsi</i>

3.2. 刺激の作成

3.2.1. 録音

刺激音を作成するにあたって、まず日本語教育または音声学に携わる 3 名の成人東京方言話者(男性 A, 女性 B・C)に音声を録音してもらった。録音は、各刺激を「それは、____、です。」というキャリア文に入れて、発話者それぞれの最も自然な発話状態で 6 回ずつ読み上げてもらったものを、早稲田大学の防音室にて収録した(サンプル周波数 44.1kHz、量子化 16 bit)。本研究ではアクセントパターンによる促音知覚への影響を考慮しないため、全ての刺激のアクセント型を頭高型とした。

3.2.2. 刺激語の選定と作成

各実験の目的に合わせて、以下の 2 種類の刺激音を作成した。波形の観察には *Praat* (version 6.0.34) を使用し、刺激語全体の標準化および長さの加工等は *Adobe Audition CC* (version 10.1.1.11) を用いて行った。

【実験 1】協力者 A と B が発音した音声データの中から、音質がよく、フォルマントがはっきりとしているものを 1 刺激につき 3 つ選び、刺激語の部分のみを切り出した。抽出した音声全体を 65dB に標準化した。

【実験 2】協力者 C が発音した音声データのミニマルペア (/aka-/akka/, /isi-/issi/など)の中から、加工に適したもの(1 ペアにつき 1 語)計 9 語を選出した。更に、この 9 つの刺激語の先行・後続母音を 100/150 msec¹にし、CDL/EWL を 15%~60%²の間で 5%の間隔で伸縮した。得られた計 90 個の刺激語の全体を 65dB に標準化した。

3.3. 被験者及び実験方法

被験者は 10~20 代の中国語普通話話者 44 名である。被験者は全員中国の大学で日本語を主専攻とする大学生又は大学院生であり、聴力が正常であると自己申告した。被験者は日本語能力試験の得点、学習時間数(月)及び指導教官の評価等に基づき、初級学習者群(男性 4、女性 20、計 24 名)と上級学習者群(男性 4、女性 16、計 20 名)に分けられた。

作成した刺激音を実験ごとに(実験 1: 計 108 [18 語×3 回×2 話者] 刺激; 実験 2: 計 180 [9 語×10 段階×2 回] 刺激) *Praat* 上でランダムに被験者に提示した。実験開始前、被験者に知覚実験に関する説明を行い、実験参加への同意を書面で得た。

実験は雑音のない静かな部屋で被験者ごとに行われ、各被験者にヘッドセットを通して刺激語を聞いてもらった。実験 1 と 2 の間には 10 分程度の休憩を設け、各実験の進行度は被験者に任せた。所要時間は、同意書などの書類記入や説明の時間を含めて約 30 分であった。

¹ 加工音声の先行・後続母音長は録音した音声データにおける全ての単語の先行・後続母音長の平均値(102/153)を、十の位まで四捨五入した数値とした。

² 予備実験の結果によると、刺激語における CDL/EWL ≤ 15% の場合は、被験者全員による直音(非促音)の知覚率が 100% であり、CDL/EWL ≥ 60% の場合、被験者全員が「促音あり」と知覚した。このため、本研究で扱う加工音声の CDL/EWL を 15%~60% という範囲で伸縮することにした。

4. 実験結果と考察

4.1. 実験 1

まず、中国人学習者全体による促音・非促音語の平均正聴率はそれぞれ 73.86%と 89.39%であり、大きな差が見られた。独立サンプル t 検定を行ったところ、両者の差が有意 [$t=10.257$, $df=710.65$, $p < .01$] であり、中国人普通話話者による促音を含む語の知覚正聴率が顕著に低く、知覚が困難であることを示している。

次に、学習者の日本語力水準（初級と上級）と刺激語のモーラ数（2-mora：促音を含まない、3-mora：促音を含む）とを主要因にし、被験者全員の平均正聴率に対して、2元配置の反復測定分散分析を行った。その結果、学習者の日本語力水準の主効果 [$F(1,788) = 16.79$, $p < .01$]、刺激語のモーラ数の主効果 [$F(1,788) = 99.77$, $p < .01$] と学習者の日本語力水準と刺激語のモーラ数の交互作用 [$F(1,788) = 70.11$, $p < .01$] はいずれも0.01水準で有意であった。したがって、学習者の日本語レベルも促音・非促音語の知覚正聴率に影響を及ぼし、学習者の日本語能力が上達するにつれて、促音・非促音語の正聴率が高くなるという傾向が見られた。

下図1は異なる母音条件の下での各被験者グループによる促音・非促音語の平均正聴率の結果である。語中母音の聞こえ度が促音・非促音語の正聴率と関係があるかどうかを確認するため、語中母音種（/a/, /e/, /i/）を主要因として、初・上級学習者それぞれの促音・非促音語の平均正聴率に対し、1元配置の分散分析を行った。いずれの場合でも0.05水準で有意差は認められなかったことから、語中母音の聞こえ度は促音・非促音語の正聴率に影響がないと示唆された。

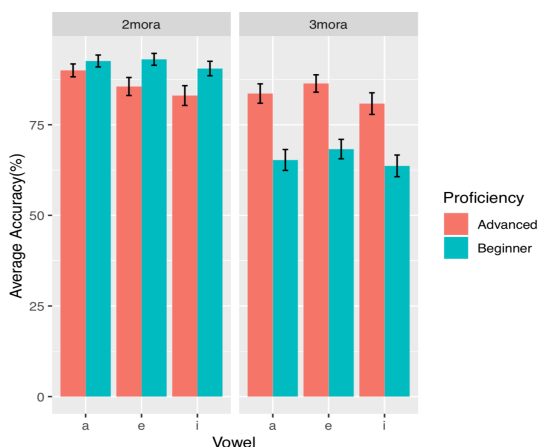


図 1: 異なる母音条件における促音・非促音の平均正聴率

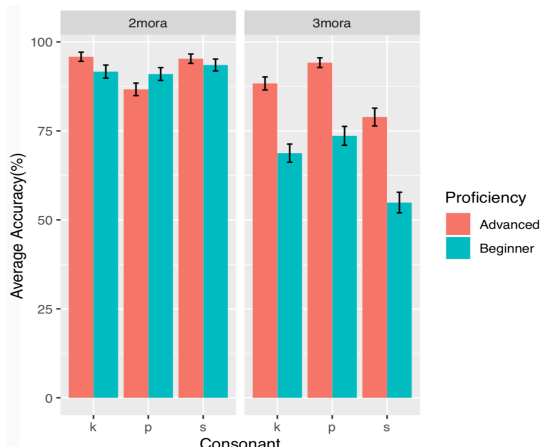


図 2: 異なる子音条件における促音・非促音の平均正聴率

上図2は三つの語中子音状況における促音・非促音の平均正聴率を示している。初・上級学習者による促音・非促音語の知覚正聴率を従属変数とし、語中子音（/k/, /p/, /s/）を因子とした1元配置の分散分析を行った結果、初級学習者による非促音語の知覚 [$F(2,213) = 0.55$, $p > .05$, $n.s.$] を除いて、全ての場合で語中子音の主効果は0.01水準で有意であった [初級: $F_{促音}(2,213) = 12.93$, $p < .01$; 上級: $F_{非}(2,177) = 12.16$, $p < .01$; $F_{促音}(2,177) =$

15.566, $p < .01$]. 次に、有意差が出た事例に対して、事後検定 (post-hoc test) として、テューキーのHSD検定を行った。その結果、非促音の正聴率については、語中子音が/k/と/s/である場合より、/p/の方が有意に低い ($p < .01$) が、/k/と/s/の間には有意な差がなかった ($p = .96, n.s.$)。そして、促音の正聴率の場合、初級と上級グループとも/s/ < /p/ ($p < .01$, /s/ < /k/ ($p < .01$) で有意であったが、/k/ < /p/ ($p_{初} = .41, n.s.$; $p_{上} = .09, n.s.$) は0.05水準で有意差が認められなかった。

4.2. 実験2

初級学習者と上級学習者による2モーラ・3モーラ語の平均知覚率と、それを基にロジスティック回帰分析による近似曲線を図3に示す。

図3を見ると、各グループの学習者による2系列 (2モーラ/3モーラ) の平均知覚曲線の傾きが急であると観察される。つまり、中国人普通話話者は明確な知覚判断境界を持って促音と非促音を聞き分けており、範疇的知覚を行っていると言えよう。

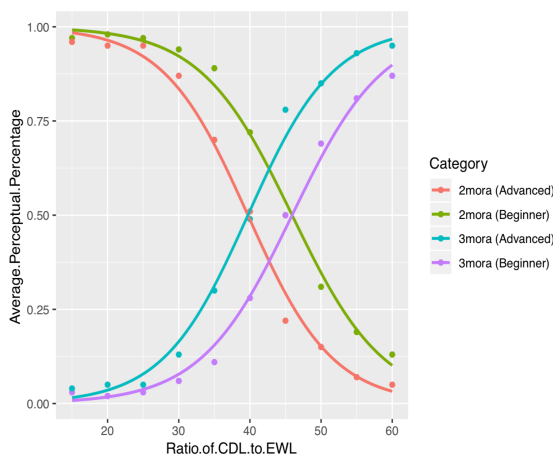


図 3: 各被験者群による2・3モーラ語の平均知覚率

各系列におけるロジスティック関数上で、平均知覚率が50%である点に対応するCDL/EWLの値をそれぞれの促音・非促音カテゴリーの知覚範疇境界閾値 (Boundary Point, 以下BPと略す) とした。図4は、学習者群ごとの平均知覚判断境界の閾値を表す。

図4にあるように、初級学習者による知覚判断境界値は上級学習者よりばらつきが大きく、より不安定な傾向を示していることが分かった。平均判断境界値については、初級学習者の場合が46.29%であり、上級学習者の場合39.66%であった。

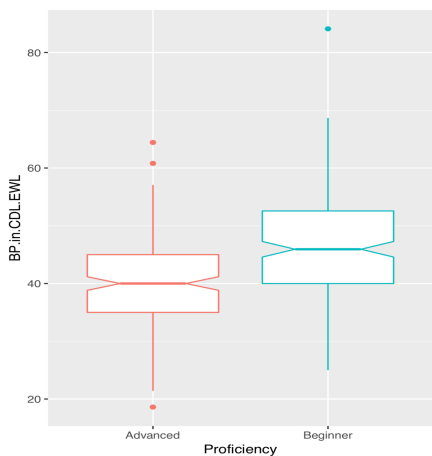


図 4: 各被験者群の平均知覚範疇境界値 (%)

次に、学習者の日本語能力水準を説明変数とし、判断境界値を目的変数として、独立サンプル t 検定を行ったところ、学習者群間に有意差が見られた [$t = 8.390, df = 394, p < .01$]. 即ち、促音・非促音を聞き分ける際には、中国人普通話話者の日本語能力が高くなるほど、より短い持続時間 (CDL/EWL の比率) で促音だと知覚できると言える。

さらに、各実験参加者群の刺激条件毎の知覚判断境界値を図5に示す。それぞれの判断境界値を従属変数とし、先行・後続母音種と語中子音種を独立変数として、2元配置の分散分析を行った。その結果、先行・後続母音の主効果 [$F(2, 387) = 0.97, p > .05, n.s.$] と交互作用 [$F(4, 387) = 1.06, p > .05, n.s.$] は有意でなく、語中子音種による主効果 [F

(2,387) = 22.19, $p < .01$] しか有意ではなかった。TukeyのHSD検定の結果から、両学習者群において、/s/ > /k/ ($p < .01$)、/s/ > /p/ ($p < .01$) は有意であったが、/k/と/p/の間の差異に有意差は認められなかった。つまり、語中子音が/s/である場合は、促音だと判断するにはより長い持続時間が必要とされ、促音・非促音の聞き分けが難しいことが明らかになった。

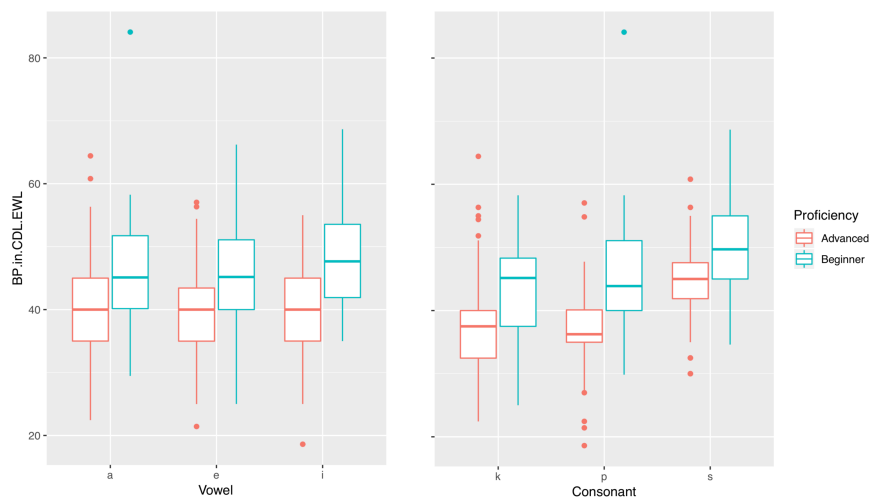


図 5: 中国人学習者による各刺激条件の知覚判断境界値(%)

5. 結論

今回の研究結果は、中国人学習者にとって日本語促音の知覚判断が問題であるという点で先行研究と一致している。更に、促音の知覚において、語中子音の絶対的持続時間という手がかりのほか、CDL/EWLという相対的時間比率をもって、より安定的な促音範疇的知覚を行っていることが明らかになった。非時間的な要因による影響に関しては、学習者の日本語能力水準、語中子音種の影響が確認された。先行・母音の聞こえ度による影響は一つのケースにしか見られなかったため、主要因として促音知覚に影響を及ぼすとは考えにくい。今後の課題としては本実験で扱わなかった/u/ /o/の二つの母音も取り入れて、母音の聞こえ度による促音知覚への影響について再考する必要があると考えられる。

6. 謝辞

本研究の音声刺激を録音して下さった日本語母語話者の方、また参加者をご紹介くださった崔健、趙文奇、趙迎結三氏、被験者の方々に感謝する。なお、本研究は中国国家留学基金の助成(201608210193)を受けている。

主な参考文献

- 大深悦子・森庸子・桐谷滋 (2005) 「促音の知覚に対する先行・後続母音長の影響」『音声研究』 9:2, 59-65.
- 川原繁人 (2013) 「日本語の特殊拍の音響—促音を中心として」『日本音響学会誌』 69:4, 191-196.
- 柳澤絵美・荒井隆行 (2015) 『フォルマント遷移とインテンシティの減衰が促音の知覚に与える影響』『日本音響学会誌』 71:10, 505-515.

中国北方方言話者による動詞活用形の東京アクセントの生成 —アクセントの習得実態及び F0 動態の考察—

王 鳳翔 (広島大学大学院)
ouousyou@gmail.com

1 導入

本稿では、筆者が生成調査を行った中国北方方言話者における 15 種類の動詞活用形のアクセントの習得実態及び F0 の特徴について論ずる。中国北方方言話者による東京語動詞のアクセントの生成について、王 (2017) では、正答率が最も低い動詞の打消し形(「ない」形)に焦点を当て、動詞の打消し形(「ない」形)と形容詞(「○○ない」例:きたない、せつない)の学習順序の関係、プロミネンス、両言語アクセントの相違、及び日本語教材におけるアクセントの解説という 4 つの側面から、日本語学習者の習得実態及び形成要因について報告した。この研究に引き続き、本稿では、残りの 15 種類の活用形に注目し、アクセントの習得実態及び F0 動態を考察する。

2 先行研究

2.1 日本語動詞活用形のアクセントの生成研究

動詞アクセントの生成研究に関しては、劉(2008)は中国語話者(北京・上海)を対象に、東京語の動詞及び複合動詞アクセントの生成調査を実施し、東京語母語話者に評価させた。学習者の音声データ及び日本語母語話者の評価に基づき、分析を行い、その上でさらに、アクセント規則の導入方法及び注意事項について検討した。生成調査の結果から、学習者は動詞辞書形のアクセントの習得が困難であること、また、動詞条件形、可能形、受身・尊敬形、使役形の正答率は、正答率の高いグループと正答率の低いグループという二つのグループに分けられること、さらに、動詞未然形(ない形)の正答率は全体的に低く、北京方言話者と上海方言話者の間に生成の誤用パターンの差異があることが示された。

2.2 中国語話者と日本語話者における日本語の音響的特徴の研究

許(2009:17)は中国の大学の日本語専攻に所属する中国北方方言話者を分析対象者とし、平板型、中高型、尾高型の 3, 4 拍名詞における相対的に低い 1 拍目のピッチ曲線の特徴を明らかにした。また、王(2003:43)によって、台湾上級学習者と東京語話者の語(名詞・動詞辞書形)アクセントのピッチ曲線の特徴が明らかにされた。他には、平野(2009)は、中国人日本語学習者と日本語母語話者の朗読音声の F0 パターンの形状を観察し、文中の文節ごとのレンジの推移を、正規化 F0 の最大値、中央値、最小値によって表現し、両者の違いを明らかにした。

以上の先行研究より、学習者と母語話者における日本語の音響的な特徴が分かった。しかし、従来の音響的な研究は、基本的に単語(名詞・動詞辞書形)あるいは文節、文を対象とする考察したものが多く、動詞活用形に焦点を当て、考察した研究は未だ行われていない。また、「テモ」形、過去形、命令形、可能形、未然形(～ず)、連用形、支配形、条件形 b(～たら)などの動詞活用形について、まだ検討されていない。そこで、本稿は、15 種類の動詞活用形の習得実態を明らかにし、その形成要因について解釈を試みることを、また、中国北方方言話者による F0 動態の一般的特徴を明らかにすることを目的とする。

3 生成調査

3.1 調査対象

本研究は、方言の影響を少なくするため、中国北方方言話者（上級日本語学習者）10名を対象とし、生成調査を実施した。

3.2 調査手順

生成調査：日本語学習者10名（1人につき180語）に調査用語を自然に読んでもらい、その発音を録音する。生成調査に要する時間は一人につき約8分である。

3.3 調査内容

アクセント式は平板式と起伏式に分類し、動詞類型による五段動詞と一段動詞という2種類に分類した。本稿で扱う動詞活用形は「テ」形、「テモ」形、過去形、命令形、可能形、未然形（～ず）、連用形、支配形、条件形a（～ば）、条件形b（～たら）、意志形、丁寧形、希望形、使役形、受身形の15種類である。

4 生成調査の結果

各動詞活用形の正答率によって、以下のような4つの型に分けられた。

I型：正答率が全体的に高い型（意志形（～よう、ろう）、丁寧形（～ます））

II型：正答率が全体的に低い型（未然形（～ず））

III型（偏り型-1）：動詞類型に関して、起伏式の正答率が高い、あるいは、平板式の正答率が高い。（「テ」、「テモ」形、過去形、命令形、条件形（～たら））

IV型（偏り型-2）：動詞類型に関わらず、起伏式の正答率が高い、もしくは、平板式の正答率が高い。（連用形、支配形、希望形、可能形、受身形、使役形、条件形（～ば））

5 習得実態の考察

生成結果によって、動詞類型と拍数に関わらず、平板式の正答率が起伏式より低く、学習者には平板式動詞活用形の習得が困難であること、また、多くの学習者が平板式動詞活用形のアクセントを起伏式のように発音してしまうことが明らかになった。特に、連用形、支配形、希望形、使役形、受身形、可能形において顕著に見られる。発音誤用例：「出な' がら→寝な' がら」、「出そ' うだ→寝そ' うだ」、「飲め' る→呼べ' る」。これは、日本語のアクセントと中国語のアクセントの表れ方が異なるからである（天沼・大坪・水谷1978）。日本語のアクセントと中国語のアクセントは両方「高さ」を用いて表されるが、日本語のピッチの変化は、各モーラ（音節）の間で生じる。これに対して、中国語は、一つの音節の中でピッチがいくつかのパターンで変化する。そのため、中国語母語話者はピッチを平坦に保つことが困難であると多くの研究者が指摘している（蔡1983、劉2008など）。また、モーラ言語である日本語では、1モーラの「み(実)」、2モーラの「はし(端)」、3モーラの「さくら(桜)」、4モーラの「たけのこ(筍)」、5モーラの「つまみもの(つまみもの)」、6モーラの「くつしたどめ(靴下留め)」などの平板型の語彙がある。このようにモーラ数が6モーラまで次第に増加しながら、ピッチを依然として平坦に保つのは中国語ではあまり見られないことである。音節言語である中国語には、1音節の「一」[i⁵⁵]、2音節の「今天」[tɕin⁵⁵tiɛn⁵⁵]、3音節の「公交车」[kuŋ⁵⁵tɕiau⁵⁵tɕh⁵⁵]、4音節の「稀稀拉拉」[ci⁵⁵ci⁵⁵la⁵⁵la⁵⁵]のような第一声によって構成された語彙が最大4音節までであり、4音節以上、かつ第1声の語彙は殆ど存在しない。それゆえに、中国人母語話者はモーラ数の長

い語彙を発音する際に、母語干渉のため、必ずどこかにアクセント核を置く傾向がある。これも中国語母語話者の平板式動詞活用形の正答率が低く、ピッチを平坦に保つことが困難である原因の一つだと考えられる。なお、今回の希望形「～たい」の結果では、起伏式の正答率が100%であるのに対し、平板式の正答率は0%である。実際には、現在では、母語話者が希望形「～たい」を発音する際にアクセントの揺れがあり、一部の母語話者も-2型アクセントで平板式動詞の希望形を発音する。従って、-2型と0型がいずれでも正解である。しかし、日本語学習者の生成結果では、このようなアクセントの揺れは見られない。日本語学習者はピッチを平坦に保つ平板型で発音するより、全員-2型アクセントで発音していた。このことは、中国人がピッチを平坦に保つことが困難であることを示している。

また、上記のⅠ型～Ⅳより、一つのアクセント規則が適用される活用形の正答率は二つのアクセント法則が適用される活用形よりはるかに高いことが見られる。それは、日本語学習者の立場から考えると、二つのアクセント法則を一つにまとめて暗記するのは両方覚えるより相対的に容易だからである。一方、日本語教育の立場から考えると、そもそも動詞活用形を学習する際に、平板式動詞と起伏式動詞の動詞活用形のアクセント法則を教師から教わっていないからである。このため、日本語学習者は、起伏式と平板式動詞の動詞変化のアクセント法則が同じであると思ったまま、統一的なアクセントパターンで動詞活用形を発音してしまっていると考えられる。

さらに、今回の生成結果から、多くの日本語学習者が動詞活用形のアクセント法則を知らず、動詞辞書形のアクセントに従って動詞活用形を発音してしまうことが見られる。例えば、「食べる」のアクセントは-2型であり、半数以上の学習者はその「テ」形、過去形、未然形「～ず」、条件形「～たら」を発音する際に、「食べ’る」と同様なアクセント型(-2型)で発音してしまうことがある。発音誤用例:「食べ’る→食べ’て」、「食べ’る→食べ’た」、「食べ’る→食べ’ず」、「食べ’る→食べ’たら」。

6 F0 動態の考察

本節では、音響的な特徴に着目し、Praat でアノテーションを行い、F0 曲線を抽出した。抽出した動詞活用形 F0 の動態を考察した上で、中国北方方言話者における一般的な特徴を明らかにする。その結果は以下のようなものである。

1) 丁寧形「～ます」

動詞類型、拍数に関わらず、丁寧形「～ます」形の正答率が全体的に100%に達する。表1の結果より、丁寧形「～ます」形を発音する際に、中国人日本語学習者の語末の「す」[su]の母音[u]のF0が欠如しない比率よりも、欠如する比率が高いことが見られる。

表1 丁寧形における語末「す」の母音[u]のF0の実現状況

	協力者1	協力者2	協力者3	協力者4	協力者5	協力者6	協力者7	協力者8	協力者9	協力者10
F0 曲線が欠如	67%	100%	0%	33%	100%	83%	100%	100%	100%	100%
F0 曲線が完全	33%	0%	100%	67%	0%	17%	0%	0%	0%	0%

では、丁寧形を発音する際に、中国人日本語学習者において、F0の欠如する比率が高いのはなぜかという点、主に二つの原因があると考えられる。

一つ目の原因は、目標言語である日本語では、狭母音の無声化が引き起こされるため、

母音のF0が欠如する。共通日本語における母音の無声化について、加藤・安藤(2016:136)は「無声子音と休止のあいだに狭母音がある場合に、この狭母音が無声化する」と述べている。また、斎藤(1997:95)も共通語などの狭母音が無声子音と句末のポーズとの間で起こることがよくあると述べている。例えば:バス[bas^u]、かし(菓子)[kaci]、あき(秋)[aki]、～です[des^u]、～ます[mas^u]。学習者が目標言語である日本語を習得する際に、特に、この点に注意し、「～ます」形の「す」の母音を無声化して発音することがあるのではないかと考え、それゆえに、F0の欠如する比率が全体的に高いという結果が出たと思われる。

一方、音声学的視点からすれば、母音の無声化の実態は狭母音が脱落し、前接の子音のみが発音される場合と狭母音脱落せず、無声母音として発音される場合という二種に分かれると佐久間(1929:229-234)が指摘した。この二種の母音の無声化の実態について、今回の調査より例を挙げると、図1の「頼みます」のサウンドスペクトログラムが示すように、「頼みます」の「す」のところで、母音[u]を表現した振幅(音圧)がまったく見られず、摩擦子音[s]の振幅(音圧)しか見えない。このことは、狭母音脱落し、前接の子音のみを発音していることを示す。このような母音脱落するのは、摩擦のあとで母音が無声化する

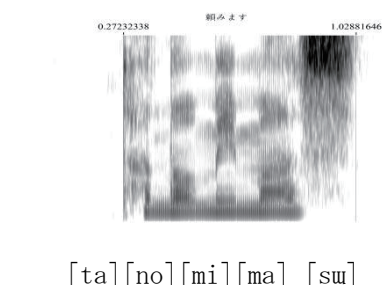


図1 狭母音が脱落するタイプ

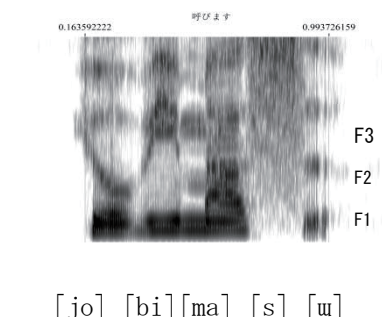


図2 狭母音が脱落しないタイプ

ると、母音の構えをする時間がないからである(加藤・安藤 2016:136)。図2の「呼びます」のサウンドスペクトログラムが示すように、「す」のところで、母音[u]を表現した振幅(音圧)が見られ、スペクトログラム上の母音を表現した第1フォルマントと第2フォルマント(色の濃い部分)もはっきり見られる。このことは、狭母音脱落せず、無声母音として発音していることを示す。

今回の調査データより、丁寧形を発音する際に、語末の「す」のF0が欠如する場合、殆どの日本語学習者は狭母音脱落し、前接の子音のみを発音する。これに対し、狭母音脱落せず、無声母音として発音する学習者は少数であることが分かった。また、同一話者でも、「狭母音脱落し、前接の子音のみを発音する」か「狭母音脱落せずに無声化として発音する」かは揺れがあることも観察された。二つ目の原因は中国語における「軽声」に関わると考えられる。中国語

における軽声音節は声調を有する一般の音節に比べ、弱くて短く発音される弱化音節である(平井 2012:81)。中国語の軽声音節において、母音の無声化という音声変化の現象がある。軽声音節における母音の無声化現象について、平井(2012:81-82)は母音の舌位の高低に関係なく生じると指摘した。例えば、舌の位置が高い母音[i]、[u]、[y]が無声化する語彙:「意思」[isi]、「豆腐」[tɔufu]、「进去」[tɛitɕy]。舌の位置が低い母音[a]が無声化する語彙:「菩薩」[p^husa]、「顔色」[iənsɔ]。これらの軽声音節における母音は発音された際に、わずかな舌の位置と唇の形の特徴のみが残されている。さらに、唇の形の特徴が脱落する場合もある。このような場合は、母音が既に発音されないと認められる(林・王 1992:167)。また、摩擦音に後続する母音が弱化する無声化現象が生じやすいことは、あらゆる言語に見られる現象であると述べている(平井 2012:82)。すなわち、中国語は日本語

と同様に摩擦音に後続する母音が無声化しやすい。中国語における軽声音節の影響を受けたため、中国人日本語学習者は「～ます」形を発音する際に、語末の「す」の母音[u]が無声化しやすいと考えられる。

2) 一段動詞の命令形と意志形

生成調査の結果より、一段起伏式動詞の命令形と意志形の正答率が比較的に高いことが分かった。音響分析の結果では、拍数に関わらず、一段起伏式動詞の命令形の持続時間が意志形とほぼ同じであることが明らかになった。同一話者の「出’よ」と「出よ’う」のアクセントは正確だが、持続時間も（「出’よ」0.434 s、「出よ’う」0.471 s）ほぼ同じである。また、半数くらいの日本語学習者が一段起伏式の命令形を意志形と同様のアクセント型(-2型)で発音することがある。このように発音すると、命令形を意志形に聞き間違えやすくなる。こうなるのは中国人日本語学習者が長音と短音の区別がつかないからである。音節を単位とする中国語においては、母音の長短が弁別的に働かないため、中国人日本語学習者にとって、日本語の長音と短音の産出と知覚が難しいことが予測される。この点に関しては、これまで、多くの日本語教育の研究者は実験や調査によって中国人日本語学習者にとって、日本語の長音と短音の産出と知覚が困難であることを検証した(戸田2003、栗原2004など)。

3) 連用形「～ながら」、語尾が「む」、「ぶ」である五段動詞の「～テ」形(飲んで)、「～テモ形」(飲んで)、

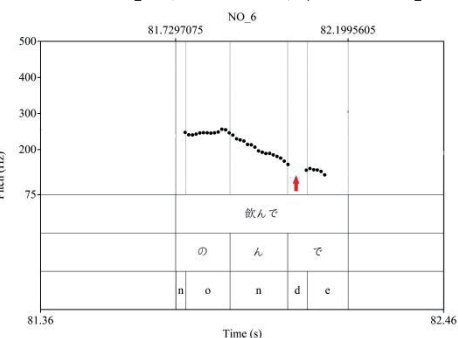


図3 「～テ」形

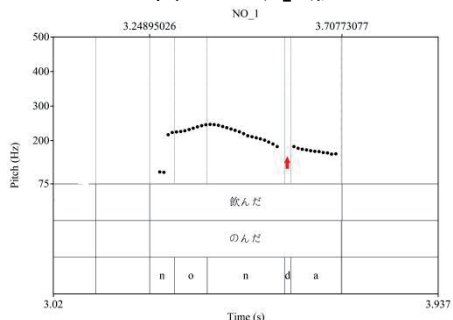


図4 「～た」過去形

音響分析によって、多くの中国人学習者が発音した連用形「～ながら」、及び語尾が「む」、「ぶ」である五段動詞の「～テ」形(飲んで)、「～テモ形」(飲んで)、「～た」過去形(飲んだ)、「～たら」条件形(飲んだら)のF0においては、有声音「が」[ga]、「で」[de]、「だ」[da]のところでは、F0が欠如することがしばしばあることが明らかになった。例えば、図3、4に示したように、動詞活用形「飲んで」[nonde]、「飲んだ」[nonda]における有声破裂子音[d]のところのF0が欠如する。F0が欠如するということは、中国人学習者が日本語の有声破裂子音を発音できないことを示している。その原因について、母語の転移ということが考えられる。中国語では有気音・無気音が弁別の特徴になっているのに対して、日本語では有声音・無声音が弁別の特徴となっている(柳・西郡2006:75)。

実際に中国語(標準語)では、日本語のような有聲無気音も存在するが、それは無声無気音の異音として存在するだけで、有聲・無声は弁別の特徴となっていない。また、王(2017)は中国人日本語学習者を対象とする日本語語頭有声破裂音の生成実験によって、統計的には、*0.05の水準で語頭有声破裂音/b/($F(2, 34)=.103, p=.902$)、/d/($F(2, 34)=2.255, p=.120$)、/g/($F(2, 32)=2.364, p=.110$)のVOT値において地域差が見られず、どの地域話者にも完全有声音を発音することが困難であることを指摘した。これらより、中国人学習者が日本語の

有声破裂子音を発音する際に、常に中国語の無気音を代替して発音することが予想できる。それゆえに、有声音部分のF0が欠如するのは当然のことであろう。なお、今回のF0動態を考察することによって、中国人学習者(北方方言話者)が語中・語末における完全有声音を発音することが困難であることも確認できる。

7 まとめ

本稿では、中国北方方言話者を対象として15種類の動詞活用形の生成調査を実施し、学習者の習得実態及びF0の特徴を明らかにした。そのうち、動詞活用形のF0動態の考察においては、本稿で扱う特徴を除き、他には、「語頭F0曲線の下降」、「F0曲線の急降」、「ピッチ曲線の下降幅が大きい」という特徴がある。これらの特徴は動詞のみならず、先行研究により、名詞、文節、文にも観察され、共通的な特徴である。また、学習者の習得実態及びF0の特徴を明らかにすることによって、学習者自身の動詞活用形アクセントの問題に注意が喚起でき、より標準的な日本語を発音させることができると考えられる。

参考文献

- 天沼寧・大坪一夫・水谷修(1978):『日本語音声学』くろしお出版
- 平井勝利(2012):『教師のための中国語音声学』白帝社
- 平野宏子・広瀬啓吉・河合剛・峯松信明(2009):「母語話者と中国語話者の日本語朗読音声の基本周波数パターンの比較」日本音響学会誌 65 巻 2 号 pp. 69-80
- 加藤重広・安藤智子(2016):『音声学講義』研究社
- 栗原通世(2004):「中国北方方言話者の日本語長音の知覚特徴」言語科学論集 第8号 pp. 1-12
- 許霖(2009):「中国人日本語学習者における語頭第1拍の音響的特徴:中国北方方言話者の語頭ピッチ曲線を中心に」名古屋大学言語文化研究会 ことばの科学 22, pp. 5-22
- 王睿来・林良子・磯村一弘・新井潤(2017):「中国語母語話者による日本語名詞アクセントの習得:知識・産出・知覚の関係から」『中国語話者のための日本語教育研究』, 7, pp. 61-75.
- 王晓青(2003):「台湾日本語学習者への日本語語アクセントの指導法の基礎的研究:中国語の四声と日本語語アクセントの対照研究の観点から」広島大学 博士論文
- 王鳳翔(2017):「中国語母語話者による日本語語頭破裂音の生成—子音の調音位置・地域差とVOTの関係—」日本言語学会2017年度全国大会(第155回)予稿集
- 王鳳翔(2017):「東京語動詞の打消し形のアクセントの生成—中国北方方言話者によるアクセントの特徴とその要因—」日本音声学学会2017年度全国大会(第31回)予稿集
- 林焘・王理嘉(1992):『语音学教程』北京大学出版社
- 柳悦・西郡仁朗(2006):「中国・上海の学習者による日本語の有声子音・無声子音の知覚と学習—「説明」と「反復練習」の学習効果—」『日本語研究』26、東京都立大学・中国首都大学 日本語研究会 75-83.
- 劉佳琦(2008):「東京語の動詞・複合動詞アクセントの生成について」,『2006年清華大学日本言語文化国際フォーラム論文集』,清華大学出版社, pp. 424-439
- 佐久間鼎(1929):『日本語音声学』京文社
- 蔡全勝(1983):「中国人に見られる日本語アクセントの傾向」,『在中華人民共和国日本語研修センター紀要 日本語教育研究論纂』,第1号, pp. 26-31.
- 斎藤純男(1997):『日本語音声学入門』三省堂
- 戸田貴子(2003):「外国人学習者の日本語特殊拍の習得」音声研究 第7巻第2号 pp. 70-83

中国語母語話者を対象とした日本語アクセントの知覚訓練の効果

王 睿来 (神戸大学)・林 良子 (神戸大学)
磯村 一弘 (国際交流基金)・新井 潤 (国際交流基金)
wang_ruilai@yahoo.co.jp

1. はじめに

第二言語習得において、自然な発音を身につけるためには、韻律が重要であるとしばしば指摘される。日本語においては、アクセントは韻律の基礎であるが、学習者がこれを習得するためには、効果的な訓練方法を確立し、教育に応用することが期待される。そこで、本研究では日本語アクセントの知覚訓練を行い、その効果を調べた。この訓練を通じて、中国語を母語とする学習者がどの程度日本語アクセントを適切に発音できるようになるのかを検討した。

2. 先行研究

中国語を母語とする日本語学習者を対象としたアクセントの訓練・指導に関する研究は、西郡・八山 (1996)、崔・吉田 (2007a)、崔・吉田 (2007b) などがある。西郡・八山 (1996) では、日本語学習初級段階でのアクセントの説明と知覚練習がアクセント知覚に効果があることが確認された。崔・吉田 (2007a) では、VT 法を用いる指導法とアクセント記号を見ながら発音練習する指導法はともにアクセントの生成に効果があることが確認されたが、指導法による効果の差は見られなかった。崔・吉田 (2007b) は、VT 法と「リピート・各自練習」の二つの方法を用いてアクセント指導を行い、VT 法のほうが効果的と結論づけた。

上記の研究における訓練法・指導法は、アクセントの知覚か生成のどちらか一方に効果があることを示したものであるが、知覚と生成の両方について効果があるかどうかは検討されていない。また、訓練・指導の効果が未訓練語に及ぶかについても検討されていない。

一方、第二言語においては、知覚訓練は知覚に効果があるだけでなく、生成にも訓練効果が転移するということが、単音レベル (Bradlow et al, 1997 ; Bradlow et al, 1999) においても、韻律レベル (Wang et al, 2003 ; 楊・山田, 2016) においてもこのことが確認されている。さらに、知覚訓練による生成への転移は、非訓練語においても起こることも報告されている (Bradlow et al, 1997 ; Wang et al, 2003)。そこで、日本語アクセントについても、知覚訓練を行うことにより、同じような結果が得られるのかについて検討する価値があると考えられる。

3. 本研究の目的

以上の知覚訓練の先行研究に基づき、本研究では中国語を母語とする日本語学習者を対象にアクセントの知覚訓練を行い、次の2点を明らかにすることを目的とする。

- 1) 知覚訓練により、学習者のアクセント知覚能力は向上するのか。
- 2) 知覚訓練により、学習者のアクセント生成能力も向上するのか。

4. 方法

本研究では、2週間計8回にわたるアクセント知覚訓練を実施した。訓練の前後に行ったアクセント知覚テストとアクセント生成テストの結果を比較した。

4.1. 実験協力者

実験協力者（以下、協力者）は、中国の大学で日本語を主専攻とする1年生24名（男性4名、女性20名；17歳～20歳，平均年齢18歳）であった。全員が北方方言話者で、日本語学習歴は6か月程度であった。日本滞在経験がある者はいなかった。協力者は知覚訓練を受ける知覚訓練グループ（12名）、テストのみを受けるコントロールグループ（12名）に分けた。

4.2. 実験対象語

実験対象語は1拍～4拍の各アクセント型の単語で、合計168語を使用した。168語中84語は訓練中に呈示した訓練用単語で、残りの84語はテストのみで呈示した般化テスト用単語であった。訓練用単語は「ま」「た」「ば」をそれぞれ拍数分用いて作った無意味語42語（例：まま、たたた）と協力者の未知語42語（例：沼、鮑）であった。般化テスト用単語はすべて協力者にとっては未知語であった。

表 1: 実験対象語

訓練用単語		般化テスト用単語	
無意味語 42語	未知語1 42語	未知語2 42語	未知語3 42語
知覚テスト	生成 Test1	生成 Test2	生成 Test3

4.3. 実験の流れ

実験前に、日本語のアクセントについて解説を行った。解説はアクセントとは何か、アクセントの型、アクセントの制約、アクセントの記号に関する内容であった。実験は「Pre-test → 知覚訓練 → Post-test」という流れで行い、Pre-test と Post-test は共通のテストを行った。

表 2: 実験の流れ

実施順序	グループ	
	知覚訓練	コントロール
1	Pre-test	Pre-test
2	知覚訓練	(訓練なし)
3	Post-test	Post-test

4.4. 知覚訓練

知覚訓練では、PowerPoint と回答用紙を使用した。回答用紙は「回答欄」、「正解欄」、「○×欄」からなった。訓練は、まず PowerPoint のスライドに、助詞「が」付きの訓練用単語が平仮名で呈示され、刺激音声は1度流れる。協力者は刺激音声を聞き、回答用紙の「回答欄」のアクセント核があると判断した箇所にアクセント記号「ㄱ」を記入する。アクセント核がないと判断した場合は、単語の最後の仮名の右上に「ー」を記入する。その後、PowerPoint は次のスライドに移り、アクセント記号が書かれた正解が呈示され、音声はもう一度呈示される。協力者は正解を回答用紙の「正解欄」に書き写し、自分自身の回答が正

しかなかったかどうかについて「○」「×」で「○×欄」に記入する。知覚訓練は2週間で8回行った。訓練時間は毎回約30分、合計約4時間であった。毎回の訓練は、訓練用単語の無意味語42語と有意味語42語の計84語について、男性1名の刺激音声と女性1名の刺激音声を使用して行った。刺激音声は、男性2名と女性2名の計4名分を使用した。

4.5. テスト

テストについて、Pre-test と Post-test では共通の知覚テストと生成テストを行った。知覚テストは、訓練用単語の無意味語(表1)を対象語とした。刺激音声は東京方言を母方言とする日本人男性が読み上げた音声であった。協力者は助詞「が」付きの対象語の音声を聞き、アクセント核があると判断した箇所にアクセント記号「 \uparrow 」を記入した。アクセント核がないと判断した場合は、対象語の最後の仮名の右上に「 $\bar{}$ 」を記入した。生成テストは3つのテストからなった。Test1では、訓練用単語の未知語1(表1)を助詞「が」付きでアクセント記号とともに呈示し、協力者に読み上げてもらった。Test2では、般化テスト用単語の未知語2(表1)を用い、Test1と同じように読み上げてもらい、Test3では、般化テスト用単語の未知語3(表1)を含めた「対象語+が+述語」という短文(例:網戸があります。)を、対象語のアクセント型とともに呈示し、読み上げてもらった。Test1とTest2では単語をすべて平仮名で呈示し、Test3では漢字がある場合は漢字に振り仮名をつけて呈示した。テストは知覚テストから生成テスト(Test1→Test2→Test3)の順で行った。テストを終了するのに、30分程度要した。

4.6. 生成テストにおける発話のアクセント聴覚判定

生成テストにおけるアクセント聴覚判定は、第一著者と東京方言を母方言とする日本人A・Bの3名で行った。3名は全員日本語音声学に関する知識を持っており、且つ日本語教育の経験者でもある。聴覚判定はアクセント核の有無と位置に着目し、まず、筆者と日本人Aが行った(判定結果の一致率:91.7%)。その後、判定が異なるものを抽出し、日本人Bが再判定を行った。

4.7. 分析方法

知覚テストは、協力者による回答が対象語のアクセントと同じであれば1点を、異なれば0点を付与した。生成テストは、協力者に生成されたアクセントが対象語のアクセントと同じと判定されれば1点を、異なると判定されれば0点を付与した。知覚テストと生成テスト(Test1・Test2・Test3)はそれぞれ42点満点であるが、分析に当たっては、パーセンテージに換算した正答率・正用率を使用した。

5. 結果と考察

結果と考察について、正答率・正用率を中心に知覚テストと生成テストの順番で呈示していく。

5.1. 知覚テスト

表3にアクセント知覚テストの結果を示す。ここでは知覚訓練を通して、学習者のアクセント知覚能力が向上することが示された。全体の正答率は、コントロールグループでは

Pre-test と Post-test はともに 70.0%程度であったが、知覚訓練グループでは Pre-test は 60.7%であったのに対して、Post-test は 96.4%であった。この知覚訓練の効果を検討するため、グループという被験者間要因とテストという被験者内要因（Pre-test と Post-test）の 2 要因の分散分析を行った結果、テストの単純主効果は知覚訓練グループにおいて有意 ($p < .001$) で、交互作用 ($F(1,22) = 17.0, p < .001$) も有意であった。

以上の結果から、知覚訓練は学習者のアクセント知覚能力の向上に寄与していることが示された。この結果は単音レベルの Bradlow et al (1997) や Bradlow et al (1999)、韻律レベルの Wang et al (2003) や楊・山田 (2016) と一致するが、本研究を通して、韻律レベルにおいては、日本語アクセントについても同様の効果が確認されたとと言える。

表 3: アクセント知覚テストの正答率

知覚訓練グループ		コントロールグループ	
Pre-test(%)	Post-test(%)	Pre-test(%)	Post-test(%)
60.7	96.4	69.0	75.8

5.2. 生成テスト

表 4 にアクセント生成テストの結果を示す。この結果から、知覚訓練を通して、学習者のアクセント生成能力も向上することが示された。アクセント生成テストの全体の正用率は、コントロールグループでは Pre-test と Post-test の 6 つのテストの正用率がすべて約 60.0%であったが、知覚訓練グループでは、Pre-test の 3 つのテストの正用率がすべて 60.0%程度であったのに対して、Post-test の Test1・Test2・Test3 はそれぞれ 95.8%、92.1%、84.9%となった。知覚訓練の生成への効果を検討するため、グループという被験者間要因とテストという被験者内要因（Pre-test の Test1・Test2・Test3 と Post-test の Test1・Test2・Test3）の 2 要因の分散分析を行った。その結果、単純主効果について、グループの単純主効果は Post-test の Test1・Test2・Test3 に有意で（すべて $p < .001$ ）、テストの単純主効果は知覚訓練グループのみに有意 ($p < .001$) であった。交互作用も ($F(5,110) = 14.4, p < .001$) 有意であった。さらにテストの単純主効果について多重比較を行ったところ、「Pre-test の Test1 < Post-test の Test1 ($p < .001$)」「Pre-test の Test2 < Post-test の Test2 ($p < .001$)」、「Pre-test の Test3 < Post-test の Test3 ($p < .001$)」、「Post-test の Test3 < Post-test の Test1」 ($p < .005$) において有意であることが示された。

以上の結果から、知覚訓練の効果が生成面に転移したということが示唆された。そして、訓練語 (Test1) だけではなく、訓練しなかった語 (Test2・Test3) にも転移したことが示された。これはアクセントの知覚と生成に相関があるためであると考えられる (王・林・磯村他, 2018)。知覚訓練の効果が知覚面だけではなく、生成面にも転移するということが、先行研究 (Bradlow et al, 1997 ; Bradlow et al, 1999 ; Wang et al, 2003) において示されているが、本研究を通して、日本語アクセントにおいても同様の結果が確認された。ただし、Post-test の Test3 の正用率は Post-test の Test1 より有意に低かった。本研究の知覚訓練の刺激は「単語+助詞」であったため、このような知覚訓練は「単語+助詞」レベル (Test1・Test2) までは効果的であるが、「単語+助詞+述語」という短文レベル (Test3) になると訓

練の効果が直接的には出にくくなるためであると考えられる。ただし Test3 の Pre-test に比較すれば Post-test では正用率が上がっており、訓練効果は見られたと言える。

表 4: アクセント生成テストのテスト別の正用率

テスト	知覚訓練グループ		コントロールグループ	
	Pre-test(%)	Post-test(%)	Pre-test(%)	Post-test(%)
Test1	63.1	95.8	61.1	62.9
Test2	65.5	92.1	62.1	66.1
Test3	60.5	84.9	54.4	60.1

続いて、知覚訓練グループによるアクセント生成テストの拍数別の正用率とアクセント型別の正用率について見ていく。なお、テスト別ではなく、Pre-test の Test1・Test2・Test3 を Pre-test 全体として、Post-test の Test1・Test2・Test3 を Post-test 全体として合わせて集計した。Pre-test 全体と Post-test 全体の正用率は表 5 の通りである。

表 5: 知覚訓練グループによるアクセント生成テストの正用数と正用率

拍数	アクセント型	語数	Pre-test (%)	Post-test (%)
1 拍	0 型	108	74(68.5)	99(91.7)
	1 型	108	85(78.7)	104(96.3)
	合計	216	159(73.6)	203(94.0)
2 拍	0 型	108	94(87.0)	106(98.1)
	1 型	108	75(69.4)	108(100.0)
	2 型	108	66(61.1)	96(88.9)
	合計	324	235(72.5)	310(95.7)
3 拍	0 型	108	95(88.0)	103(95.4)
	1 型	108	59(54.6)	102(94.4)
	2 型	108	54(50.0)	92(85.2)
	3 型	108	72(66.7)	98(90.7)
	合計	432	280(64.8)	395(91.4)
4 拍	0 型	108	91(84.3)	99(91.7)
	1 型	108	55(50.9)	104(96.3)
	2 型	108	31(28.7)	83(76.9)
	3 型	108	60(55.6)	98(90.7)
	4 型	108	42(38.9)	83(76.9)
	合計	540	279(51.7)	467(86.5)
全体の合計		1512	953(63.0)	1375(90.9)

各拍のアクセント型別の語数=実験対象語 3 語×協力者 12 名×3 テスト=108 語

拍数別の正用率について、Pre-test では、1 拍語と 2 拍語がそれぞれ 73.6%、72.5%と同じ程度であったが、2 拍語から 4 拍語にかけて下がる傾向が見られた。Post-test では、1~4 拍語はすべて 85.0%以上の正用率で、1 拍語~3 拍語はすべて 90.0%以上であった。つまり、知覚訓練を通して、1~4 拍語はすべて大幅に改善された。

アクセント型別の正用率について、Pre-test では、1 拍語の 1 型、2 拍語~4 拍語の 0 型は比較的高かったが、他のアクセント型は低かった。Post-test では、2 拍語の 2 型、3 拍語の 2 型、4 拍語の 2 型と 4 型を除き、他のアクセント型の正用率はすべて 90.0%以上であった。

4 拍語の 2 型と 4 型はともに 76.9%と特に低かった。これらの結果から、知覚訓練を通して、改善されにくいアクセント型があるが、学習者がほとんどのアクセント型を正しく生成できるようになったと言える。

最後に、知覚訓練グループによるアクセント正用率の個人差について検討する。Pre-test では、学習者の正用率は 24.6%～83.3%と個人差が大きかったが、Post-test では 73.0%の学習者 2 名と 84.9%の学習者 1 名を除き、他の 9 名は全員 90.0%以上と個人差が縮小した。上昇率については、12.7%～60.3%と全員上昇した。これらの結果から、学習者 3 名を除き、他の学習者は訓練後アクセントをほぼ正しく生成できるようになったと言える。

6. おわりに

本研究では、知覚訓練が学習者のアクセント知覚能力と生成能力に与える影響について検討を行った。その結果、知覚訓練はアクセント知覚能力の向上に効果があり、知覚訓練の効果が生成面に転移し、さらに訓練語だけではなく、非訓練語の生成にも転移した。以上のことから、アクセントの知覚訓練は第二言語の音韻習得上、知覚面のみならず生成面にも効果的な訓練法であることが本研究の結果から示された。

参考文献

- Ann R. Bradlow., David B. Pisoni., Reiko Akahane-Yamada & Yoh'ichi Tohkura.(1997) "Training Japanese listeners to identify English /r/ and /l/: IV. Some effects of perceptual learning on speech production",The Journal of the Acoustical Society of American 101(4),2299-2310.
- Ann R. Bradlow., Reiko Akahane-Yamada., David B. Pisoni & Yoh'ichi Tohkura.(1999) "Training Japanese listeners to identify English /r/ and /l/: Long-term retention of learning in perception and production",Perception & Psychophysics 61 (5),977-985.
- 西郡仁朗・八山京子 (1996) 「北京語母語話者による東京語アクセントの聞き取りの習得：日本語学習者初級段階における詳説と練習の効果」『音声言語の韻律特徴に関する実験的研究 平成 8 年度研究報告書』, 81-88.
- 王睿来・林良子・磯村一弘・新井潤 (2018) 「中国語母語話者による日本語名詞アクセントの習得：知覚と生成の関係に着目して」『ことばの科学研究』19, 81-96.
- 崔春福・吉田光演 (2007a) 「VT 法による日本語の単一語アクセントの指導法の考察」『広島大学大学院総合科学研究科紀要 I, 人間科学研究』2, 11-21.
- 崔春福・吉田光演 (2007b) 「中国語母語話者を対象とした日本語の複合語アクセントの指導法：VT 法の指導効果をめぐって」『欧米文化研究』14, 71-84.
- Wang,Y., Jongman,A., & Sereno,J.A.(2003) "Acoustic and perceptual evaluation of Mandarin tone productions before and after perceptual training", The Journal of the Acoustical Society of American 113(2),1033-1043.
- 楊姝怡・山田玲子 (2016) 「日本語母語話者を対象とした中国語四声の知覚訓練と語彙訓練の効果：訓練前後の比較」『日本音声学会第 30 回全国大会予稿集』, 32-37.