

中国語を母語とする日本語学習者による感情音声の知覚

李 歆玥 (神戸大学大学院)・羅 米良 (大連外国語大学)・林 良子 (神戸大学)
lixinyue12200@163.com・luomiliang@gmail.com・rhayashi@kobe-u.ac.jp

1. はじめに

音声も表情と同様に、喜び、怒り、悲しみなどの感情の知覚に文化を超えた共通性があるということがたびたび指摘されている(Tickle 2000 他)。一方で、感情の知覚は母語話者と非母語話者の間で異なる傾向があることも報告されている。感情音声の表出には、声の高さ・長さ・強さ・声質などの複合的な音響特徴が関与していることが知られており、各言語によってその音響特徴も異なる(Johnstone 2001)。Campbell & Erickson (2004)では様々な発話音声に対し、自由に評価してラベリングするという実験を行ったところ、異なる母語話者間では同一音声に異なったラベリングを行うことが観察され、このような知覚の相違が感情音声知覚の誤解を引き起こす可能性があるとして述べている。日本語教育においては、感情音声の聞き取り教材(西端 1996)がわずかながら開発されているものの、感情は普遍的なものであるため指導する必要がないという認識が通用しており、日本語教育の現場において感情の表出、理解が扱われることがほとんどないとされる(中林 2011)。そのため、日本語学習者による感情音声の産出、知覚についてはこれまでにあまり多くの研究が行われてこなかった。

本研究では、このような背景をもとに、中国語を母語とする日本語学習者(以下 JFL)は日本語音声に含まれる感情を一体どの程度知覚できるのか、日本語学習歴により感情音声の知覚が変化するのか、また、学習者に混同されやすい感情パターンは何かについて明らかにすることを目的とする。

2. 先行研究

2.1 日本語母語話者による感情音声の産出・知覚

重野(2001)では「東京」「河原崎さん」「11時半」「さようなら」「そうですか」に含まれる「幸福」「驚き」「怒り」「嫌悪」「恐れ」「悲しみ」という6つの感情発話の産出について分析した。その結果、基本周波数とエネルギーの標準偏差が、「怒り」では他の感情より高く、「嫌悪」では最も小さく、大きな差異が認められた。田川(2001)では、第1拍目の後に下降が来る「見ないの」という音声の「平静の問いかけ」と「非難の問いか

け」について分析を行った。「非難の問いかけ」の冒頭部分から上昇部分までのピッチ曲線は「平静の問いかけ」より急な傾斜が見られ、「非難の問いかけ」の全体時間長と各音節の時間長は「平静の問いかけ」より長いと述べている。2つの研究のみを挙げたが、これだけでも日本語の感情音声産出には、発話のピッチパターンや持続時間、声の大きさが大きく関与していることが分かる。

2.2 非日本語母語話者による感情音声の産出・知覚

エリクソン・昇地 (2006) は、日本語話者が「喜び」「悲しみ」「怒り」「驚き」「疑い」の5種類の感情を込めて発話した「バナナ」を日本人母語話者、アメリカ英語母語話者、韓国語母語話者に聞いてもらい、どの感情に知覚されるのかについて検討を行った。その結果、アメリカ英語母語話者の正答率は他の母語話者に比べ全体的に高かった。日本語母語話者の正答率が最も低く、これは日本人被験者が語の意味と感情音声による情報の間に差異を感じたためであろうと述べている。この際には、「怒り」「喜び」「驚き」がよく混同され、その原因はピッチ曲線と声の大きさが類似していることであるとした。さらに、「悲しみ」と「疑い」も混同されるが、その原因はF0最小値が低いことと、声の大きさが小さいこととされた。また、韓国語母語話者は「怒り」について正答率が特に高いことも示された。

中林 (2011) では、「韓国」「毎日」(頭高型)、「現金」「片仮名」(中高型)、「カラオケ」「温泉」(平板型)の6つの4拍語を刺激語とし、問い返し疑問文に含まれた「嫌」と「驚き」の感情について、ロシア人日本語学習者を対象に分析を行った。その結果、「嫌」という感情の認識が困難であることがわかった。また、アクセント核の有無によって感情の聴取の難易度が異なる可能性、さらに正答率の低い学習者では特定のアクセント型からの感情の聴取が困難である可能性が示唆された。

このように非母語話者が日本語音声から感情を知覚する際、母語話者と同様の傾向も認められるが、聞き手の母語が日本語感情音声の知覚に影響する可能性も示された。

3. 日本語母語話者による判定実験

3.1 刺激音の収録

多くの先行研究においては、エクマンが提案した基本六感情 (Ekman 1992) が用いられている。本研究においては他の感情とよく共起するとされる「驚き」と「恐れ」を研究対象から除外し、通言語的に知覚されると報告されている「平静」(Calm)を追加した (Tickle 2000)。これにより、「喜び」「怒り」「悲しみ」「平静」の4感情を対象とした。

刺激語は、重野 (2001) を参考に、単語自体が何らかの感情情報を示さず複数の感情を表

現しやすいとされる固有名詞と一般動詞、「小野さん」「ドイツ」「12 時」「分かった」を選択した。

これらの刺激語を 4 名の日本語母語話者（男女各 2 名、以下 JN）に、できるだけ自然かつ明確に感情を表出するように指示し、各感情とも 1 文につき 3 回発話してもらった。1 回目と 3 回目を除き、残り第 2 回目の録音を刺激音とした。このように、計 64 トークン（4 語×4 感情×4 名）の刺激音を収録した。

3.2 刺激音の評価と選択

刺激音の妥当性を評価するため、収録した 64 トークンをランダムに並べ替え、5 名の JN に聞かせ、4 つの感情から 1 つだけ選ぶように指示した。その結果、母語話者によってそれぞれ高い正答率が得られたため（「喜び」70.0%、「怒り」76.3%、「悲しみ」68.8%、「平静」88.8%）、64 トークン全てを聴取実験の刺激音とした。

4. 中国人日本語学習者を対象とした聴取実験

4.1 聴取実験

3.2 で得られた刺激音を用いて聴取実験を行った。被験者は中国語を母語とし、日本語を専攻する、日本滞在経験のない大学 1 年生 14 名、2 年生 14 名（日本語能力試験 N3 合格者）、3 年生 14 名（N2 合格者）および院生 14 名（N1 合格者）であり、合計 56 名であった。64 トークンを 16 例ずつランダムに並べ替え、4 試行を準備した。それぞれの刺激音提示の前に 260ms 間の長さの合図音を提示し、その 600ms 後に刺激音を提示した。刺激音の提示後、3s 間の解答記入時間を設けた。引き続き次の刺激音を提示した。試行間に休憩時間 15s を入れ、4 試行すべて終了するのに約 6 分間がかかった。

4.2 実験結果：中国人日本語学習者の知覚

表 1 に学年別の知覚結果を示す。Scherer et al (2001)では 9 カ国の母語話者を対象としたドイツ語の聴取実験を行い、9 カ国の母語話者がチャンスレベル以上の正答率で感情を正しく知覚できたと報告している。本研究の結果では、1、2 年生の正答率は 45.7%、55.3% 程度であり、中国人初級学習者では日本語感情音声を正確に知覚できていない可能性が示された。各感情の知覚の傾向をより分かりやすく示すために、図 1 に JFL の平均正答率を示す。Scherer et al (2001)の指摘と同様に、「平静」の知覚が最も正答率が高く、次に「怒り」「悲しみ」という順に正答率が下がった。また、エリクソン・昇地 (2006) に指摘されたように、「喜び」の正答率が最も低くなることが分かった。さらに、図 1 に示すように、1 年生から院生まで外国語の習得がすすむにつれ、正答率が上がることも示唆され

た。

表 1 各学年の感情知覚 confusion matrix (%)

1年生	喜び	怒り	悲しみ	平静	2年生	喜び	怒り	悲しみ	平静
喜び	36	27	16	21	喜び	45	27	5	22
怒り	17	51	15	17	怒り	19	54	13	14
悲しみ	8	15	44	33	悲しみ	5	8	52	35
平静	18	7	22	53	平静	9	6	16	69

3年生	喜び	怒り	悲しみ	平静	院生	喜び	怒り	悲しみ	平静
喜び	55	24	6	15	喜び	59	29	5	7
怒り	12	69	7	12	怒り	17	69	8	6
悲しみ	2	3	64	31	悲しみ	1	3	65	31
平静	1	6	12	81	平静	1	2	14	83

表 1 の縦欄は刺激語の感情、横欄は回答された感情を示す。

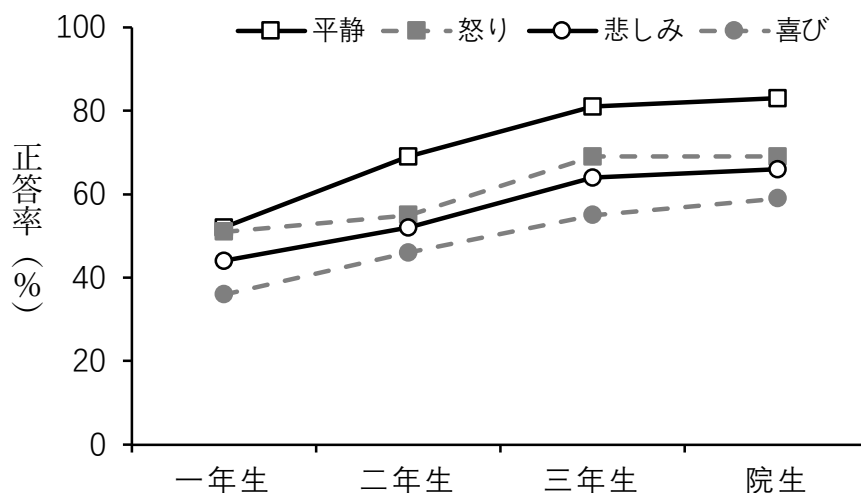


図 1 平均正答率

4.3 日本語感情音声の韻律的特徴

表 2 に「小野さん」のトークンを対象に、平均基本周波数(F0)、平均最高基本周波数(F0max)、平均最低基本周波数(F0min)、平均ピッチレンジ(F0range)、平均持続時間(Dur)およびインテンシティの平均最大値(Int)を示す。

表 2 「小野さん」の音響的特徴

	F0(Hz)	F0max(Hz)	F0min(Hz)	F0range(Hz)	Dur(ms)	Int(dB)
喜び	436	439	297	142	672	83
怒り	338	345	196	149	660	82
悲しみ	237	261	212	49	623	73
平静	289	305	234	71	666	80

「喜び」の発話においては、どの学年の学習者も平均 27% の高い確率で「怒り」との混同が観察された。これは、日本語の「喜び」の音響特徴が、F0 最高値が高く、発話持続時間が長く、インテンシティが大きいことに起因していると考えられる。これは、中国語の「怒り」(劉 2011) に似ており、このことが混同の原因であると考えられた。日本語の「怒り」は、ピッチレンジが最も大きく、ピッチ曲線の下降が顕著であり、声質特徴としてきしみ声 (creaky voice) も観察された。エリクソン・昇地 (2006) によれば、「怒り」は持続時間が一番短く、正答率も一番低いが、本研究の結果とは異なっていた。このことは「怒り」の度合いと関係があるのではないかと考えられる。鶴・武田 (2007) において、「激怒」の持続時間が「軽い怒り」より長いと述べられている。よって、感情の度合いも韻律的特徴に反映し、さらに知覚に影響を与えたと考えられる。

また、「平静」と「悲しみ」も混同が見られたが、この原因は、低い基本周波数と幅の狭いピッチレンジという特徴が類似しているためと考えられる。「悲しみ」の音声は、平坦なピッチ曲線が特徴であり、F0 最高値と最低値も「平静」より低く、「平静」と区別している可能性が示唆された。

5. まとめ

本研究では、日本語母語話者が「喜び」「怒り」「悲しみ」「平静」の感情を込めた発話に対して、中国人日本語学習者がどのように知覚するかについて検討した。学習歴が浅い被験者の場合には、感情知覚の正答率が必ずしも高くはないが、学習時間の増加につれ、正答率も上がっていく可能性が示唆された。また、4 つの感情音声の中では「平静」の正答率が一番高く、「喜び」ではの知覚が最も困難であることが観察された。

次に、感情音声の音響的特徴についても分析を行った。「喜び」では、平均基本周波数、発話持続時間およびインテンシティともに一番高く、「悲しみ」では平均基本周波数、発話持続時間、インテンシティ、ピッチレンジともに一番低かった。さらに、「悲しみ」では、

平坦なピッチ曲線が観察され、低い基本周波数、狭いピッチレンジを示す「平静」と混同されることも観察された。今後は感情の種類を増やしてさらに検討していくことで、日本語学習者が日本語感情音声を知覚・産出する際に何が問題となるのかを明らかにしていきたい。

謝辞

本研究は科学研究費（基盤 B）課題番号 17H02352 の研究成果の一部である。

引用文献

- Campbell, N. & Erickson, D. (2004) “What do People hear? A Study of the Perception of Non-verbal Affective Information in Conversational Speech,” *Journal of the Phonetic Society of Japan* 8(1), 9-28.
- Ekman, P. (1992) “An Argument for Basic Emotions,” *Cognition And Emotion*, 1992,6(3/4), 169-200.
- Johnstone, T. (2001) “The effect of emotion on voice production and speech acoustics,” PhD Thesis, the University of Western Australia.
- Scherer, K., Banse, R., Wallbott, H. (2001) “Emotion Inferences From Vocal Expression Correlate Across Languages and Cultures,” *Journal of Cross-Cultural Psychology* 32(1), 76-92.
- Tickle, A. (2000) “English and Japanese Speakers’ Emotion Vocalization and Recognition: A Comparison Highlighting Vowel Quality,” *ISCA Workshop on Speech and Emotion*.
- <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.409.4169> (2017年7月30日閲覧)
- エリクソン ドナ・昇地 崇明(2006)「性差、および母語が感情音声の知覚に与える影響—日本語、韓国語、英語母語話者を対象として—」音声文法研究会（編）『音声文法』東京:くろしお出版, 31-46.
- 重野純(2001)「音声に含まれた感情の認知」*青山心理学研究* 1, 37-43.
- 田川恭識(2001)「『平静の問いかけ』と『非難の問いかけ』の弁別に対する F0 パタンの影響」*聴覚研究会資料* 31(10), 39-55.
- 鶴真紀子・武田昌一(2007)「アナウンサーが発声した『怒り』の韻律的特徴と聴覚的印象の主成分分析を用いた比較」*久留美信愛女学院短期大学研究紀要* 30, 65-70.
- 中林律子(2011)「日本語音声に表れる感情の知覚—ロシア人日本語学習者の知覚の個人差に着目して—」*音声研究* 15(3), 14-25.
- 西端千香子(1996)「感情・態度を持つ音声の知覚・表出訓練のための CAI 用テキストの作成—促音が含まれる発話を中心に—」感情・態度を表す日本語音声の表出診断・訓練プログラムの構築に関する研究, 平成7年度科学研究費補助金, 一般研究(B)研究成果報告書, 21-40.
- 劉艳(2011)《普通话的情感语音韵律分析》, 南京师范大学硕士学位论文.

日本語母語話者による英語音声の 音節とストレス位置知覚に及ぼす諸要因

江口 小夜子（国際電気通信基礎技術研究所(ATR)，神戸大学大学院）

山田 玲子（国際電気通信基礎技術研究所(ATR)，神戸大学）

1. はじめに

日本語母語話者は英語音声の知覚において、英語と日本語の音声学的な違いから韻律の知覚が困難な場合があり、様々な観点からの研究が報告されている。例えば、Beckman (1986) や江口 (2015) は、英単語音声のストレス位置を判断させた結果、英語母語話者は、音の高さ、長さ、強さ、母音音質の4つの音響的特徴を手がかりとして区別するのに対し、日本語母語話者は、主に高さを手がかりにすることを示した。また、Tajima and Akahane-Yamada (2004)は、日本語母語話者に英単語の音声を聴覚呈示し、単語の音節数を数えさせた結果、音節数のカウントが困難であることを報告した。一方、韻律知覚には、音節数、子音の数、位置、その他様々な要因が影響している可能性があり、例えば Tajima and Akahane-Yamada (2004)では、音節構造の複雑さや語内の子音の数に影響して正答率が下がることが示されている。しかし、各要因の影響の度合いや差異は明らかになっていない。そこで本研究では、日本語母語話者を対象に、音節およびストレス位置の知覚実験を行い、子音構成、重子音の位置、音節数、ストレスの位置、母音構成、黙字、音節主音の子音の影響を系統的に比較した。本稿では、調査した要因のうち、影響の度合いが大きかった子音構成、重子音の位置、音節数、ストレスの位置を中心に報告する。

2. 方法

2.1. 実験参加者

日本語母語話者 17名(男性 11名，女性 6名；19歳～40歳，平均年齢 24歳)を実験参加者とした。実験参加者の TOEIC スコアは 330 点から 940 点に分布しており、英語習熟度の幅は広がった。アンケートにより全員 1 年以上の海外滞在経験がないこと、聴力や言葉の障害がないことを確認した。

2.2. 課題

Syllable Count Task と Stress Identification Task の 2 種類の課題を用いた。Syllable Count Task では、聞こえた英単語の音節数を 1～12 の選択肢ボタンから回答した。Stress Identification Task では、ストレスの位置を英単語の綴り（アルファベット）上をマウスクリックして回答した。刺激の呈示、反応の取得は PC 上の実験プログラムで制御した。刺激音の呈示はヘッドフォンを用いた。実験参加者にはできるだけ速く回答するように教示した。

2.3. 刺激

①子音構成，②重子音の位置，③音節数，④ストレス位置，⑤母音構成，⑥黙字，

⑦音節主音的子音の要因について調べるため刺激語を選定し、6つの刺激セット(合計1008語)を作成した(表1)。これらの語をアメリカ英語母語話者4名(男性2名, 女性2名)が発音したものを刺激音とした。Syllable Count Taskでは, 1008語(刺激セット1~6)を, Stress Identification Taskでは, 717語(刺激セット1, 3, 4)を使用した。それぞれランダムに3つのブロックに分けたため, ブロックA~Fの合計6ブロックとなった。表2にブロックと課題およびトライアル数の関係を示す。

表1. 刺激セットと調査した要因, 語数の関係

刺激セット	調査する要因	語数	SCT	SIT
1	①子音構成	360	✓	✓
2	②重子音の位置	63	✓	
3	③音節数	213	✓	✓
	④ストレス位置		✓	✓
4	⑤母音構成	144	✓	✓
5	⑥黙字	108	✓	
6	⑦音節主音的子音	120	✓	

表2. 各ブロックで使用した課題とトライアル数

ブロック	課題	トライアル数
A	Syllable Count	336
B	Syllable Count	336
C	Syllable Count	336
D	Stress Identification	239
E	Stress Identification	239
F	Stress Identification	239

※SCT = Syllable Count Task で使用した刺激セットに✓マーク

SIT = Stress Identification Task で使用した刺激セットに✓マーク

2.4. 手続き

6つのブロック (A~F) を実験被験者毎にランダムな順序で実施した。また, 各ブロックの問題呈示順序も実験参加者毎にランダムに出題した。

3. 結果

3.1. ①子音構成の影響

刺激セット1は4種類の子音構成(Syllable Complexity)と3種類の音節数(1 σ ~3 σ ; 本稿では, 刺激語を構成する音節数を σ で示す)を組み合わせた条件を満たす252語で構成された。子音構成条件は下記のとおり。

- ・ SC0 : 母音の前後に単独 (連鎖しない) 子音がある CVC, CVCVC, CVCVCVC
- ・ SC+1 : 2つの子音連鎖を含む CCVC, CVCC, CCVCVC, CVCVCC, etc.
- ・ SC+2 : 3つの子音連鎖を含む CCCVC, CVCCC, CCCVCVC, CVCVCCC, etc.

3.1.2. Syllable Count Task の結果

各実験参加者の刺激条件毎の正答率を従属変数とし, 子音構成(SC0, SC+1, SC+2)と音節数(1 σ ~3 σ)を要因とした2要因の分散分析を行った。その結果, 子音構成および音節数の主効果はなく, 交互作用[F(4,64)=2.763, p<.05]が有意だった(図1)。下位検定の結果も合わせてまとめると, 1音節語のみ, SC0, SC+1 > SC+2 の順に正答率が低かった。つまり, 3重子音の刺激に対して, 正答率が低くなることが示唆された。

次に, 回答と正解の差分(「回答した音節数」-「正解音節数」)を従属変数とし, 同様の分散分析を行った結果, 子音構成の主効果[F(2,32)=17.460, p<.001], および音節数の主効果[F(2,32)=22.578, p<.001]が有意だった(図2)。下位検定の結果を合わせてまとめると, 回答と正解の差分は, SC0 < SC+1 < SC+2 の順で, プラスの方向へ有意に

大きくなった。したがって、子音構成が複雑なほど、実際の音節数より多く数えることが示されたと言える。また、 $3\sigma < 2\sigma < 1\sigma$ の順で、プラスの方向へ有意に大きくなった。つまり、音節数が少ないほど、実際の音節数より多く数えることも示された。

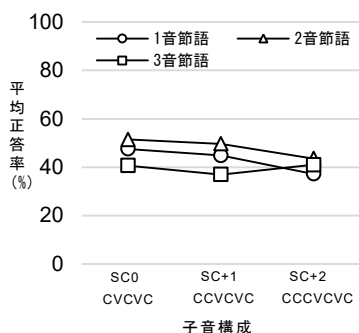


図 1: Syllable Count Task における刺激条件毎の正答率

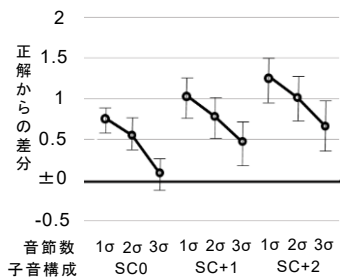


図 2: Syllable Count Task における正解と回答の差分

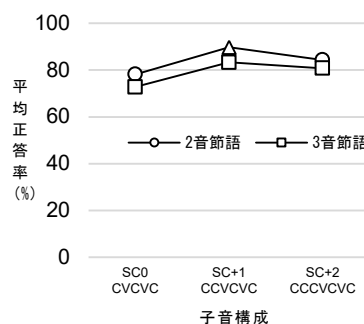


図 3: Stress Identification Task における刺激条件毎の正答率

3.1.3. Stress Identification Task の結果

各実験参加者の刺激条件毎の正答率を従属変数とし、子音構成(SC0, SC+1, SC+2)と音節数(2σ, 3σ)を要因とした 2 要因分散分析を行った。その結果、交互作用はなく、子音構成の主効果[F(2,32)=8.846, p<.001], 音節数の主効果[F(1,16)=5.226, p<.05]が有意だった(図 3)。下位検定の結果を合わせてまとめると、SC0 の正答率が他の条件より有意に低かった。つまり、子音構成が複雑になると正答率が高くなることが示唆された。

3.2. ②重子音の位置の影響

3.2.1. Syllable Count Task の結果

刺激セット 1, 2 から、2 種類の重子音(2 重子音, 3 重子音), その出現位置(語頭, 語末)および 3 種類の音節数(1σ~3σ)を組み合わせた条件を満たす 144 語を抽出して用いた。

3.2.1. Syllable Count Task の結果

各実験参加者の刺激条件毎の正答率を従属変数とし、重子音(2 重子音, 3 重子音), 位置(語頭, 語末), 音節数(1σ~3σ)を要因とした 3 要因の分散分析を行った。その結果、主効果および 2 次の交互作用はなく、単純交互作用が全て有意だった(重子音と位置[F(1,16)=7.696, p<.05], 重子音と音節数[F(2,32)=5.257, p<.05], 重子音の位置と音節数[F(2,32)=6.845, p<.005])。下位検定の結果を合わせてまとめると、1 音節語は、2 重子音, 3 重子音ともに語頭の正答率が語末よりも有意に低かったが、他の条件については統一的な傾向は認められなかった。

次に、回答と正解の差分(「回答した音節数」-「正解音節数」)を従属変数とし、分散分析を行った結果、主効果, 単純交互作用, 2 次の交互作用が全て有意だった(重子音[F(1,16)=19.042, p<.001], 位置[F(1,16)=26.862, p<.001], 音節数[F(2,32)=37.996, p<.001], 重子音と位置[F(1,16)=9.495, p<.01], 重子音と音節数[F(2,32)=4.048, p<.05], 重子音と位置と音節数[F(2,32)=13.182, p<.001])(図 4, 5)。下位検定の結果を合わせてまとめると、正解からの差分は、2 重子音と 3 重子音ともに、語頭 > 語末の順、また、 $1\sigma > 2\sigma > 3\sigma$ の順で低かった。したがって、重子音は語末よりも語頭にある場合、および音節数が

少ない場合に、実際の音節数より多く数えることが示唆された。

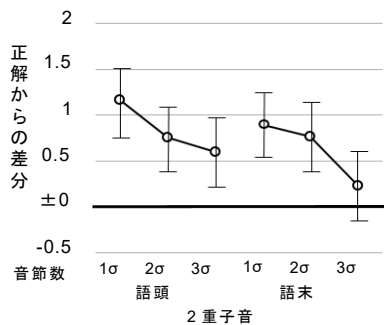


図 4: Syllable Count Task における 2 重子音の正解と回答の差分

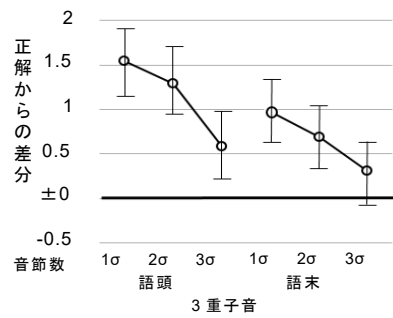


図 5: Syllable Count Task における 3 重子音の正解と回答の差分

3.3. ③音節数の影響

刺激セット 3 は音節数(1σ~6σ)の条件を満たす 213 語で構成された。

3.3.1. Syllable Count Task の結果

各実験参加者の刺激条件毎の正答率を従属変数とした F 検定を行ったところ、音節数(1σ~6σ)の効果は有意だった[F(5,80)=6.568, p<.001] (図 6)。多重比較の結果をまとめると、2σ > 1σ, 3σ > 4σ > 5σ > 6σ の順で正答率が低くなった。したがって、1 音節語以外の条件では、音節数が増えるにつれ、正答率が低くなることが示唆された。

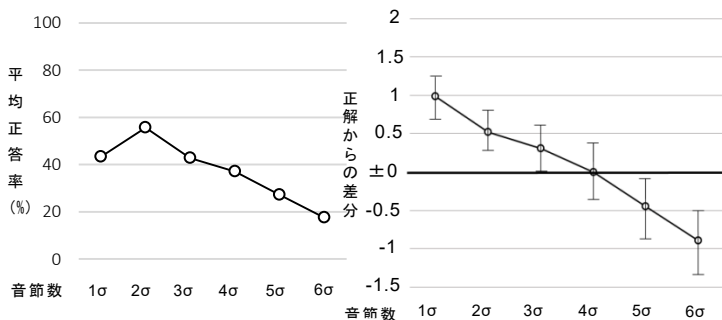


図 6: Syllable Count Task における刺激条件毎の正答率

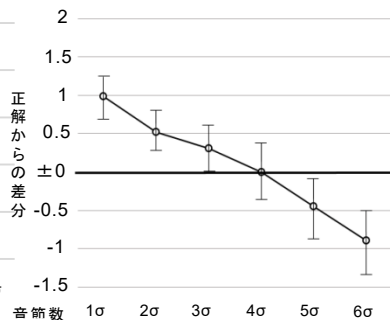


図 7: Syllable Count Task における正解と回答の差分

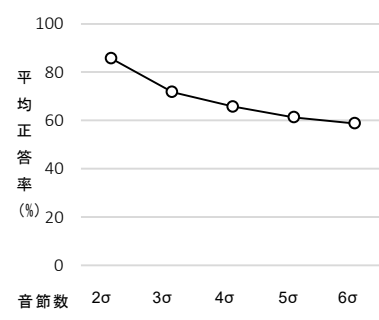


図 8: Syllable Count Task における刺激条件毎の正答率

次に、回答と正解の差分（「回答した音節数」－「正解音節数」）を従属変数とした F 検定を行った結果、音節数の効果は有意だった[F(5,80)=39.994, p<.001] (図 7)。多重比較の結果をまとめると、6σ < 5σ < 4σ < 3σ < 2σ < 1σ の順で、正解からの差分はプラスの方向に大きくなった。また、1σ~3σの差分はプラスであったが、5σ、6σの差分はマイナスであった。つまり、音節数が少ない1~3音節語では、実際の音節数より多く数えるのに対し、音節数の多い5~6音節語の音節数は実際の音節数より少なく数えることが示唆された。

3.3.2. Stress Identification Task の結果

各実験参加者の刺激条件毎の正答率を従属変数とし、音節数(2σ~6σ)を要因とした F 検定を行った。その結果、音節数の効果が有意だった[F(4,64)=30.523, p<.001] (図 8)。多重比較の結果をまとめると、2σ > 3σ > 4σ > 5σ > 6σ の順で正答率が低くなった。

したがって、音節数が増えるにつれ、正答率が低くなることが示唆された。

3.4. ④ストレス位置の影響

刺激セット 3 から、2～6 音節の 201 語に対する結果を分析した。刺激条件はストレスの位置であった。音節数ごとに刺激条件の水準数は異なっており、2 音節語のストレス位置条件は、1 音節目と 2 音節目の 2 水準、3 音節語のストレス位置条件は 1 音節目、2 音節目、3 音節目の 3 水準であった。6 音節語のみ、1 音節目と 6 音節目にストレスがある単語の数が不十分であったため、それらを除外した 4 水準であった。

3.4.1. Syllable Count Task の結果

各実験参加者の刺激条件毎の正答率を従属変数とし、刺激条件(2σ～6σ)ごとにストレス位置を要因とした F 検定を行ったが、有意差は認められなかったことから、ストレス位置の影響がないことが示唆された。

次に、回答と正解の差分(「回答した音節数」－「正解音節数」)を従属変数として、刺激条件(2σ～6σ)ごとにストレス位置を要因とした F 検定を行った結果、2, 4, 5, 6 音節語においてストレス位置の効果が有意であった(2σ[F(1,16)=9.781, p<.01]; 4σ[F(3,48)=9.127, p<.001]; 5σ[F(4,64)=9.613, p<.001]; 6σ[F(3,48)=3.994, p<.05])。多重比較の結果をまとめると、ストレスの位置が語末位置に近づくほど、正解からの差分がプラスの方向へ大きくなることが示された。

3.4.2. Stress Identification Task の結果

各実験参加者の刺激条件毎の正答率を従属変数とし、刺激条件(2σ～6σ)ごとにストレス位置を要因とした F 検定を行った。その結果、全ての条件においてストレス位置の効果が有意だった(2σ[F(1,16)=11.620, p<.005]; 3σ[F(2,32)=26.847, p<.001]; 4σ[F(3,48)=26.176, p<.001]; 5σ[F(4,64)=19.087, p<.001]; 6σ[F(3,48)=4.087, p<.05])(図 9)。多重比較の結果をまとめると、ストレスの位置が語末位置に近づくほど、正答率が有意に低くなることが示唆された。

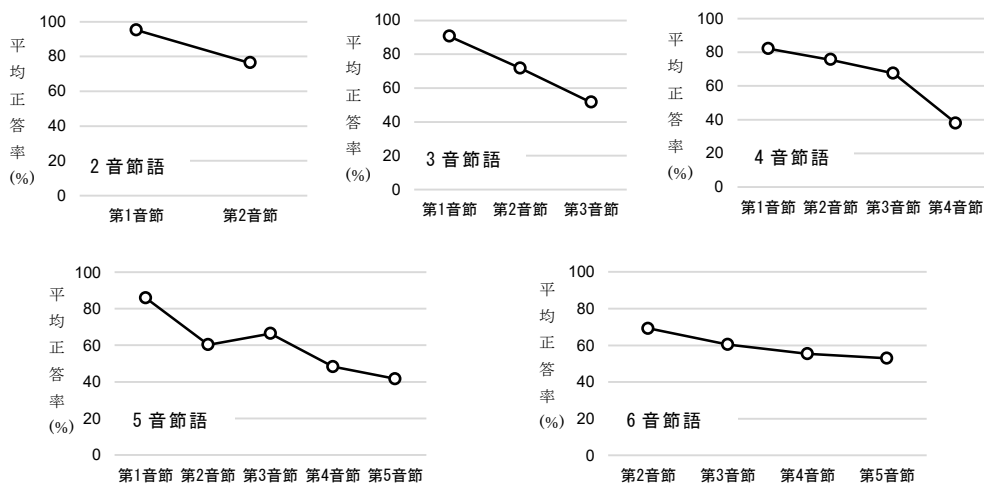


図 9. Stress Identification Task における
2 - 6 音節語の刺激条件毎の正答率

4. 考察

日本語母語話者の全刺激音に対する正答率は、音節知覚では 39.3%，ストレス位置知覚では 69.5%とともに低かった。課題が異なるため両知覚の難易度の直接の比較はできないものの、日本語母語話者は、ストレス知覚のみならず音節数のカウントも困難であることが示唆された。

表 3. 音節知覚およびストレス位置知覚における各要因の影響

	音節知覚	ストレス位置知覚
子音構成	1 音節語の 3 重子音で困難	子音構成が複雑な場合に正答率が高い
重子音の位置	1 音節語の 2 重子音, 3 重子音において, 重子音が語頭にある場合に困難	
音節数	音節数が多くなるほど困難	音節数が多くなるほど困難
ストレス位置	影響なし	ストレス位置が語末位置に近づくほど困難

各要因の影響について、まとめた結果を表 3 に示す。音節知覚では、日本語母語話者は、音節数が多くなるほど、音節知覚が困難になることが分かった。また、1 音節語の刺激において、子音構成および重子音の位置の影響が認められたが、その他の条件では影響がなかったことから、子音構成、重子音の位置の影響の度合いは比較的小さいと考えられる。次に、ストレス位置知覚では、母音の前後に単独子音がある条件の正答率は低く、2 重子音、3 重子音など子音構成が複雑な場合に正答率が高くなることが示された。また、音節数が多くなるほど、およびストレス位置が語末位置に近いほど知覚が困難になることが分かった。子音構成やストレス位置が影響する理由について明らかにするためには、今後さらなる検討が必要である。正解との差分のデータからは、調査した要因全てにおいて、音節数が少ないほど、実際の音節数より多く数えることが分かった。

以上の結果から、総じて、今回対象とした要因は複雑に音節知覚およびストレス位置知覚に影響していることが示された。今後は、本研究の結果から明らかになった要因に焦点をあて、韻律の学習方法について検討したい。

参考文献

- Beckman, M. E. (1986) *Stress and non-stress accent*. Dordrecht: Foris.
- 江口小夜子 (2015) 「英語音声のストレス知覚における母音音質の影響 -英語母語話者と日本語母語話者の比較-」『第 29 回日本音声学全国大会予稿集』 68-72.
- Tajima, K. and Akahane-Yamada, R. (2004) “Production and perception of syllable structure in second-language speech.” *The 18th International Congress on Acoustics, Proc. ICA 2004, IV*, 3321-3324.

中国語母語話者による日本語特殊拍の知覚 —特殊拍間の混同—

張 林姝（神戸大学大学院） 林 良子（神戸大学） 山田 玲子（国際電気通信基礎技術研究所<ATR>・神戸大学）

1. はじめに

日本語特殊拍に関する研究は多数報告されているが（皆川他 2002, 戸田 2003, Kubozono et al. 2011 等）、それらのほとんどは特殊拍と自立拍の混同を対象にしたものである。しかし、特殊拍に分類される 3 つの音素である長音/R/、撥音/N/、促音/Q/の間、即ち特殊拍間でも混同が起こる可能性がある。

岡田（2003）は、生成における/R/と/N/の混同について報告した。上野（2014）は、母語話者が/N/を/R/に（例えば、「原因」を「ゲーイン」に）発話する場合があると述べている。知覚に関して、黒崎（2002）は、「テーイン」を単独で聞くと正しく「テーイン」に聞き取る確率は 60%だが、「コンビニのテーインをしている」という文では、「テンイン」に聞き取る確率が高くなり、「テーイン」と聞き取る確率は 0%になると報告した。また、本橋（2005）は英語母語話者が促音を長音に聞き取る傾向にあることを示した。

しかし、これらの特殊拍間の混同を取り上げた先行研究では、限られた対の混同しか扱っておらず、使用された刺激語も限られている。本研究では、特殊拍の混同を系統的に検討するために、/R/、/N/、/Q/、自立拍<I>^注の相互の関係について日本語母語話者と中国人日本語学習者を対象に 2 つの知覚実験を通して検討した。実験 1 では、特殊拍で対立する無意味語 2 つの同-異判断実験を行った。実験 2 では、文内の、特殊拍が含まれる無意味語を対象に、同定実験を行った。

2. 実験 1

2.1. 方法

2.1.1. 刺激音

4 種類の拍（R/N/Q/I）で対立する 3 音節 4 拍の無意味語/ta?paka/、/tapa?ka/、/tapaka?/（?=R/N/Q/またはなし）を用いた。アクセント型 4 種類（平板/頭高/中高 I/中高 II）と特殊拍の生起位置 3 種類（第 1 音節/第 2 音節/第 3 音節）を組み合わせた 12 条件のうち、「ターパカ/タンパカ/タツパカ/タパカ」など 4 種類の拍の対立を実現できる 6 セットの 24 語を素材とした（表 1）。ただし、自立拍の語はセット間で重複があり、重複を除くと 21 語だった。

東京方言母語話者（女性、40 代）が防音室でこれらの語をそれぞれ 5 回読み上げたものを、IC レコーダーを用いて、サンプリング周波数 44.1kHz、量子化精度 16bit の音声ファイルとして保存した。各語 5 回の発話から時間長が中央値を示す発話を刺激音に使用した。刺激語の長さは、2 音節語では平均 365.95msec、3 音節語では平均 534.66msec だった。

同一セット中の異なる 2 語を組み合わせた 6 対（例えば、ターパカ-タンパカ/タンパカ-タツパカ/タツパカ-ターパカ/タパカ-ターパカ/タパカ-タンパカ/タパカ-タツパカ）を 2 種

注：本文では自立拍を<I>（Independent morae）と記す。

類の順序 (AB/BA) で連結し 1 セットから 12 個 (6 対×2 種類の順序) の刺激音を作成した。刺激語と刺激語の間には 1000msec の無音区間を挿入した。6 セット、合計 72 個 (12 個×6 セット) の刺激音に加え、ダミーとして同じ語をつなげた刺激語対の刺激音を作成した。

2.1.2. 実験参加者

日本語母語話者 15 名 (20~40 才) と中国人日本語学習者 15 名 (20 才) が参加した。全員健聴者で、神戸大学の大学生と大学院生であった。中国人日本語学習者は全員日本語能力試験 N1 の合格者だった。

2.1.3. 手続き

静かな部屋でヘッドホンを通して刺激音をランダムに呈示した。実験参加者には、ヘッドホンから聞こえてくる対となる 2 つの単語が同じか否かを判断し、できる限り早くキーボードの割り当てられたキーを押して回答するように教示した。刺激の呈示と反応の取得は SuperLab (ver.5) を用いた。

表 1 実験 1 の刺激語

分析に用いた語を太枠で示した。日本語の音韻規則に沿わない語は背景に色を付けて示した。

拍の種別	平板型			頭高型			中高型 (核が 2 拍目)			中高型 (核が 3 拍目)		
	1*	2	3	1	2	3	1	2	3	1	2	3
R	ターバ カ	タバ ーカ	タバカ ー	ター バカ	タバ ーカ	タバ カー	ター バカ	タバ ーカ	タバ カー	ター バカ	タバ ーカ	タバ カー
N	タンバ カ	タバ ンカ	タバカ ン	タン バカ	タバ ンカ	タバ カン	タン バカ	タバ ンカ	タバ カン	タン バカ	タバ ンカ	タバ カン
Q	タツバ カ	タバ ツカ	タバカ ツ	タツ バカ	タバ ーカ	タバ カツ	タツ バカ	タバ ツカ	タバ カツ	タツ バカ	タバ ツカ	タバ カツ
I	タバカ カ	タバ カ	タバカ カ	タバ カ	タバ カ	タバ カ	タバ カ	タバ カ	タバ カ	タバ カ	タバ カ	タバ カ

*特殊拍が何音節目にあるかを示す。

2.2 結果

正答率は、日本語母語話者は 98.3%、中国人日本語学習者は 96.9% と高く、グループ間の差がなかった。そこで反応時間を従属変数とし、母語 (日本語母語話者/中国人日本語学習者)、アクセント型 (平板型/中高型/頭高型)、対の種別 (「特殊拍-特殊拍」対/「特殊拍-自立拍」対) を要因とした 3 元配置分散分析を行った。その結果、母語とアクセント型の交互作用 [F(2, 116)=4.16, p<.05]、アクセント型と対の種別の交互作用 [F(2, 116)=18.34, p<.001]、母語の主効果 [F(1, 58)=18.32, p<.05] が有意だった。中国人日本語学習者は日本語母語話者より反応時間が長く、下位検定の結果、日本語母語話者で平板型と中高型に差があったが、中国人日本語学習者はアクセント型による差がなかった (図 1)。また、「特殊拍-特殊拍」対と「特殊拍-自立拍」対の差が中高型のみで現れた。以下、アクセント型ごとに、3 元配置分散分析 (母語・生起位置・対の種別) を行った。

平板型では、3 つの要因の交互作用がなく、母語の主効果のみ有意で [F(1, 28)=9.74, p<.005]、中国人日本語学習者の反応時間は日本語母語話者より長かった (図 2)。対の種別について、交互作用、主効果ともになかったことから、特殊拍同士の弁別の反応時間は、特殊拍と自立拍の弁別の場合と差がなかったといえる。頭高型では、母語と対の種

別の交互作用 [F(1, 28)=5.86, p<.05]、母語の主効果[F(2, 116)=4.16, p<.05]、生起位置の主効果[F(1, 28)=11.30, p<.005]が有意だった (図 3)。日本語母語話者の「特殊拍-特殊拍」対に対する反応時間は「特殊拍-自立拍」対より長かったが、中国人日本語学習者では、差がなかった。中国人日本語学習者の反応時間は拍の種別にかかわらず、日本語母語話者より長かった。特殊拍が第 2 音節にある際の反応時間は第 1 音節の場合より長かった。

頭高型では、母語と対の種別の交互作用 [F(1, 28)=5.86, p<.05]、母語の主効果[F(2, 116)=4.16, p<.05]、生起位置の主効果[F(1, 28)=11.30, p<.005]が有意だった (図 3)。日本語母語話者の「特殊拍-特殊拍」対に対する反応時間は「特殊拍-自立拍」対より長かったが、中国人日本語学習者では、差がなかった。中国人日本語学習者の反応時間は拍の種別にかかわらず、日本語母語話者より長かった。特殊拍が第 2 音節にある際の反応時間は第 1 音節の場合より長かった。

中高型では、生起位置の要因とアクセント核の位置の要因が共起しているが、交互作用はどの要因の間にもなく、母語 [F(1, 28)=7.70, p<.01]、生起位置 [F(1, 28)=8.12, p<.01]、対の種別 [F(1, 28)=24.93, p<.001]、すべての主効果が有意だった (図 4)。中国人日本語学習者の反応時間が日本語母語話者より長かった。第 2 音節の反応時間が第 1 音節より長かった。また、「特殊拍-自立拍」対の反応時間が「特殊拍-特殊拍」より長かった。

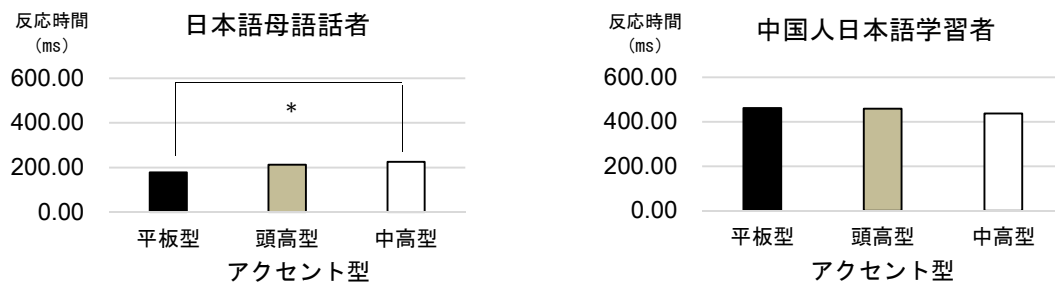


図 1 実験 1: 弁別課題におけるアクセント型と反応時間の関係。

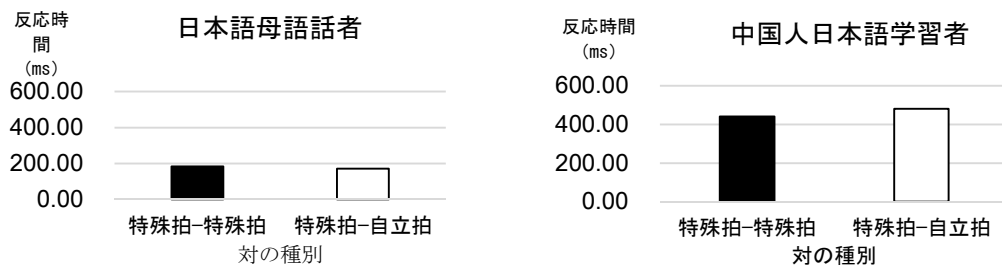


図 2 実験 1: 平板型における対の種別と反応時間の関係。

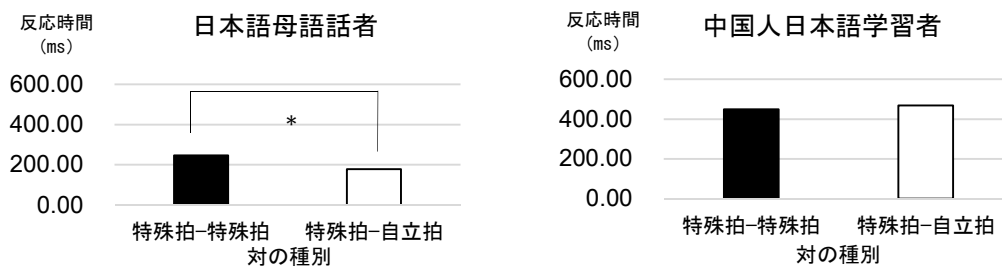


図 3 実験 1: 頭高型における対の種別と反応時間の関係。

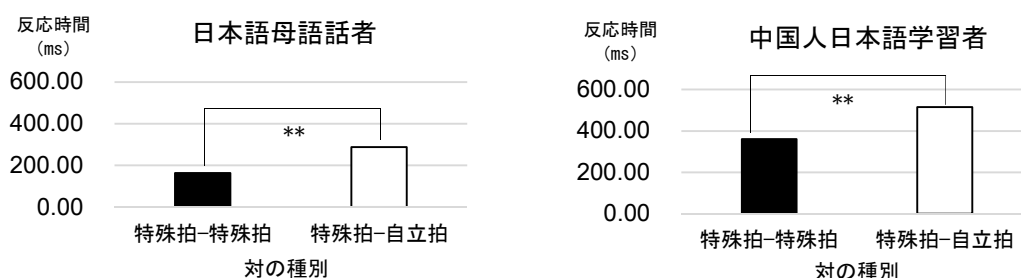


図4 実験1:中高型における対の種別と反応時間の関係

以上のすべての分析で、母語の主効果が有意だったことから、拍で対立する2語の弁別の反応時間は、中国人日本語学習者は日本語母語話者より長いことが示された。次に、拍の生起位置の主効果が有意だったことから、特殊拍が第2音節に生起する場合は第1音節に生起する場合より長いことが分かった。また、対の種別については、反応時間が平板型では「特殊拍-特殊拍」＝「特殊拍-自立拍」、頭高型では「特殊拍-特殊拍」>「特殊拍-自立拍」、中高型では「特殊拍-特殊拍」<「特殊拍-自立拍」だったことから、「特殊拍-特殊拍」は「特殊拍-自立拍」より知覚困難である可能性が示唆された。

3. 実験2

3.1. 方法

3.1.1. 刺激音

拍4種類(R/N/Q/I)を1セットにし、3音節4拍の無意味語を作成した。特殊拍の前後に必ず/ka/と/ta/を入れ、特殊拍が第1(語頭)、2(語中)、3(語尾)のいずれかの音節に現れるようにした。アクセント型は全て平板型とした。刺激語は合計11語となった(表2)。

これらの刺激語をキャリア文の中に入れ、刺激文とした。キャリア文は、刺激語が外来語と解釈されるもの、例えば「去年オーストラリアで新型ロボット _____が開発されました。」を、5種類作成した。1つの語を2種類のキャリア文に入れたので、刺激文は合計22文となり、ダミー語の入ったダミー文16文を加え、合計38文を用いた。全38文を東京母語話者1名(女性、40代)が防音室で読みあげたものを、ICレコーダーを用いて録音した。なるべく自然に発話するように教示した。サンプリング周波数44.1kHz、16bitの精度でリニアwavファイルとして保存した。

3.1.2. 実験参加者

日本語母語話者15名(20~40才)と中国人日本語学習者15名(20才)が参加した。全員健聴者で、神戸大学の大学生と大学院生であった。中国人日本語学習者は全員日本語能力試験N1の合格者だった。

3.1.3. 手続き

実験参加者は静かな部屋で、ヘッドホンを通して刺激文を聴取し、刺激文の再生終了3秒後に画面上に呈示される4つの選択肢から、聞いた刺激文の中に入っている無意味語と同じものを選び、なるべく速くそれに対応するキーを押すように教示した。選択肢は刺激語とそれと対立する3語だった。刺激、選択肢呈示は、SuperLab(Ver.5)を用いた。

表 2 実験 2 の刺激語

日本語の音韻規則に沿わない語は背景に色を付けて示したが、それらは刺激語に使わなかった。

拍の種別	第 1 音節(語頭)	第 2 音節(語中)	第 3 音節(語尾)
R	カータパ	パカータ	パタカー
N	カンタパ	パカンタ	パタカン
Q	カッタパ	パカッタ	パタカッ
I	カタパ	パカタ	パタカ

3.2. 結果

正答率を従属変数とし、母語と生起位置を要因とした分散分析を行なった結果、交互作用が有意で [F(3, 168)=5.375, p<.005]、中国人日本語学習者のみ、正答率が生起位置によって異なった (図 5)。生起位置ごとに、母語 (日本語母語話者/中国人日本語学習者) と拍の種別 (R/N/Q/I) を要因に配置分散分析した結果を以下に記す。

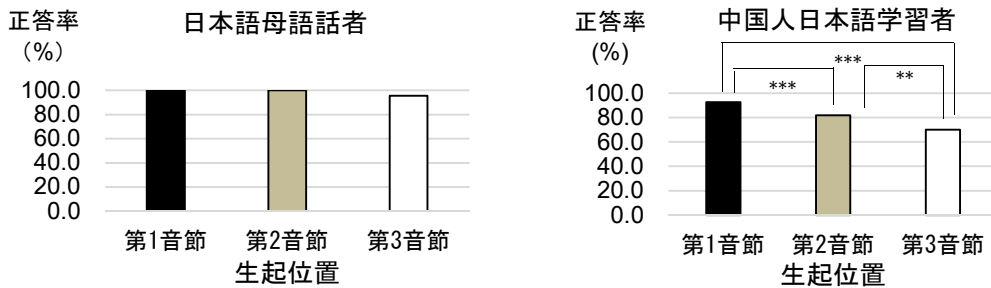


図 5 実験 2: 同定課題における特殊拍の生起位置と正答率の関係。

第 1 音節では、交互作用がなく、母語の主効果のみ有意だった[F(1, 28)=4.295, p<.05]。中国人日本語学習者の正答率は日本語母語話者より有意に低かった。第 2 音節では、母語と拍の種別の交互作用 [F(3, 84)=3.961, p<.05]が有意だった。日本語母語話者は拍の種別に関わらず、ほぼ 100%の正答率を示したが、中国人日本語学習者は拍の種別による正答率の変動が大きく、/R/における正答率は他より低かった。また、日本語母語話者と中国人日本語学習者の差が/R/と/N/のみにあった。第 3 音節では、母語と拍の種別の交互作用[F(2, 116)=4.08, p<.05]が有意だった。日本語母語話者は拍の種別に関わらず、ほぼ 100%の正答率を示した。中国人学習者は拍の種別による正答率の変動が大きく、特殊拍の正答率は自立拍より低かった。日本語母語話者との差が特殊拍のみにあった。

回答パターンは図 6~8 に示す。誤答パターンを見ると、第 1 音節で、日本語母語話者も中国人日本語学習者も特殊拍間の混同はほぼなかった (図 6)。第 2 音節で、日本語母語話者には混同がなかったが、中国人日本語学習者には/R/と/Q/の互いの混同があった (図 7)。第 3 音節で、日本語母語話者には混同がなかったが、中国人日本語学習者には/N/から/R/への混同があった (図 8)。

日本語母語話者の同定率 (%)

刺激語 \ 選択肢	R	N	Q	I
R	100.0	0.0	0.0	0.0
N	0.0	100.0	0.0	0.0
Q	0.0	0.0	100.0	0.0
I	0.0	0.0	0.0	100.0

中国人日本語学習者の同定率 (%)

刺激語 \ 選択肢	R	N	Q	I
R	96.7	0.0	3.3	13.3
N	0.0	96.7	0.0	0.0
Q	0.0	0.0	90.0	0.0
I	3.3	3.3	6.7	86.7

図 6 実験 2: 第 1 音節における拍の種別と回答パターンの関係。

刺激語 選択肢	R	N	Q	I
R	100.0	0.0	0.0	0.0
N	0.0	100.0	0.0	0.0
Q	0.0	0.0	100.0	0.0
I	0.0	0.0	0.0	100.0

刺激語 選択肢	R	N	Q	I
R	63.3	3.3	10.0	3.3
N	3.3	90.0	3.3	3.3
Q	20.0	0.0	83.3	0.0
I	13.3	6.7	3.3	93.3

図7 実験2:第2音節における拍の種別と回答パターンとの関係。

刺激語 選択肢	R	N	I
R	100.0	3.3	3.3
N	0.0	96.7	6.7
I	0.0	0.0	90.0

刺激語 選択肢	R	N	I
R	80.0	16.7	23.3
N	0.0	53.3	0.0
I	20.0	30.0	76.7

図8 実験2:第3音節における拍の種別と回答パターンとの関係。

以上の結果から、拍の同定において、中国人日本語学習者は語尾に行くほど、特殊拍同定の正答率が下がり、特殊拍間の混同も起こることが分かった。

4. 考察

実験1の結果からは、弁別の際に「特殊拍-特殊拍」のほうが「特殊拍-自立拍」より反応時間が長い場合があることが示された。実験2の結果からは、同定時に、RとQを相互に混同する、NをRと混同するなど、特殊拍間で混同される場合があることが示された。これらの結果から、中国語母語話者は特殊拍同士も混同していることが明らかになり、日本語の学習の過程で「特殊拍-自立拍」の混同のみでなく、特殊拍同士の混同を解決するための訓練も行う必要性が示唆された。より良い学習方法を提案するためには、今後、生成の面からも考察する必要がある。

参考文献

- 岡田祥平 (2003) 「撥音から長音への「言い間違い」現象について:『日本語話し言葉コーパス』を資料として (第17回全国大会発表要旨)」『音声研究』4:3, 117.
- 上野善道 (2014) 「フンイキ>フィンキの変化から音位転換について考える」『生活語の世界』8-19.
- 黒崎典子 (2002) 「母音に前接する撥音について:日本語母語話者にとっての知覚の難易」『神奈川大学言語研究』25, 11-22.
- 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』7:2, 70-83.
- 皆川泰代・前川喜久雄・桐谷滋 (2002) 「日本語学習者の長/短母音の識別におけるピッチ型と音節位置の効果」『音声研究』6:2, 88-97.
- 本橋美樹 (2005) 「英語話者による促音の認識」『言語文化と日本語教育』30, 95-98.
- Kubozono, H., & Takeyasu, H., & Giriko, K., & Hirayama, M. (2011) "Pitch cues to the perception of consonant length in Japanese." *International Congress of Phonetic Sciences August2011*. (pp, 17-266).

後続子音の調音方法の日本語話者による英語の母音の同定と弁別への影響

野澤 健 (立命館大学)

t-nozawa@ec.ritsumei.ac.jp

1. はじめに

一般に、母語以外の言語の母音の知覚実験を行う場合、前後の子音の影響を最小限にするため/hVt/のような構造が用いられる。また、子音の影響を考慮する場合、先行する子音の調音位置の影響が中心であった(Nozawa & Wayland 2012 など)。

このような現状を踏まえて、母音の音響的特性に大きな影響を与えることは知られていても、それが非母語話者の母音の知覚にどのように影響するかはあまり研究されて来なかった後続子音の調音方法に焦点を当て、母音の知覚へのその影響について述べていく。具体的には、鼻音/n/と接近音/l/の影響である。鼻音は母音の F1 領域に鼻音のフォルマントを起こすため、母音の高さに関する情報が影響を受ける可能性がある。英語の母語話者を対象にした実験では、後続の鼻音により鼻音化する文脈(/bVnd/)で鼻音化した母音の高さの知覚は影響されなかったが、本来鼻音化しない文脈での鼻音化した母音/bVd/は、対応する口腔母音よりも低く知覚されるという報告がある(Krakow et al. 1988)が、アメリカ英語の一部の方言で鼻音の前で/e/-/i/中和する(pen/pin merger)こと(Labov et al. 2005)や音素として鼻母音を持つ言語では鼻母音の音素数は口腔母音の音素数よりも少ないこと(Beddor, 1993)などからも鼻音は母音を不明瞭にすると考えられる。また、アメリカ英語固有の現象として、鼻音の前で/æ/が緊張母音化(tensing)し、[eə]のような響きを持つことが挙げられる。図 1 に同じ話者の発話による pat と pan の波形とスペクトラムを示す。Pan の方が F2 が大きく曲がっているのがわかる。

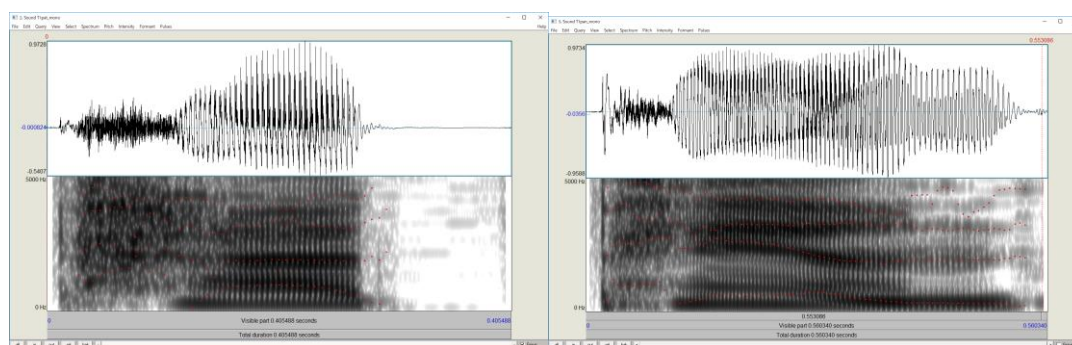


図 1; アメリカ英語母語話者の発話による pat と pan の波形とスペクトログラム

一方、母音に後続する/l/は、先行する母音の F2 を継続的に降下させることが知られている(Olive et al. 1993)。図 2 に同じ話者の発話による Pete と peel の波形とスペクトラムを示す。Peel では、F2 が連続して下降しているのがわかる。

このように/n/や/l/の前では、母音が閉鎖音の前とは異なる動きを見せるので、非母語話者である日本語話者の英語の母音の同定・弁別に大きな影響を及ぼすのではないかと考えられる。

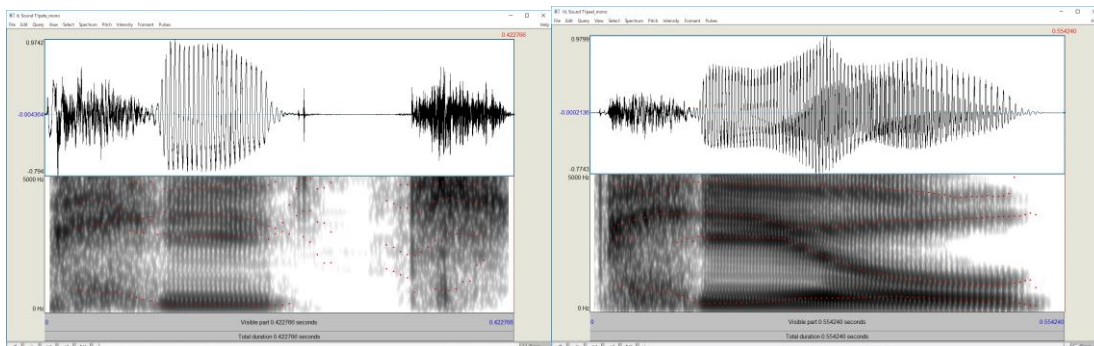


図2：アメリカ英語母語話者の発話による Pete と peel の波形とスペクトログラム

2. 実験

2.1. 音声刺激

音声刺激は、これまでの実験(Nozawa & Wayland 2012 等)で使用した音声及び同時期に録音した音源を使用した。4人のアメリカ英語の母語話者(女性)の発話をアラバマ州オーバーンでサンプル周波数 44.1KHz で録音した。4人の話者の内訳は、ニューヨーク州出身2名、カリフォルニア州出身1名、ウィスコンシン州出身1名である。これらの話者が /i, ɪ, e, æ, ɑ, ʌ/ の6母音を /hVt/, /pVt/, /pVn/, /pVl/ の子音環境の中で発したものを音声刺激として使用できるように編集したものを実験に使用した。6母音×4子音環境×4話者=96の音声刺激を使用することとなった。

2.2. 実験参加者

実験参加者は、日本の大学に通う日本語話者33名(34名が参加したが、最後まで参加できなかったもので、33名のデータについて述べる)である。内訳は男性5名、女性28名で、実験参加時の平均年齢は19.4(18-22)才である。英語力を示す指標としての TOEIC (R) の得点は、自己申告で330点から990点で平均628点(5名が受験経験がない)である。英語以外の外国語の学習歴があると回答したものが8名、海外での生活経験があるものが10名、うち1名は海外(マレーシア)生まれであった。全員、聴覚は正常である。

2.3. 実験方法

実験参加者は、以下の3つの実験に参加した。実験用の音声はすべてコンピュータで再生され、実験参加者はヘッドフォンを通じて音声を聴いた。

2.3.1 知覚同化実験

実験参加者は、1回の試行に1つの音声刺激を聴き、音声刺激の母音に最も近い日本語の母音を解答用紙の選択肢の中から選び、選んだ日本語の母音としての良し悪しを7段階(1=良くない、7=良)で評価した。選択肢は、/p/で始まる子音環境のものは「パ」、「ピ」のよう

にパ行音のカタカナで表示し、/hVt/についてはハ行音のカタカナで表示した。子音環境ごとに分けて、音声は提示し、実験参加者は/pVt/では/p/で始まり/t/で終わる語を聴くが最後の/t/は無視して、その直前までの音を日本語で表記するとどれが一番近いかを選ぶように指示を受けた。各音声刺激は順番を変え 2 回提示された。こうして、6 母音×4 話者×2 回=48 回の試行が各子音環境ごとに作成された。実験参加者が解答を終えたのを確認して、実験参加者がマウスをクリックして次の試行に移った。ITIは 1,000ms に設定した。

2.3.2 同定実験

実験参加者は、1 回の試行に 1 つの音声刺激を聴き、聞こえた音声を表す選択肢にカーソルを動かし、解答した。選択肢は各子音環境の中で 6 母音で、表 1 に示しているように綴り字を用いた。この中には実験参加者には馴染みのない語や実在しない語が含まれるが、実験参加者には語の意味に左右されずに聞こえたと選択肢を選ぶように指示した。実在しない語については、het は head と同じ母音を持つ語で語尾が d ではなく t になったものであるとか、pul は pulse から s の音を除いたもので、pull とは異なるなどの説明を加えた。各音声刺激は、順番を変え 2 回提示された。こうして、6 母音×4 話者×2 回=48 回の試行が各子音環境ごとに作成された。これで各母音の同定能力は 8 つの試行で測定されることとなった(4 話者×2 回)。ITIは、1,000ms に設定した。また、10 秒間解答しないともう一度同じ音声刺激を聴くことが許された。4 つの子音環境の提示順は参加者ごとに変えられた。

表 1: 同定実験に用いた選択肢一覧

	/i/	/ɪ/	/ɛ/	/æ/	/ɑ/	/ʌ/
hVt	heat	hit	het	hat	hot	hut
pvt	Pete	pit	pet	pat	pot	putt
pVn	peen	pin	pen	pan	pon	pun
pVl	peel	pill	pell	pal	pol	pul

2.3.3 弁別実験

AXB 方式を採用した。実験参加者は、1 回の試行に 3 つの音声刺激を聴き、2 番目の音声刺激が最初の音声刺激と最後の音声刺激のどちらと同じカテゴリーの母音を含むか、つまり AAB か ABB かを選択することにより解答した。選択肢は 2 つ用意されていて、AAB の場合は First、ABB の場合は Last を選択するように指示された。6 つの母音のペア、/i/-/ɪ/、/ɛ/-/ɪ/、/æ/-/ɛ/、/æ/-/ɑ/、/æ/-/ʌ/が設定された。それぞれの子音環境で各母音のペアの弁別能力は 12 の試行で測定された。例えば/i/-/ɪ/の場合、/i/-/i/-/ɪ/の試行が 3 つ、/i/-/ɪ/-/ɪ/の試行が 3 つ、/ɪ/-/i/-/i/の試行が 3 つ、/ɪ/-/ɪ/-/i/の試行が 3 つの計 12 の試行で弁別能力が測定された。各試行の 3 つの音声刺激はすべて異なる話者の発話によるもので構成された。ISI、ITIともに 1,000ms に設定した。同定実験同様 10 秒間解答しないと同一試行を再度聴くことが許された。4 つの子音環境の提示順は参加者ごとに変えられた。

奇数番号の参加者は、同定実験を先に、偶数番号の参加者は弁別実験を先に行った。

2.4. 結果

2.4.1 知覚同化実験

各子音環境での 6 母音が日本語のどの母音のカテゴリーの音として分類されたかを表 2 に示す。選択率を見ると、常に同じ日本語の母音のカテゴリーに結びつけられた母音はほとんどない。すべての子音環境において最も多く結びつけられた日本語の母音が同じなのは、/i/、/e/だけである。ただし、/pVn/、/pVI/では/i/が「イ」と結びつけられた率は 50%を下回っていて、/hVt/、/pVt/とは明らかに異なることがみてとれる。/æ/、/a/、/ʌ/はいずれの子音環境においても 50%以上の率で同じ日本語の母音とは結びつけられてはいない。図 1 と 2 で見たフォルマント周波数の変動の影響は/pVn/での/æ/や/pVI/での/a/と/ʌ/などに見られる。

表 2: 知覚同化実験の結果:最も多い回答の選択率()内の数字は日本語の母音として良し悪しの

判定の平均

	/i/	/ɪ/	/e/	/æ/	/a/	/ʌ/
hVt	ヒ—49.6 (4.8)	ヒ 75.0 (3.9)	へ 60.6 (3.8)	ハ 45.5 (3.3)	ハ 42.4 (3.9)	ハ 39.0 (3.6)
pVt	ピ 53.0 (5.2)	ピ 77.3 (3.8)	ぺ 50.4 (4.0)	パ 31.1 (3.8)	パ 41.3 (4.7)	プ 38.6 (3.9)
pVn	ピ 62.5 (4.5)	ピ 40.5 (3.5)	ぺ 62.5 (4.2)	ペ 33.7 (3.6)	ポ 24.2 (4.5)	パ 47.4 (4.5)
pVI	ピ 37.1 (3.9)	ピ 39.4 (3.5)	ぺ 42.8 (3.1)	ペ 27.7 (3.2)	ポウ 27.7 (4.3)	ポ 34.8 (4.1)

2.4.2 同定実験

6 母音の各子音環境での平均正答率を図 3 に示す。4 つの子音環境と 6 母音を被験者内要因にした反復測定分散分析の結果、子音環境の主効果 $[F(3, 96)=47.46, p<.001]$ 、母音の主効果 $[F(5, 160)=40.80, p<.001]$ 、子音環境と母音の交互作用 $[F(15, 480)=11.38, p<.001]$ といずれも有意であった。Bonferroni で調整した多重比較の結果、/e/を除くすべての母音の子音環境の影響を受けていることがわかった。/i/と/æ/は共に/pVn/、/pVI/での正答率は/hVt/、/pVt/よりも有意に低かった($p<.001$)。/ɪ/についても同様であったが、更に/pVn/では/pVI/よりも有意に正答率が低かった($p=.024$)。/a/は、逆に/hVt/の方が/pVt/、/pVn/よりも有意に正答率が低かった(それぞれ $p=.032$ 、 $p=.039$)。また、/ʌ/は、/pVt/の方が他の 3 つの子音環境に比べて有意に正答率が高かった。

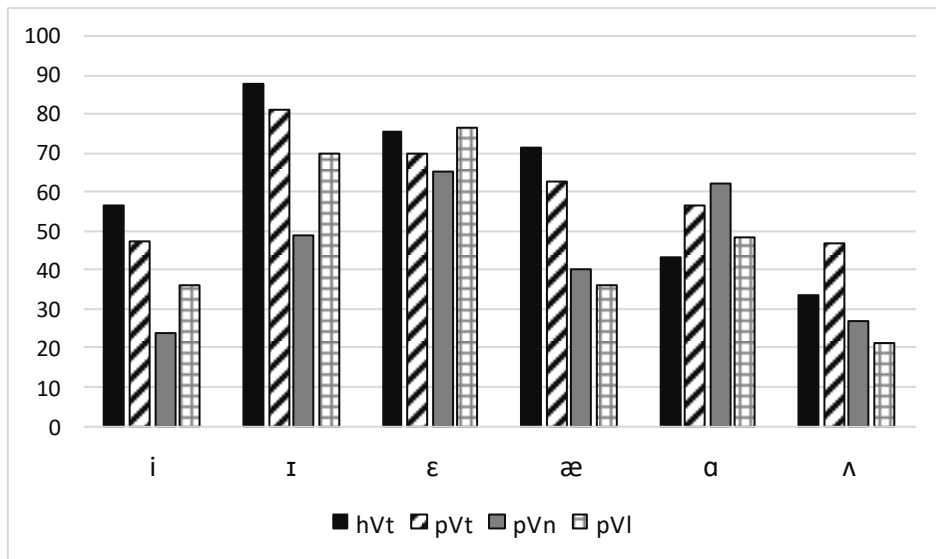


図 3: 同定実験の正答率(%)

2.4.2 弁別実験

6 母音の各子音環境での平均正答率を図 4 に示す。

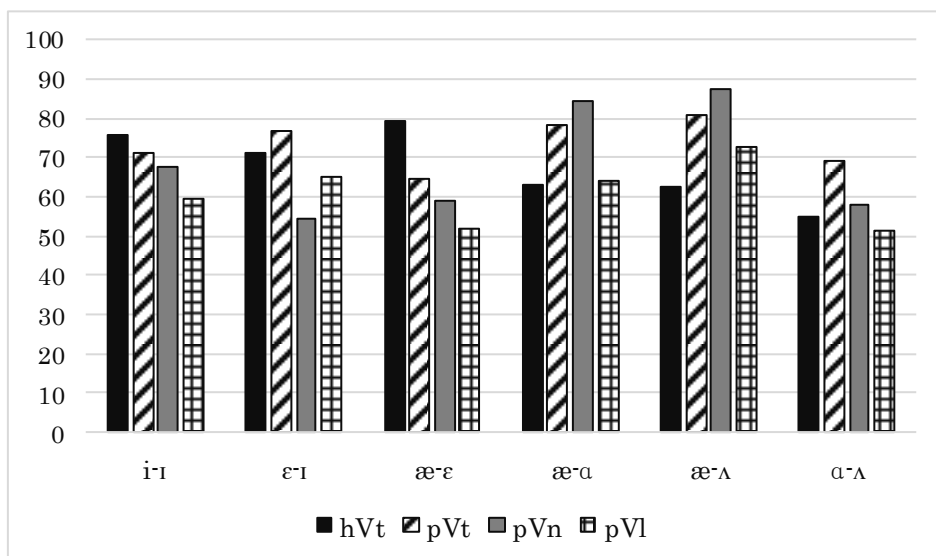


図 4: 弁別実験の正答率(%)

同定実験同様、4 つの子音環境と 6 つの母音のペアを被験者内要因にした反復測定分散分析を行った。その結果、子音環境の主効果 $[F(3, 96)=13.46, p<.001]$ 、母音のペアの主効果 $[F(5, 160)=21.63, p<.001]$ 、子音環境母音のペアの交互作用 $[F(15, 480)=12.96, p<.001]$ のいずれも有意であった。全体を通して、/pVt/での正答率が他の 3 つの子音環境に比べて $p<.05$ の水準で高く、逆に/pVl/での正答率は他の 3 つの子音環境よりも $p<.05$ の水準で低かった。Bonferroni で調整した多重比較の結果、以下に示すように全ての母音のペアの正答率に少なくとも

$p < .05$ の水準で有意差があった。

表 3: 弁別実験の多重比較の結果

/i/-/ɪ/	/ɛ/-/ɪ/	/æ/-/ɛ/	/æ/-/ɑ/	/æ/-/ʌ/	/ɑ/-/ʌ/
pVl < hVt, pVt	pVt > pVn, pVl hVt > pVn	hVt > 他全 pVt > pVl	hVt < pVt, pVn pVl < pVt, pVn	hVt < pVt, pVn pVl < pVn	pVt > 他全

3. 結論

以上見てきたように、母音に後続する/n/と/l/は日本語母語話者の英語の母音の知覚に大きく影響することがわかった。3つの実験結果がどのように関連しているかは、表や図を見ただけではわかり辛い。例えば/æ/は、冒頭で述べたように鼻音の前では前寄り・高めの位置に移動する。この影響が表2にも表れている。/æ/は一般に「ア」に近い母音とされているが、鼻音の前では「エ」に近い母音に聞こえるため、同定実験での正答率は低くなる。/æ/が/ɛ/に近づくため、弁別実験での/æ/-/ɛ/の正答率は下がり、逆に/ɑ/や/ʌ/とは離れるため、/æ/-/ɑ/と/æ/-/ʌ/の正答率は高くなる。音節末の/l/には音響的には母音との分節化が難しいが、実験参加者がどこまでを母音として、どこからを/l/として知覚したのかという問題が残る。

4人の話者の発話を音声刺激として使用したが、4人の音声刺激に対して同じ反応を示した訳ではない。知覚の実態を解明するには個々の音声刺激に対する回答を分析する必要がある。今回は日本語の母語話者であること以外に特に条件を設けずに参加者を募り、実験結果に大きな個人差が生じた。個々の実験参加者ごとに実験結果の関連を精査する必要がある。本研究は、科学研究費補助金C(16K02650)の補助を受けたものである。

参考文献

- Beddor, P. S. (1993) The Perception of Nasal Vowels, In Huffman, M.K. & Krakow, R. A. (eds.) *Nasals, Nasalization, and the Velum (Phonetics and Phonology Volume 5)* pp. 171-196. Academic Press.
- Krakow, R. A., Beddor, P. S., Goldstein, L. M. & Fowler, C. A. (1988) Coarticulatory Influence on the Perceived Height of Nasal Vowels, *Journal of the Acoustical Society of America* 83, 1146-1158.
- Labov, W., Ash, S. & Boberg, C. (2005) *Atlas of North American English: Phonetics, Phonology & Sound Change*. Mouton de Gruyter.
- Nozawa, T. & Wayland, R. (2012) Effects of Consonantal Contexts on the Discrimination and Identification of American English Vowels by Native Speakers of Japanese, *Journal of the Japan Society of Speech Sciences* 13, 19-39.
- Olive, J. P., Greenwood, A., & Coleman, J. (1993) *Acoustics of American English Speech*. New York: Springer-Verlag.

日本語母語話者による中国語声調の知覚と生成

楊 姝怡 (神戸大学大学院)・山田 玲子 (ATR/神戸大学)
syuiyang@gmail.com

1. はじめに

中国語には四つの声調の区別があるが(Chao, 1948), 非母語話者は知覚と生成の両面で声調を混同する場合がある(Chuang *et al.*, 1972; Kiriloff, 1969; Wang *et al.*, 1999; Wang *et al.*, 2003; Yang & Akahane-Yamada, 2016; 他). また, その混同の割合は声調によって異なる. Wang *et al.* (1999) はアメリカ英語母語話者を対象として, 四つの声調のうちの2つを組み合わせたミニマル・ペアの単音節語(例えば, bēi, bèi)を刺激語とした二肢強制選択課題の知覚実験を行った. その結果, 正答率は一声, 三声>四声>二声の順に低下した. また, Wang *et al.* (2003)ではアメリカ英語母語話者に単音節語を発音させ, その音声を中国語母語話者に一声~四声, または「どれでもない」の5つの選択肢から判定させた. その結果, 発話の発音明瞭度(話者が発音した声調と判定者が選んだ声調が一致した率)は一声>四声>二声>三声の順に低下した.

本研究では日本語母語話者を対象として, 中国語声調の知覚実験と発音の収録を行い, 声調による難易度およびその知覚と生成の関係について検討した.

2. 方法

2.1. 実験参加者

日本語母語話者 30 人(男性 14 人, 女性 16 人; 18~28 歳, 平均 21 歳)が参加した. 全員中国語学習経験はなかった. また, 中国語を使う地域での一か月以上の居住経験もなかった.

2.2. 刺激音

4 つの声調で対立する中国語単音節語(例えば, bā, bá, bǎ, bà)を 1 組とし, 36 組 144 語を刺激語とした. 刺激語を 1 人の中国語母語話者(女性, 23)が読み上げ, サンプリング周波数 44.1kHz, 量子化精度 16bit でデジタル化し, 音声ファイルとして格納したものを音声刺激とした.

2.3. 手続き

実験開始時に, 中国語-日本語のバイリンガルである実験者が中国語声調の説明が日本語で記されたプリントを見せ, 説明を読み上げるとともに 4 つの声調の音声を聞かせつつ, 声調の違いを教示した. この教示の後, 実験を開始したが, 実験では刺激の提示, 反応の取得はすべてパソコンで制御した.

知覚実験では, 実験参加者は静かな部屋でパソコンに向かってヘッドホンを付けて座り, キーボードを使って回答した. 各トライアルでは 1 つの刺激語に対してそれと声調で対立する 3 語をディストラクタとし, 4 つの選択肢がピンインでパソコンのディスプレイ上に表示された(図 1). 選択肢はトライアル毎にランダムな順に配置された. 実験参加者には聞

こえた音声と一致した選択肢を選ぶように教示した。刺激音はトライアルの冒頭で 1 回だけ再生され、繰り返し聞くことはできなかった。144 の刺激音をランダムな順で呈示した。

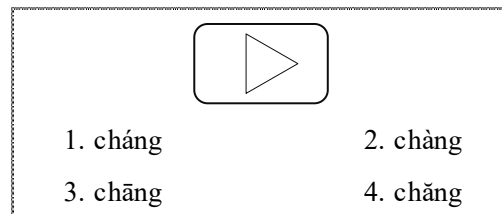


図 1: 知覚実験の画面例

音声の収録においては実験参加者は刺激音を聴取し、それをまねして発音した。実験参加者には声調の変化に注意して発音するように、あらかじめ教示した。その発音をサンプリング周波数 44.1kHz、量子化精度 16bit でデジタル化し、音声ファイルとして格納した。後に 6 人の中国語母語話者(女性 3 人, 男性 3 人; 22~33 歳, 平均 25 歳)がその音声を声調で対立する 4 種類のピンインのうちのどれに聴こえるか、前項の知覚実験と同じ強制選択課題を用いて評定した。一人の評定者は 4,320 全音声ファイル(144 刺激語×30 人)を話者(実験参加者)ごとのブロックに分けて評定した。

3. 結果

知覚実験について、参加者ごと声調ごとの平均正答率を求めた。声調(一声～四声)を要因とし、平均正答率を従属変数として F 検定を行った結果、声調によって正答率は有意に変動した[F (3,116) = 12.681, p<0.0001](図 2 左)。一声～四声の正答率はそれぞれ 74.6%, 58.4%, 96.8%, 67.9%であった。

生成では、各発話音声について 6 人の評定者のうち刺激語と同じ語を選択した割合を求めた。本研究ではこの割合を発音明瞭度と呼ぶ。声調(一声～四声)を要因とし、発音明瞭度を従属変数として F 検定を行った結果、声調(一声～四声)の効果は有意だった[F (3,116) = 29.270, p<0.0001](図 2 右)。一声～四声の正答率はそれぞれ 96.3%, 87%, 78%, 98.8%であった。

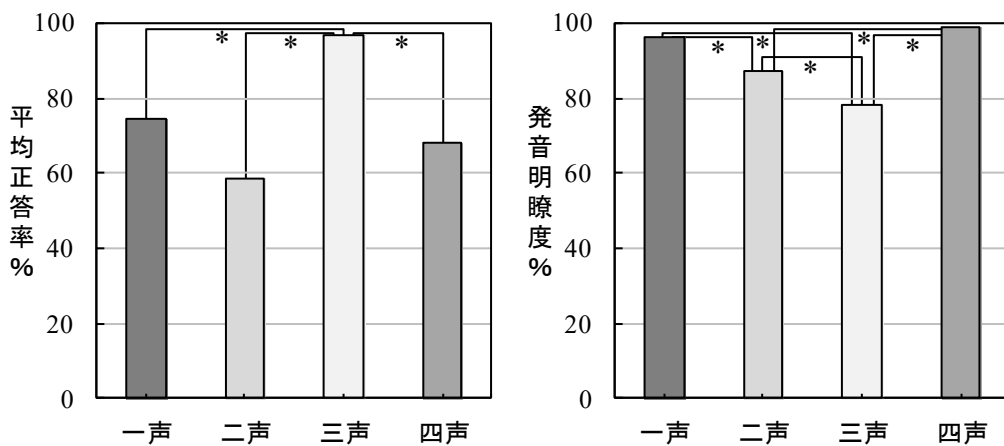


図 2: 知覚実験(左)と生成(右)の結果 * $p < 0.001$

4. 考察

知覚実験の結果では、平均正答率は三声が一声、二声、四声より有意に高かった(三声 > 一声, 二声, 四声)。生成の明瞭度は三声が一声、二声、四声より有意に低く、四声、一声 > 二声 > 三声の順に低下した。知覚と生成の結果を比較すると、三声は知覚の正答率は最も高いが発音明瞭度は最も低く、一声と四声は知覚の正答率は低いが発音の明瞭度は 100% に近かった。知覚は容易だが発音が困難、またはその逆の音が存在することが示唆されたといえる。

また、知覚の結果は Wang *et al.* (1999) の知覚実験の結果とは、全体的な傾向は類似していたが、本研究では三声の正答率が顕著に高かった (図 3 左)。生成の結果は Wang *et al.* (2003) のアメリカ英語母語話者の生成の難易度と全体的な傾向は類似していたが、本研究では声調間の差が小さかった。

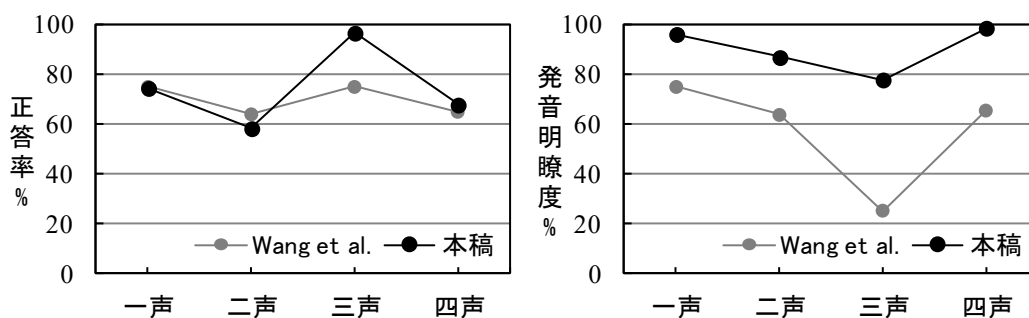


図 3: Wang 等によるアメリカ英語母語話者の結果と日本語母語話者を対象とした本研究結果の比較。左が知覚の比較(Wang *et al.*, 1999), 右が生成の比較(Wang *et al.*, 2003)。

Wang 等と本研究と比較すると、課題が多少異なるので直接の比較はできないものの、アメリカ母語話者と日本語母語話者で声調の効果について生成、知覚とも類似した傾向を示したことは、声調の難易度は両母語話者に共通した傾向である可能性を示した。また、本研究では中国語を全く知らない日本語母語話者を対象としたが、特に生成について高スコアを示したことは興味深い。声調のみならず中国語音声に関する知識を持たないは、ピッチの変動を真似することで一時的に明瞭度が上昇している可能性がある。母語である日本語がピッチアクセントであることも関係しているかもしれない。今後、中国語学習過程との関連を検討することにより、効果的な発音習得方法について検討したい。

参考文献

- Chao, Y. R. (1948). *Mandarin primer: An intensive course in spoken Chinese*. Harvard University Press.
- Chuang, C. K., & Hiki, S. (1972) "Acoustical features and perceptual cues of the four tones of standard colloquial Chinese", *The Journal of the Acoustical Society of America* 52(1A), 146-146.
- Kiriloff, C. (1969) "On the auditory perception of tones in Mandarin", *Phonetica* 20(2-4), 63-67.
- Wang, Y., Spence, M. M., Jongman, A., & Sereno, J. A. (1999) "Training American listeners to perceive Mandarin tones", *The Journal of the Acoustical Society of America* 106(6), 3649-3658.
- 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 America* 113(2), 1033-1043.
- Yang, S., & Akahane-Yamada, R. (2016) "Effects of perception and vocabulary training of Mandarin tones for native speakers of Japanese: Pre-, post-, and retention test comparison", *The Journal of the Acoustical Society of America* 140(4), 3392-3392.

自然談話における終助詞「ね」と「よ」のイントネーション

張海霞(東京学芸大学大学院) 白勢彩子(東京学芸大学)
 m161419k@st.u-gakugei.ac.jp shirose@u-gakugei.ac.jp

1. はじめに

日本語の会話において、終助詞の「ね」と「よ」は高頻度に出現し、話者の意図を伝達する重要な役割を担っている。イントネーションにより、「ね」と「よ」は意味機能に変化があるが、二つの関係を体系的に議論した研究は多くない。そこでイントネーションと意味機能の関係をより詳しく記述するため、我々は文に上昇調、平坦調、下降調三パタンの文末イントネーションを付与し、意味判定実験を実施した(Zhang and Shirose,2016)。その結果、「ね」は下降調を伴うと「不自然」と高く判断され、「よ」は下降調を伴うと、「困惑」「非難」のような不快な意味に判断される傾向があることが分った。

本稿では、前回の聴取実験を踏まえ、自然談話における終助詞「ね」と「よ」及びイントネーションパタンの出現状況を調べ、意味機能の関係を検討する。

2. 先行研究

終助詞「ね」と「よ」の意味機能に関しては、従来多くの研究が行われてきた。金水(1993)「ね」は話し手は聞き手と知識を共有するため、「確認」「同意を求める」の意味機能を持ち、「よ」は聞き手の知らない知識を教えるために用いられると述べた。メイナード(1992)は、話し手と聞き手のそれぞれの情報量の状況から、終助詞「ね」と「よ」の使用条件を呈示した。金水とメイナードのように、情報帰属あるいは情報量の視点から、終助詞「ね」「よ」の意味機能を議論した研究が多いが、イントネーションとの関係を論じた研究は多くない。終助詞の意味機能は終助詞の元々の意味とイントネーションとの相互が関わると述べた片桐(1997)、自己内省で終助詞「ね」と「よ」の意味を分析した小山(1997)音声合成で「ね」の機能の境界を論じた杉藤(2001)、ピッチ曲線でイントネーションを分類し、意味機能を分析したデイヴィッド(2013)がある。しかし、自然談話において、終助詞「ね」と「よ」のイントネーションがどのように現れるについての調査はほぼない。

本稿は自然談話を前提として、終助詞「ね」「よ」のイントネーション様子を呈示しつつ、意味機能との関係も検討する。

3. 方法

3.1 資料

『千葉大学 3 人会話コーパス』は日常場面に近い自然な会話を収めた、大学生・院生・ポスドクを含む 3 人からなる友人同士 12 組の雑談、7599 発話のデータである。今回の調査ではこのコーパスを利用し、終助詞「ね」と「よ」の生起頻度を調査した。

3.2 対象となる文の選別基準

『日本語国語大辞典』で、「終助詞は文の終わりにあって、命令・疑問・反語・願望・禁止などの意味を決定し、または陳述の意味を強めたり、感動を表わしたりする助詞」と定義していることから、文の中間で使用される文を対象外とする。今回の調査でも、複合助詞、接続助詞との接続、疑問詞、相槌表現も対象外とした。具体的な標準は表 1 に示す。対象となる文は表 2 のような文である。

表 1 対象外とする基準

基準	例
A 文間で使用される「ね」「よ」	私はね
B 複合助詞	よね、かね
C 単独使用	ね
D 接続詞と接続する場合	思うけどね
E 疑問詞	いつ…ね
F 相槌表現	そうだね

表 2 対象となる文

番号	例
1	水冷売ってるよ
2	今週末ぐらいだったね
3	大昔だね
4	怖かったよ
5	すごいね

3.3 文末イントネーションの分類基準

音響分析ソフトウェア Praat に基づくピッチ曲線の視察および聴覚印象により、終助詞「ね」「よ」のイントネーションパターンを分類した。

今回の調査では、終助詞「ね」と「よ」部分のイントネーションを上昇調、平坦調、下降調、複合音調に分類した。上昇調と平坦調の区別について、10Hzで分けた。「ね」と「よ」内の音調は10Hz以上上昇した場合、上昇調をし、10Hz以内で変動する場合、平坦調とした。複合音調も10Hzを基準にし、2か所以上で10Hz以上変動すると、複合音調とした。ピッチ曲線を抽出されない場合もあり、知覚による判断とした。以下は上昇調(図1)、平坦調(図2)、下降調(図3)、複合音調(図4)のサンプル(Praat分析画面)である。

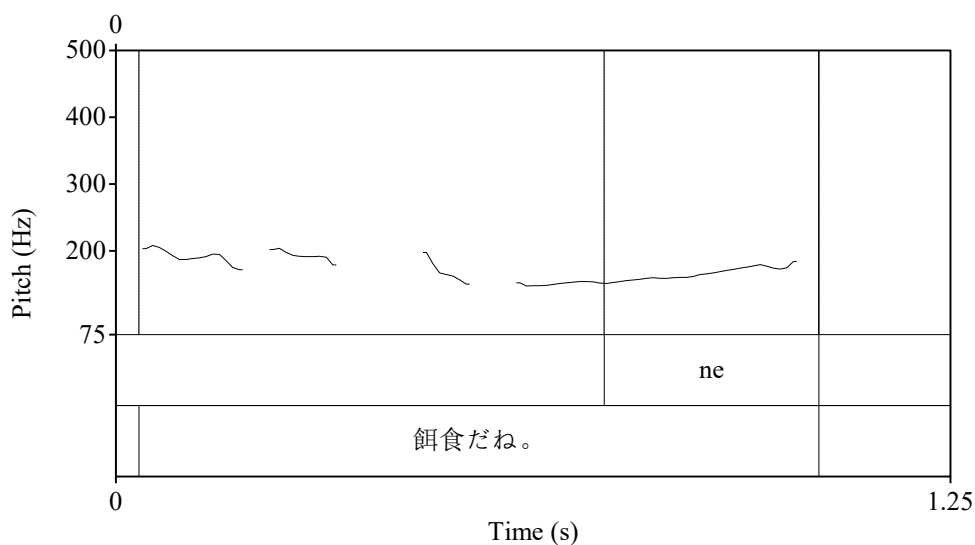


図1 上昇調

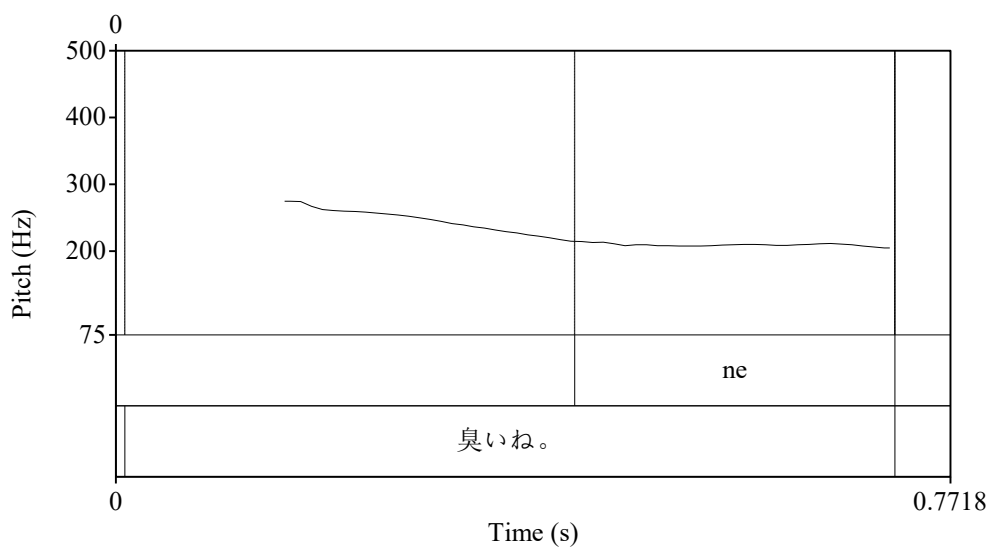


図2 平坦調

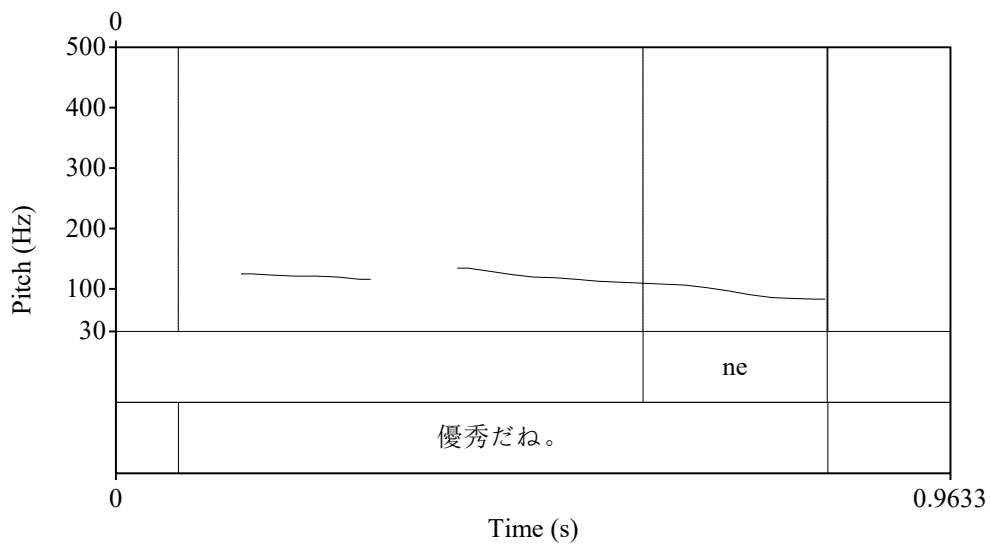


図 3 下降調

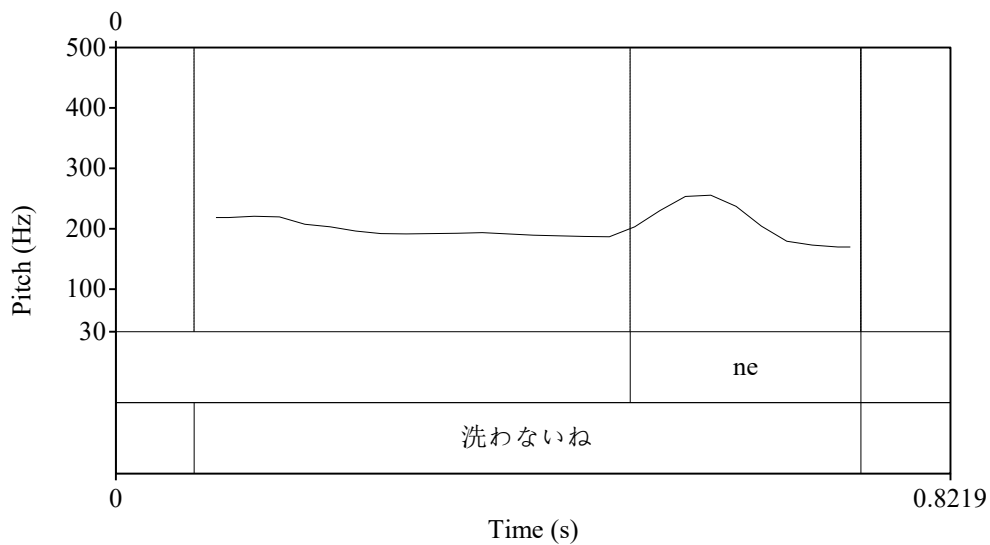


図 4 複合音調

4 結果と考察

3.2の選考基準により、「ね」と「よ」、それぞれ84発話と145発話が抽出された。終助詞「よ」の使用数は「ね」より約1.7倍多く生起していることが分かった。イントネーションの分類の結果、終助詞「ね」と「よ」の出現数をそれぞれ図5と図6に示す。

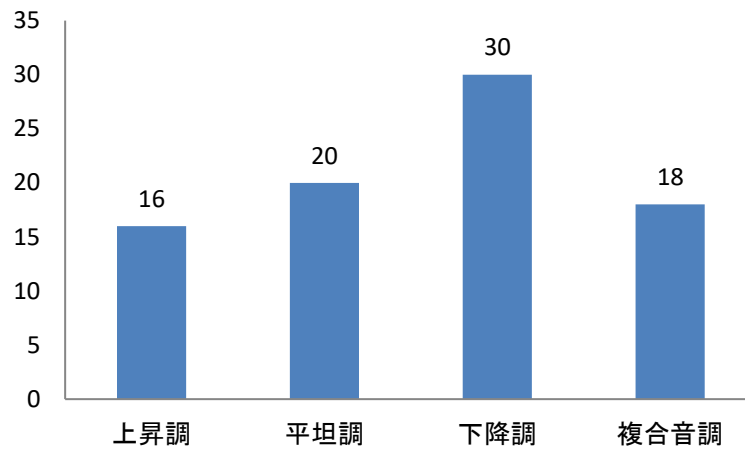


図5 終助詞「ね」のイントネーションパタンの出現数

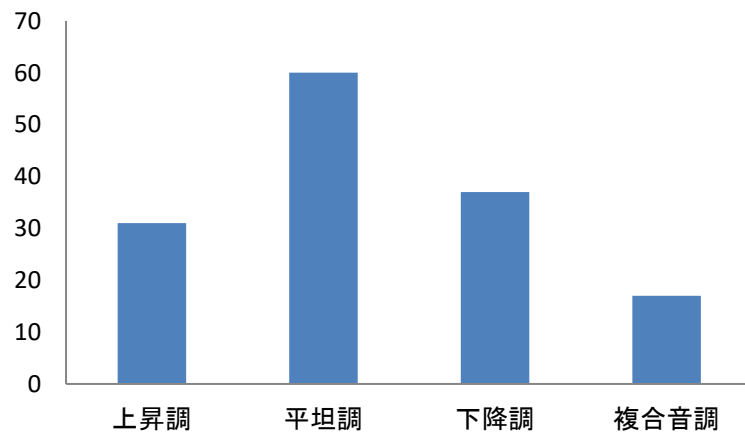


図6 終助詞「よ」のイントネーションパタンの出現数

終助詞「ね」の出現状況について、上昇調 16(19.1%)、平坦調 20(23.8%)、下降調 30(35.7%)、複合音調 18(21.4%)という結果となった。下降調は他の音調より多く出現することが分った。前回の聴取実験では、下降調になると、不自然(61.5%)と判断されやすく、イントネーションによる意味機能の変動が大きくないという結果があった。ただし前回の聴取実験では、下降調になると、「感心」というような意味が高く判断される傾向があった。今回の調査で対象とした自然談話では、「感心」のような自己感情を表す意図が多く出現し、このことが下降調の生起に反映したと考えられるだろう。

終助詞「よ」の出現状況については、上昇調 31(21.4%)、平坦調 60(41.4%)、下降調 37(25.5%)、複合音調 17(11.7%)という結果となった。複合音調はあまり出らず、平坦調が主に出現したことが分った。前回の聴取実験では、「よ」の上昇調、平坦調、下降調について、不自然と判断されにくく、どれでも許容して使用いられるという結果があった。下降調になると「困惑」「非難」のような不快な意味になる傾向があり、「よ」の意味機能はイントネーションにより変動しやすいとの結果があった。今回の調査結果から見ると、平坦調を主に出現することで、話し手と聞き手が談話をスムーズに進められるように、意味機能を抑え、イントネーションを採用したと考えられる。

5 まとめ

コーパス分析の結果から、自然談話において、終助詞の「よ」は「ね」より多く使用されていることが分った。終助詞「ね」について、非上昇調が主に出現し、「感心」というような自己感情を表す発話が多かったと考えられた。「よ」については、平坦調が主に出現することは、自然談話を順調に続けるため、意味機能を抑えたと考えられるだろう。自然談話における、終助詞の意味機能とイントネーションの相互関係の詳細については、今後の課題としたい。

参考文献

- 金水 敏 (1993) 「終助詞ヨ・ネ」『言語』第 22 巻第 4 号, pp.118-121.
- メイナード・泉子・k (1993) 『会話分析』 pp.101-109,くろしお出版.
- 伊豆原 英子 (1993) 「「ね」と「よ」再考-「ね」と「よ」のコミュニケーション機能の考察から」 『日本語教育』 80 号, pp.103-114.
- 伊豆原 英子 (2001) 「「ね」と「よ」再再考」『愛知学院大学教養部紀要』第 49 巻第 1 号, pp.35-49.
- 小山 哲春 (1997) 「文末詞とイントネーション」『文法と音声』 pp.97-119,くろしお出版.
- 片桐 恭弘 (1997) 「終助詞とイントネーション」『文法と音声』 pp.235-256 ,くろしお出版
- 杉藤 美代子 (2001) 「終助詞「ね」の意味・機能とイントネーション」『文法と音声Ⅱ』 pp.3-16,くろしお出版.
- 大島 デイヴィッド 義和 (2013) 「日本語におけるイントネーション型と終助詞機能の相関について」『国際開発研究フォーラム』 43, pp.47-63
- Zhang and Shirose (2016) A study of Japanese modal particles “ne” and “yo”, Journal of Acoustical Society of America, 140-4, P3394.

F0, long-term formants and LTAS in Korean-English Bilinguals

Sylvia Cho, Murray J. Munro (Simon Fraser University)
sylvia_cho@sfu.ca, mjmunro@sfu.ca

1. Introduction

Voice quality measures such as speaking fundamental frequency (SFF), long-term formants (LTF), and long-term average speech spectra (LTAS) have been useful for speaker identification of monolingual speech samples (Nolan & Grigoras, 2005). However, the existing studies on SFF, LTFs, and LTAS either lack bilingual data or have reported obscure or confusing results. It is thus largely unknown whether the acoustic properties of speech samples from a bilingual speaker differ according to the language being spoken. Such comparisons have important implications for the practice of forensic speaker identification, as bilinguals may or may not behave like their monolingual counterparts in their productions. The current study therefore aims to contribute to this line of work by (1) extending SFF comparisons to Korean-English bilinguals, (2) conducting a comparison of bilingual LTFs that employs LTF histograms (as well as mean values), and (3) confirming whether LTAS differences between monolingual speakers of Korean and English are also found in Korean-English bilinguals.

2. Production Experiment

2.1. Methods

2.1.1 Participants

Ten Korean-English bilingual speakers (5M 5F) were recruited at Simon Fraser University. All subjects spoke standard Seoul Korean and Western Canadian English. Their age ranged from 18 to 36 (Mean age: 25), and their age of arrival (AOA) in Canada ranged from 0 to 12 years (Mean AOA: 5.8). Each subject was screened prior to the recording to ensure that they met the following criteria: (1) they were judged to be proficient Korean-English bilinguals by one of the authors who is proficient in both languages (2) they did not report any hearing or speech impairments; (3) they had no signs of sickness from colds or other respiratory problems on the day of the recording.

2.1.2 Procedures

A total of four speech samples were taken from each subject. The subjects were instructed to read a passage out loud (i.e., *The North Wind and the Sun*), at a comfortable pace in both Korean and English. They were also asked to produce extemporaneous narratives in both languages, in which they were instructed to describe a picture story (Derwing, Munro, Thomson, & Rossiter, 2009).

3. Results and Discussion

3.1. Speaking Fundamental Frequency (SFF)

Speaking fundamental frequency (SFF) was measured using the autocorrelation method in Praat (Boersma & Weenik, 2016). Paired-sample *t*-tests¹ were then carried out to assess the existence of a language effect on SFF. These results are shown in Figures 1 and 2.

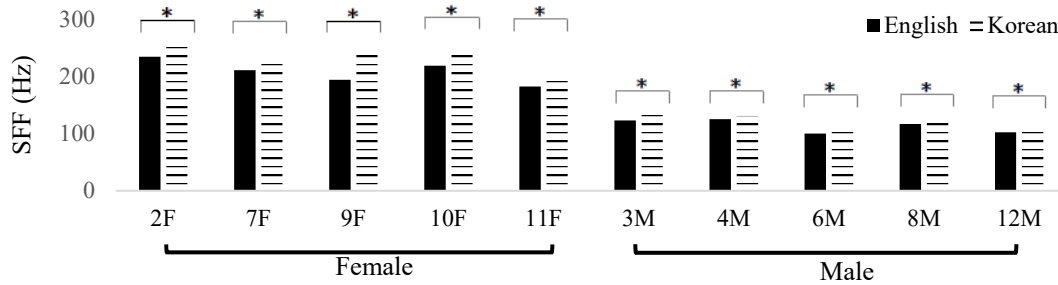


Figure 1: Average SFF (in Hz) for passage reading in Korean and English

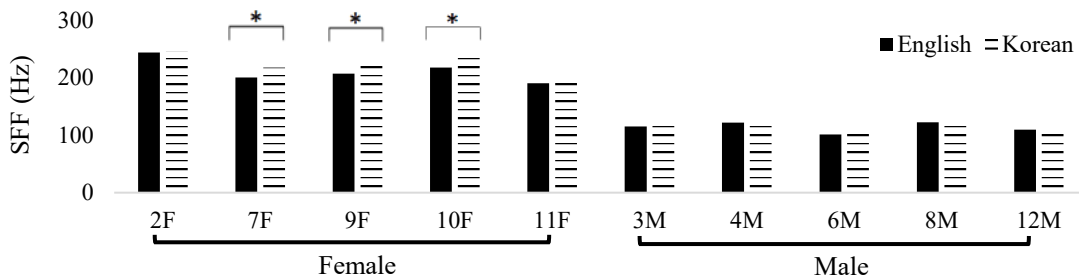


Figure 2: Average SFF (in Hz) for extemporaneous speech in Korean and English (* $p < 0.05$)

As displayed in Figures 1 and 2, significant SFF differences were found between English and Korean in the passage reading, but this language effect was not as robust in the extemporaneous data (i.e., only 3 out of 10 speakers adjust their SFF significantly). With the passage reading, there was a mean F0 difference of 15 to 45 Hz between the languages for the female subjects, and for the male subjects, there was a difference of 5 to 20 Hz. In the extemporaneous speech task, there was a maximum of 15 Hz difference in the females, and up to a 5 Hz difference in the males. This indicates an effect of speaking style on SFF, and that casual or spontaneous speech is produced with smaller between-language differences. In terms of the direction of the SFF change, there was either no statistical difference between the bilinguals' production of Korean and English, or English was produced with a lower SFF. Regarding this observed pattern, it is possible that subjects produced English with a lower

¹ Each speech sample was parsed into 20 stretches of equal duration. The mean F0 value for each stretch was obtained, and these values (i.e., 20 data points for Korean; 20 data points for English) were used in the paired-sample *t*-tests.

SFF due to a language-dependent tendency. That is, English has been associated with a relatively low SFF, compared to that of other languages (Hanley & Snidecor, 1967). This tendency seems to be more readily seen in read speech, as opposed to extemporaneous speech, which better approximates natural speaking conditions.

3.2. Long-term Formants (LTF)

LTFs were computed for speech samples of equal duration (i.e., first 30 seconds of speech sample) using formant frequencies chosen by the automatic formant tracker in Praat, which uses linear predictive coding. In terms of mean formant values, there were very small within-subject differences. This study finds that the shapes of the formant frequency distributions are more indicative of speaker identity than are the means. For example, the histograms in Figures 3 and 4 display the distributions of formant value ranges (displayed in 100 Hz bins) for two speakers that have very similar mean formant values (i.e., F1: 400 Hz, F2: 16-1800Hz, F3: 2800Hz). While their average formant frequencies are similar, Subject 4M, has a more prominent F1 peak, and his F3 peak is skewed more to the left.

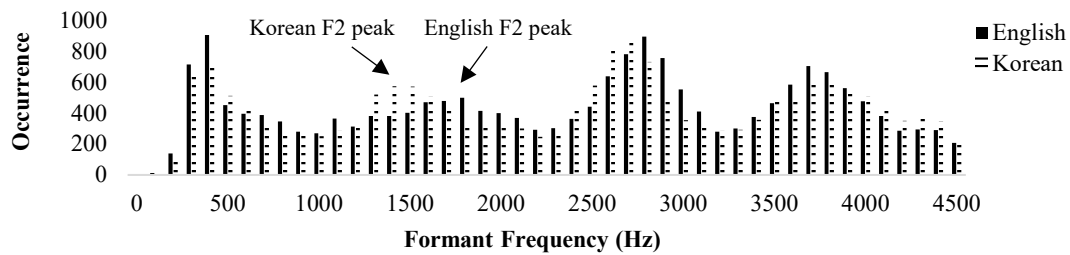


Figure 3: Read Speech LTF for Subject 3M

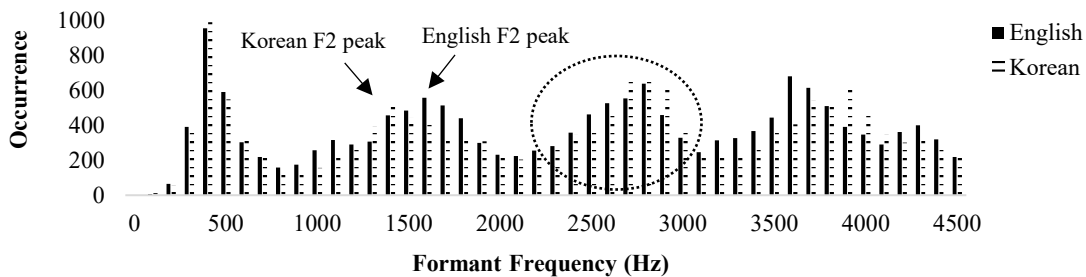


Figure 4: Read Speech LTF for Subject 4M

As shown in the histograms above, there is also a small effect of language across bilingual LTFs. That is, Korean productions are often associated with a lower F2 peak. This is consistent with acoustic studies that report slightly lower F2 values for Korean vowels in comparison to English vowels (Yang,

1996). A strong effect of speech style was not found with bilingual LTFs. Altogether, the current study finds that LTFs are mostly speaker-specific with small language-specific effects.

3.3. Long-term Average Speech Spectra (LTAS)

There appear to be no existing studies on bilingual LTAS, but a study that compared LTAS of twelve languages² found that most of the variations among languages are smaller than 3 dB in magnitude and occur within the range of 200 to 6300 Hz (Bryne et al., 1994). It was thus suggested that a “universal” LTAS can be applied across languages. A more recent study (Noh & Lee, 2012) compared Korean LTAS to that of English, and found that the LTAS of Korean speakers show significantly lower intensity levels in frequencies above 2000 Hz, except at 4000 Hz. The lower intensity levels found in Korean are thought to stem from lower occurrence of phonemes with high frequency components (i.e., fewer fricatives in Korean). This suggests that for languages such as Korean, the universal LTAS is rendered inappropriate. The current study examines whether the trends found for monolingual English and Korean speakers also hold true in bilingual speakers.

LTAS were computed using Praat to observe overall spectral properties of the continuous speech samples. Each speaker’s samples were then normalized for peak intensity (dB) to permit straightforward comparisons. This study finds that most speech spectra show maximum intensity around 500 Hz, and a decrease is seen in the higher frequencies. Similar to the monolinguals in Noh and Lee’s study, many bilingual speakers (6 out of 10) showed significantly lower intensity levels in frequencies above 2000 Hz (except at 4000 Hz) in their production of Korean (see Figure 5). There was a 3 to 5 dB difference from 2000 to 4000 Hz, and a difference of 5 to 10 dB in the higher frequencies.

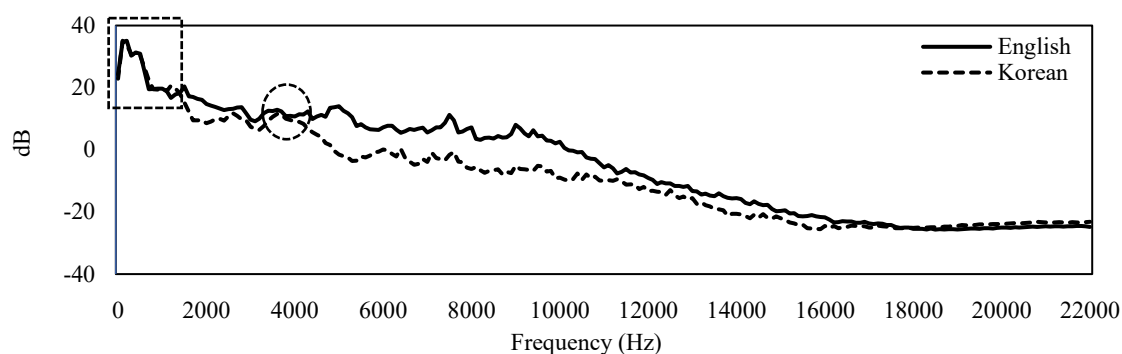


Figure 5: Extemporaneous Speech LTAS for Subject 4M

It should be noted that this pattern was not observed across all speakers. Some bilinguals (4 out of 10) had very similar intensity levels in the two languages up to 4000 Hz (see Figure 6). This suggests that

² English (several dialects), Swedish, Danish, German, French, Japanese, Cantonese, Mandarin, Russian, Welsh, Singhalese, Vietnamese

there are some within-speaker characteristics that influence bilingual LTAS – which is different from what is found across monolingual speakers (i.e., between speakers). The reasons for such a distribution are not clear at the moment, but it seems that the bilingual speakers that were impressionistically more dominant in one language (i.e., better at either Korean or English) demonstrated such patterns.

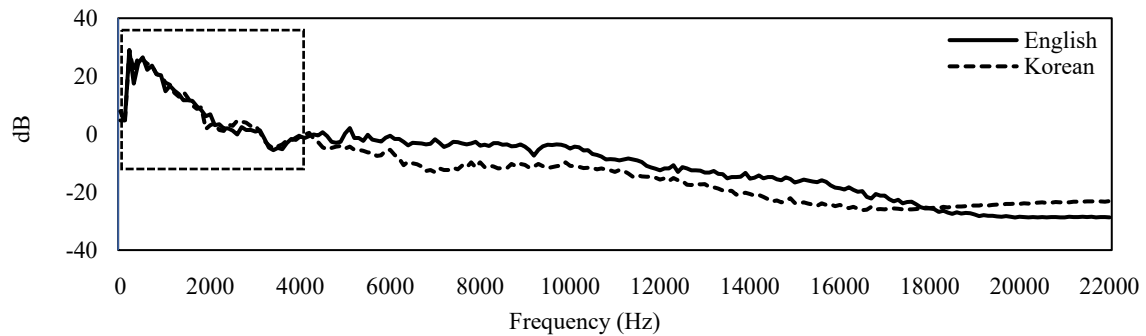


Figure 6: Extemporaneous Speech LTAS for Subject 2F

Furthermore, this study found nearly identical LTAS for bilingual speakers that had clearly lower proficiency in one of the languages. Figure 7 shows the LTAS of a speaker that was omitted from the scope of this study due to his lack of native-like proficiency in English (late-onset bilingual, AOA: 18 years old).

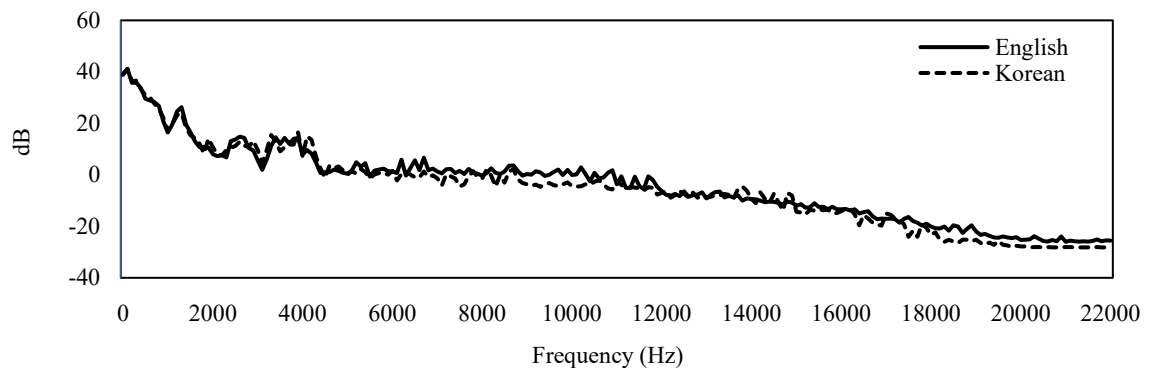


Figure 7: Extemporaneous Speech LTAS for Subject 5M

As can be observed from the figure above, energy differences seen across other Korean-English bilinguals were not found for Subject 5M, namely a subject who is not very proficient in one of the languages. It is plausible that such speakers produce the less-proficient language in much the same manner as their native language. Such findings suggest that LTAS analysis can be somewhat useful in speaker identification cases involving multilingual samples, given that the speaker lacks proficiency in one of the languages spoken. Lastly, there was little indication of an effect of speech style on LTAS distributions.

4. Conclusion

This study investigated the usefulness of SFF, LTF, and LTAS analyses in bilingual speaker identification. Among the different acoustic measurements, formant frequency estimates across long speech samples were found to be most reliable and speaker-specific. Speaking style was also found to be important; stronger SFF correlations were found between languages in extemporaneous speech. Such results suggest that multilingual samples may serve some useful purpose in speaker identification cases that involve bilingual speakers. Future research could further expand on the results obtained from this project by applying the methods to other bilingual populations, such as Japanese-English bilinguals. One previous study finds that SFF is consistently higher in Japanese for Japanese-English bilinguals (Graham, 2015), and it might be the case that, like Korean-English bilinguals, there is some effect of language in the bilingual LTFs given the small vowel inventory of Japanese.

References

- Boersma, P., & Weenink, D. (2016). Praat: doing phonetics by computer (Version 6.0.23) [Computer program]. Retrieved from <http://www.praat.org/>
- Byrne, D., Dillon, H., Tran, K., Arlinger, S., Wilbraham, K., Cox, R., ... Ludvigsen, R. (1994). An international comparison of long-term average speech spectra. *The Journal of the Acoustical Society of America*, 96(4), 2108-2120.
- Derwing, T. M., Munro, M. J., Thomson, R. I., & Rossiter, M. J. (2009). The relationship between L1 fluency and L2 fluency development. *Studies in Second Language Acquisition*, 31(4), 533-557.
- Graham, C. (2015). Fundamental Frequency Range in Japanese and English: The Case of Simultaneous Bilinguals. *Phonetica*, 71(4), 271-295.
- Hanley, T. D., & Snidecor, J. C. (1967). Some acoustic similarities among languages. *Phonetica*, 17, 141-148.
- Heeren, W., Van der Vloed, D., & Vermeulen, J. (2014). Exploring long-term formants in bilingual speakers. *Proceedings from the 2014 Conference of the International Association for Forensic Phonetics and Acoustics*. Zurich, Switzerland.
- Krebs, P., & Braun, A. (2015). Long Term Formant measurements in bilingual speakers. *Proceedings from the 2015 Conference of the International Association for Forensic Phonetics and Acoustics*. Leiden, Netherlands.
- Noh, H., & Lee, D. (2012). Cross-Language Identification of Long-Term Average Speech Spectra in Korean and English. *Ear and Hearing*, 33(3), 441-443.
- Nolan, F., & Grigoras, C. (2005). A case for formant analysis in forensic speaker identification. *International Journal of Speech, Language and the Law*, 12(2), 143-173.
- Yang, B. (1996). A comparative study of American English and Korean vowels produced by male and female speakers. *Journal of Phonetics*, 24(2), 245-261.

東京出身者の母音フォルマントの世代間比較

脇田 佳幸 (NIEMAN INTERNATIONAL SCHOOL)
wakida.13y@gmail.com

1. はじめに

本研究は『日本語話し言葉コーパス』(CSJ)中の東京出身者の母音を音響分析することで、母音のフォルマントに世代差がある可能性を指摘する。日本語(共通語)の母音の発音に世代差がある可能性については、いくつか短い指摘があるものの、それを音響分析で検証した研究はほとんどない。そこで本研究は、母音フォルマント(F1・F2)に世代差が見られるかどうかを検証する。

2. 母音の世代差に関する指摘

日本語(共通語)の母音に音変化が見られる可能性について、音声の専門家が複数指摘を行っている。例えば、ウの円唇性の度合いの変化について、杉藤美代子(1997: 3-4)は、東京の若い世代に非円唇ではなく円唇のウの発音があるとする。また、アの変化について、井上史雄(1989: 111)は、「現代日本語では母音 a が前寄りになり、おそらくその影響で、e が狭くなる傾向がみられる」と指摘する。さらに、柴田実(2005: 66)は、若い世代のアはあごの開きが不足して明るさがない、と主張している。

いずれも音声に関わる専門家が母音の変化を指摘したものである。しかし、これらを受けて、母音の世代差を実証的に検証した研究はほとんどない。そこで本研究は、日本語の母音を音響分析することで、母音に世代差があるのかどうかを実証的に示すことを目指す。

3. 使用データ

現代日本語の自発音声を収録した『日本語話し言葉コーパス』(CSJ)を用いる。CSJには講演・自由対話など様々な種類の音声が含まれるが、スタイルによる差を排除するため、「模擬講演」(CSJのデータ採取のために行われた簡単な講演)のみを使用した。また、方言による差や性別による差を排除するため、「東京で生まれ」「首都圏以外に在住歴がない」「男性」に話者を限定した。

以上の条件を満たす話者は、63人となった。そのうち、生年代が最も離れた2つの集団として、1930年～1944年に生まれた高年層9人と、1975年～1984年に生まれた若年層10人を取り出し、この2つの集団でフォルマント周波数に違いが見られるかを分析した。なお、それぞれの話者は、1回あたり10数分の講演を、平均して約3回ずつ行っている。

これらの講演の中から、実際に分析する母音を次のようにして選定する。

- ・先行子音の影響を排除するため、ポーズの直後にある母音に限定する。CSJの書き起こしテキストは、0.2秒のポーズごとに分割されているので、その冒頭のみを対象とする。
- ・後続子音による影響を小さくするため、「サ・ザ・タ・ダ・ナ・ラ行のいずれかの子音

が続く」母音に限定する。これは、母音や半母音が後続する母音を除外し、さらに、後続する子音の調音位置を統制したものである。

このようにして選定した各母音を対象に、母音区間の中央 1/3 について、F1・F2 の平均値を計測した。計測には Praat (version 6.0.09) を使用し、Formant settings (設定) は Maximum formant (Hz) の値を 5000 として、Number of formants の値は 5 とした¹。母音が無声化している場合は、声帯の安定した振動が観察されないので、使用不可能と判断して除外した。同様に、後続子音の閉鎖や摩擦が極端に弱まっている場合も、母音区間を特定できないので、使用不可能と判断して除外した。

最終的に、818 個の母音を使用サンプルと位置付けて、分析を行う。高年層と若年層それぞれの使用サンプルの個数は、表 1 の通りである。

表 1 使用サンプルの個数

	ア	イ	ウ	エ	オ	計
高年層 [*]	156	146	24	18	50	394
若年層	184	119	46	18	57	424
計	340	265	70	36	107	818

4. 分析

4.1. サンプルの偏りについて

実際の分析に入る前に、得られたサンプルの偏りについて、3つの観点から検討する。コーパスからある条件を満たすものを拾い出す、という手法を採っている以上、話者ごとのサンプル数、後続音の音環境、アクセント核の有無、などの条件は、完全には統制できない。そこで、これらの偏りが、世代間比較に際してどの程度影響するかを考察する。

①話者ごとのサンプル数の偏り

例えばアを見ると、話者ごとに見た高年層の平均は 17.3 個であるが、ある話者は 43 個も含む。サンプル数の多い話者の影響で、その世代のフォルマント周波数の分布が、大きく歪められてしまうかもしれない。対応策として、平均の 2 倍以上のサンプルを含む話者については、フォルマント周波数が特異な値を取っていないかを確認する。

②後続音の音環境の偏り

計測する母音のフォルマント周波数は、後続母音によって影響を受けると予想される。これを考慮せずに分析をしてしまうと、後続母音の違いによって生まれるフォルマント周波数の差を、世代間の差異と見なしてしまうかもしれない。後続母音の違い

¹ ただし、オの計測においては、F1 と F2 が接近するため、それらが 1 つのフォルマントとして捉えられてしまうことがあった。これを避けるため、3000Hz までに 3 つのフォルマントが安定的に示されない場合は、Number of formants を 7 に変更した。

を考慮に入れても世代間の差があるのかを検討するために、後続母音については、分散分析の要因として扱う。

③アクセント核の有無による偏り

アクセント核の有無によって、フォルマント周波数に差が出る可能性もある。そこで計測に当たっては各母音にアクセント核が有るかどうかも入力し、高年層と若年層の世代ごとに、各母音について、アクセント核が有るものと無いものの個数を比較した。両世代とも、アとイに関しては、アクセント核が有るものも無いものも多くのサンプルを含む一方、ウとエとオに関しては、アクセント核のあるサンプルの個数はいずれも一桁であり、無視できるほど少ないことがわかった。したがって、分散分析において、アとイではアクセント核の有無を要因として扱うべきであり、ウとエとオでは要因として扱う必要は無いと判断した。

結果として、①～③の偏りは、世代間比較を不可能にするものでは無いと判断できる。

4.2. 計測値での分析

ここから具体的な分析に入る。まず、計測して得られたフォルマントの数値そのものによる分析を行う。世代ごとに一括した各母音の平均値は図 1 の通りである（描画には NORM: The vowel normalization and plotting suite を使用）。

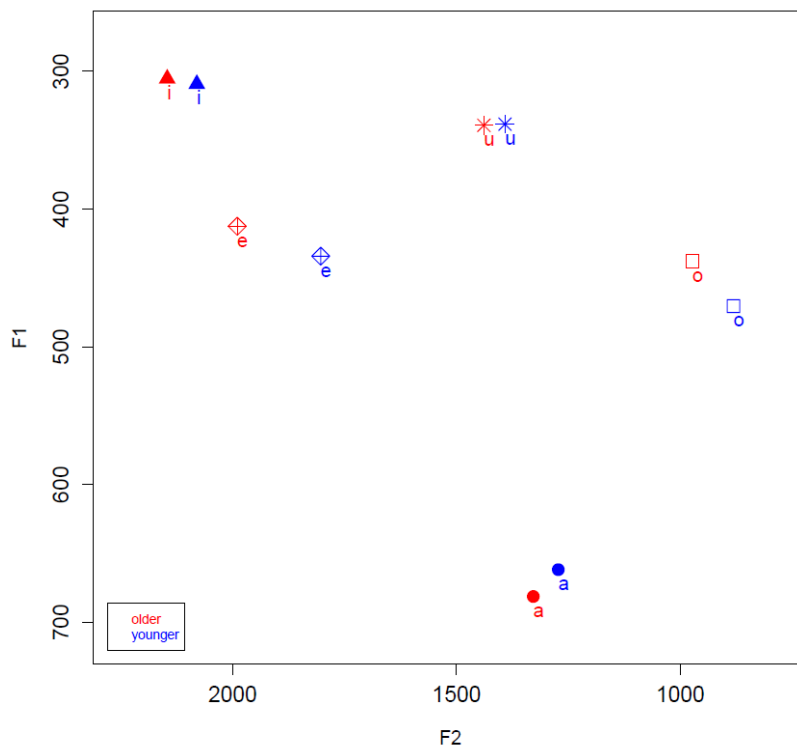


図 1 世代ごとに一括した各母音の平均値

また、世代ごとに区分した全サンプルは図 2 の通りである。

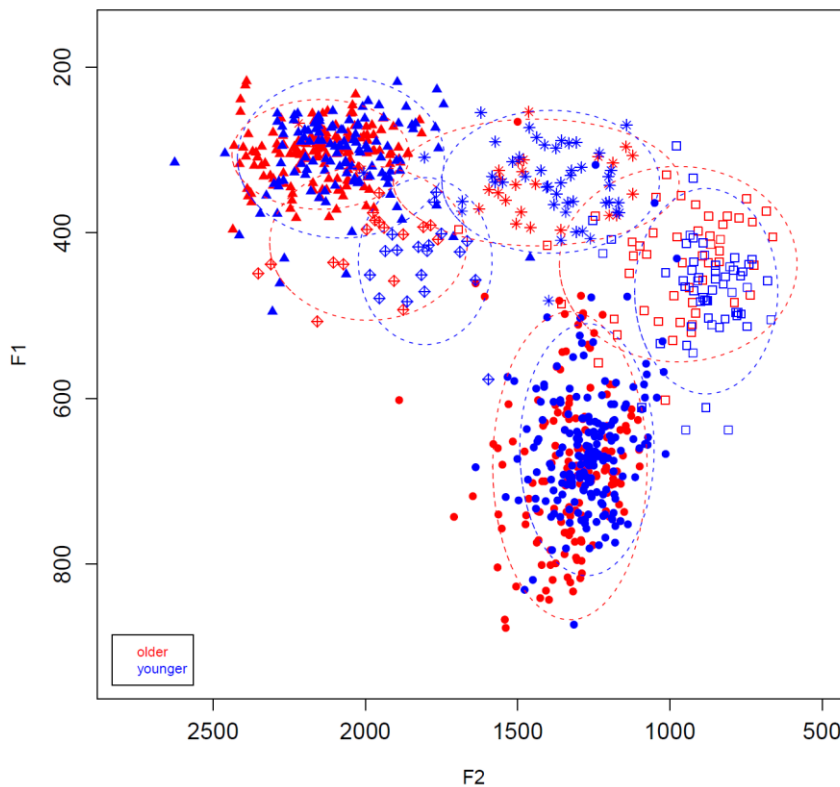


図2 世代ごとに区分した全サンプル

次に、各母音について分析を行う。アを例として説明すると、まず、各話者のアの平均値について、サンプル数が平均の2倍以上である話者が、F1・F2において最大値または最小値を取っていないかどうかを確認する（サンプル数の多い話者が極端な値を取っていないことを確認するため）。その上で、「話者ごとのサンプル数の偏りによる影響は少ない」と判断し、各世代ごと一括して扱って良いものと見なす（結果的には、すべての母音において、このように扱うことが出来ると判断した）。そして、「世代」・「後続母音」・「アクセント核の有無」の3つを要因とする分散分析を行う（ウ・エ・オについては、「アクセント核の有無」を考慮する必要がないので、「世代」・「後続母音」の2つを要因とする分散分析を行う）。

分散分析は、それぞれの要因が結果に影響を与えているかどうかを調べるものである。本研究においては、「世代」という要因（高年層か若年層か）が、フォルマント周波数の値に影響しているかどうかを調べることが最も重要である。計測値による分析の結果、「世代」という要因による有意差の有無は、表2のようになった（分析にはANOVA4 on the webを使用）。有意差があるものについてその大小を説明すると、「アのF1・F2は、若年層の方が高年層と比べて相対的に小さい」、「イ・エ・オのF2は、若年層の方が高年層と比べて相対的に小さい」ということになる。

表2 計測値による分散分析の結果

	ア	イ	ウ	エ	オ
F1	*	n.s.	n.s.	n.s.	n.s.
F2	**	**(**)	n.s.	**	**

※*は危険率 5%で、**は 1%で有意差が示されたもの。n.s.は有意差が無いもの。

※イの F2 は、「3 要因での分散分析 (2 要因での分散分析)」の順で結果を併記した²。

4.3. 正規化を施した上での分析

計測値での分析に続いて、正規化を施した上での分析を行う。母音の比較における「正規化」とは、男性と女性、あるいは成人と子供など、声道の長さに違いがあると考えられる集団を比較可能にするための、数学的な処理である。フォルマント周波数は声道の長さに影響され、声道が長くなればなるほどフォルマント周波数が小さくなるという関係がある。本研究では話者を男性に限定しているが、4.2節で示された有意差が、身長差によるものである可能性も、排除できない。そこで、本節では正規化を施した上で同様の分析を行い、それでも有意差が示されるのかを確かめる。

正規化には複数の手法があり、どの手法も長所と短所がある。それぞれの特徴について、Thomas (2011) を参考にした上で、本研究では、「バーク尺度に変換してその差を取る」という手法を採用した。

この手法により得られた数値を用いて、計測値での分析と同様に分析を行う (正規化の計算には NORM を使用)。正規化後の数値による分析の結果、世代による有意差の有無は表 3 のようになった。

表3 正規化後の数値による分散分析の結果

	ア	イ	ウ	エ	オ
F1'	n.s.	n.s.	—	*	**
F2'	n.s.	**(**)	—	**	n.s.

※イの F2 については、表 2 と同様。

※ウについては、この手法では適切な判断が出来ないため、空欄とした³。

² 世代 (2 種類)、後続母音 (5 種類)、アクセント核の有無 (2 種類) によってサンプルを分けるので、20 種類 (2×5×5) のパターンが必要である。しかし実際には、そのうちの 2 パターン (「高年層・アクセント核有・後続母音 a」と「若年層・アクセント核有・後続母音 i」) に該当するサンプルが 1 個も得られなかった。データがゼロのパターンが存在する場合、分散分析が難しい。そこで次善の策として、後続母音が a と i のサンプルをすべて除外した上での分散分析 (3 要因) と、アクセント核の有無という区別を撤廃して上での分散分析 (2 要因) を行った。

³ この手法では F3 の値を使用して正規化が行われる。この手法で正規化を施して 2 つの集団を比較する場合、F3 の値が集団間で大きく異なるということが前提になっている。本研究の計測時には予備的に F3 も記録したが、ウだけは世代によって F3 の値が大きく異なっていた。したがってこの手法では、ウについては適切に正規化できないと判断した。

5. 結論

計測値による分析（表2）と正規化後の数値による分析（表3）から、世代差が示唆されるフォルマントは、①計測値と正規化後の数値との両方で有意差が示されたもの
②計測値と正規化後の数値のどちらか一方で有意差が示されたもの、に二分できる。

①イのF2とエのF2は、計測値においても正規化後の数値においても、若年層のF2が小さくなるという傾向が見出された。

②アのF1・F2、エのF1、オのF1・F2は、計測値あるいは正規化後の数値のどちらか一方で世代差が示されており、世代差が存在する可能性がある。このグループは解釈が難しく、今後の課題としたい。特にオについては、F1とF2が接近するため計測が難しく、計測方法自体にも改良の余地があるかもしれない。

以上をまとめると、少なくともイとエのF2には世代差が存在する可能性が高い。どちらも若年層の方が相対的に小さい値を取っており、イとエが中舌化しつつあると解釈できる。

今後は、話者数を増やしたり、別の要因を設定したり、今回とは異なる正規化の手法を採用したりしても、このような世代差が示されるかを調べる必要がある。また、様々な条件を統制して分析するためには、インタビュー調査を新たに行ったり、これまでに蓄積された録音資料を用いたりすることも有効であろう。

参考文献

【使用データ及び分析に使用したサイト】

国立国語研究所・情報通信研究機構・東京工業大学「日本語話し言葉コーパス」第5刷
桐木建始（2002）“ANOVA4 on the Web” <https://www.hju.ac.jp/~kiriki/anova4/>（2017年7月31日最終アクセス）

Thomas, Erik R. and Tyler Kendall（2007）“NORM: The vowel normalization and plotting suite”
<http://lingtools.uoregon.edu/norm/index.php> (accessed July 31, 2017)

【欧文文献・邦文文献】

Thomas, Erik（2011）*Sociophonetics: An introduction*. Basingstoke: Palgrave Macmillan.

井上史雄（1989）「子音の発音の変化」 杉藤美代子（編）『講座日本語と日本語教育 2—日本語の音声・音韻（上）』109-134. 東京：明治書院.

柴田実（2005）「平成の発音変化」『月刊言語』34(1), 64-72.

杉藤美代子（1997）「日本語音声の音声学的特徴」『BME（生体医工学）』11(4), 2-6.

母音や摩擦音が後続する撥音の異音について

松井 理直 (大阪保健医療大学)
michinao.matsui@ohsu.ac.jp

キーワード：エレクトロパラトグラフィ (EPG)、撥音、変異音、同化、音節

1. 研究の目的

日本語の撥音は、その音声環境に従って様々な変異音を持つ。特に、破裂口音や破裂鼻音あるいは弾き音といった閉鎖的性質を持つ子音が撥音に後続する場合、その撥音は後続子音の調音位置に関する逆行同化を受けるという点で、全ての先行研究の見解が一致している。しかし、母音や摩擦音といった閉鎖的性質を持たない分節音が後続した撥音がどのような変異を起こすかという点については、必ずしも意見の一致が見られない。

そこで本研究は、母音や摩擦音が後続した撥音の変異について、エレクトロパラトグラフィ (EPG) を用いた調査結果の報告を行う。最初にいくつかの先行研究を概観した後、EPGを用いた実験結果を示す。結論として、撥音に母音が後続した場合には逆行同化が起こるとは限らず、前後の音声環境のうち、より口腔空間の狭い分節音の影響を受ける易いを述べる。また、口腔内の舌運動と口唇の運動が独立した影響を及ぼすことも示す。最後に、この実験結果から分かる日本語の音節構造について、若干の議論を行う。

2. 撥音変異に関する先行研究の概観

2.1 撥音が中線的接触を持つ子音に先行する場合

撥音に破裂口音や破裂鼻音あるいは弾き音といった閉鎖的性質を持つ子音が後続する場合に、撥音が起こす変異については、全ての先行研究で見解が一致している。すなわち、その撥音は後続子音から調音位置の情報を受け取り、逆行同化を起こす。

- (1) a. 両唇破裂 [p], [b], [m] が後続する場合：撥音は [m] に変異。
- b. 歯茎破裂 [t], [d], [n], [ts], [dz], 歯茎弾き音 [r] が後続する場合：撥音は [n] に変異。
- c. 歯茎硬口蓋破擦音 [tʃ], [dʒ] が後続する場合：撥音は [nʲ] あるいは [ɲ] に変異。
- d. 軟口蓋破裂 [k], [g], [ŋ] が後続する場合：撥音は [ŋ] に変異。

ただし、ラ行子音が後続する場合については、この子音自体が様々な変異音を持つことを反映して、意見が一致するとは限らない。例えば、Vance (2008) は [r] 音に先行する撥音は鼻音化接近音 [ɑ̃], [ɪ̃] に変異するという。また 益子 (2009) は、ラ行子音の異音として [l] 音も考慮に入れ、この音が撥音に後続する場合には [l̃] になると指摘している。接近音や側面接近音が調音方法の一種であることを考慮すると、これらの指摘は非常に興味深い。撥音が調音位置の逆行同化のみならず、ある種の調音方法に関する情報の逆行同化も受けることを示唆しているからである。こうしたラ行子音の問題を置いておくと、破裂音・破裂鼻音・弾き音そして歯茎側面音はいずれも中線的接触を持つ調音であり、これらの子音が撥音に後続する場合は、少なくとも調音位置に関する拘束変異が起こるとまとめておいてよいだろう。

2.2 撥音が末尾に位置する場合

これに対し、撥音が末尾に位置しており、他の分節音が後続しない環境では、撥音の変異について見解が食い違う。例えば、藤村・大泉 (1972) は [N] に変異するとし、斎藤 (2006) は「本」や「缶」では [N] だが、「銀」や「千」では [ŋ] としている。また、村木・中岡 (1989) は「口蓋垂と奥舌面の間が狭まるが、口腔内の閉鎖がない継続鼻音」と述べており、これは [ŋ] か [ɳ] に相当する音と見てよいだろう。同様に Vance (2008) も、撥音変異に接近音の鼻音化 [ŋ] があり得ることを指摘している。

末尾の撥音に関しては、吐師ほか (2016) による生理学的手法を用いた精密な研究が極めて興味深い。彼らは、X 線マイクロビームを用いた実験により、末尾に位置する撥音変異が [N] に限定されず、自由度が高く、話者間変動も大きいことを見いだした。先行研究における記述のばらつきも考慮すると、末尾に位置する撥音は自由変異であり、その自由度は、Itô & Mester (1995) が指摘した「撥音の調音位置は過小指定されている」という性質と関係している可能性がある。ただし、超音波測定を用いて撥音変異を実験的に調べた Yamane (2013) によると、こうした撥音変異の自由変異も、軟口蓋に明確な調音位置の指定を持っている [k] 音に比べて自由度が高いわけではないという。この指摘は、撥音の性質を考える上で見逃せないものの 1 つである。

2.3 撥音が摩擦子音に先行する場合

撥音に摩擦音が後続する環境における変異については、益子 (2009) の記述が非常に詳しい。益子は後続摩擦音に従って鼻母音に変異すると見ており、[s], [ç], [z], [ʒ] が後続する場合は [ĩ] に、[ç] が後続する環境では [ĩ] に、[ʃ] が後続する時には [ɳ] に変異するとしている。興味深いのは [h] 音に先行する撥音の変異で、これは [h] 音に後続する母音に従って変異音が決まるという。声門摩擦音は口腔の調音が指定されておらず、[h] 音の調音時点で口腔の構えは後続母音と同一のものになるため、これも通常の逆行同化とほぼ同様の現象であると見てよい。

摩擦音に先行する環境で撥音が鼻母音に変異するという点は、多くの先行研究が一致するところであるが、微妙な点で相違も見られる。服部 (1951) は、摩擦音に先行する撥音は後続摩擦音に同化し、鼻音化摩擦音になる (ただしその摩擦的噪音がなくなり、鼻母音になることもある) と指摘した。同じく Kochetov (2014) も、[s] 音に先行する撥音変異を EPG によって測定し、[ĩ] か [z] に変異すると述べている。これらの現象は、前述した [ĩ] 音への同化と同様、撥音は単に調音位置の逆行同化を受けるだけでなく、調音方法の逆行同化も受けることを示しているのだろう。

しかし村木・中岡 (1989) によれば、摩擦音に先行する撥音変異はもう少し複雑なようだ。彼らは、撥音に摩擦音が後続した場合、「口構えは前後の音に従い、舌は口蓋に向かって持ち上げられるが、口腔内の閉鎖がない継続鼻音 (すなわち鼻音化摩擦音か鼻母音)」に変異するという。「前後の音に従い」という点に注目されたい。すなわち彼らによれば、撥音は逆行同化のみならず、順行同化の影響を受ける可能性も持つ。これは、次節で述べる母音が後続する環境における撥音変異でも同様に問題となる。

2.4 撥音が母音に先行する場合

母音が撥音に後続する場合、ほぼ全ての研究が撥音は鼻母音に変異するとしており、この点では違いがない。大きく異なるのは、順行同化か逆行同化かという点である。服部 (1951) や 益子 (2009) は、撥音が後続母音に従って変異すると見なす。斎藤 (2006) は撥音変異の全体的な説明として「後続音に同化」と明記しているが、具体例を見ると、「線を」が [sẽ.o]

となっており、例は少ないものの順行同化もあり得ると考えているようだ。これに対し、村木・中岡 (1989) は「口構えは前後の音に従い、舌は口蓋に向かって持ち上げられるが、口腔内の閉鎖かがない鼻音 (継続鼻音すなわち鼻音化摩擦音か鼻母音)」と述べ、「全員 [zečii]」「禁煙 [kiieč]」、「安易 [aai]」といった発話のスペクトログラムを示して、順行同化の例を多く挙げている。

この点で注目すべきは、Vance (2008) であろう。彼は、撥音が母音・接近音・流音・摩擦音の前で鼻母音に変異すると共に、環境や話者によって先行母音に同化したり、「調音上さらに狭窄の狭い音」に自由変異する可能性について指摘した。次節で述べる実験結果から、本研究も「調音動態としてより狭窄が狭い」という条件こそが撥音変異の本質であると考えられる。

2.5 音節構造内のスロット表示を巡って

撥音が逆行同化の影響のみを受けるのか、あるいは順行同化も影響するのかという問題は、音韻論における撥音のスロット表示とも深い関係を持つ。一般に、日本語の長音や促音は Obligatory Contour Principle (OCP, McCarthy 1986) に違反しない形で、

長音 (例：カート)	C	V	V	C	V	促音 (例：カット)	C	V	C	C	V
	/k/	/a/	/a/	/t/	/o/		/k/	/a/	/t/	/o/	/o/

のような構造を持つ。では、同じ特殊拍である撥音はどのような表示を持つのだろうか。「ミント」に見られる逆行同化は、撥音が基底から [+nas] 素性を持っているという点を除き、基本的に促音と並行的な関係なので、(2) のような構造を持つと考えてよいかもしれない。すなわち、「官記」で逆行同化が起こるのは撥音 ([+nas] 素性) が C スロットに結びついており、かつ [cont] 素性や調音位置に関する素性を持っていないため、後続の C スロットに結びついている [cont] 素性や調音位置素性が拡張される。しかし、「官位」のように撥音の前後が母音だった場合、撥音の“C”するとは同化を引き起こす力にならない。そうすると、何が同化を引き起こすのだろうか。本研究の目的は、こうした問題も明らかにするために、母音や摩擦音が後続する場合の撥音変異の詳細を EPG を用いて調査する点にある。

(2) 撥音 (例：官記) の構造

a.	C	V	C	C	V
	/k/	/a/	[+nas]	/k/	/i/

「官位」の構造？

b.	C	V	C	V
	/k/	/a/	[+nas]	/i/

3. 摩擦音／母音に先行する撥音変異に関する実験

3.1 実験方法と分析方法

実験は被験者 4 名 (東京方言女性 1 名、関西方言女性 1 名、関西方言男性 2 名) に対し、EPG および鼻腔用マイクと口腔用マイクを装着させ、3 モーラの刺激語をランダムに 5 回ずつ発話させる形で行った。本論文で分析の対象とする刺激語のパターンは以下の通りである (これ以外のデータも採ったが、紙面の関係上、本稿では述べてない)。有意味語と無意味語が混ざっているため、EPG を装着する前に事前に全ての刺激語を順に 5 回連続で音読させ、無意味語に慣れてもらった。またアクセントは平板型で読むように指示しているが、起伏型で読まれた場合に言い直し等はさせていない。また、正中面からビデオカメラによって口唇の動きを撮影した。

- (3) a. 母音-撥音-母音: 「あんい」「あんう」... 「おんお」
 b. /k/-母音-撥音-母音: 「かんい」「かんう」... 「こんお」
 c. 母音-撥音-/h/-母音: 「あんは」「あんし」... 「おんほ」
 d. /k/-母音-撥音-/h/-母音: 「かんは」「かんし」... 「かんほ」

撥音が母音に挟まれている環境では、いずれの母音も鼻音化する可能性があるため、撥音区間の認定は、鼻腔用マイクおよび口腔用マイクから収録した音とスペクトログラムを手がかりに行った。なお、撥音が母音に挟まれている場合でも、図 refN-inter (a) の「かんだ」と類似した撥音区間を持つ図 refN-inter (b) のようなパターンは、撥音部が鼻子音 [ŋ], [N] であると見なし、図 refN-inter (C) のように撥音部であっても口腔用マイクが完全に音声を拾っている場合を鼻母音であると見なす。

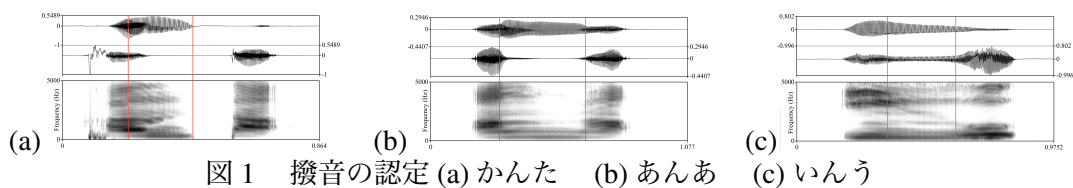


図1 撥音の認定 (a) かんだ (b) あんあ (c) いんう

EPG データは、4名の被験者に対し人工口蓋床を各被験者ごとに用意し、サンプリング周期 10 ms で収集した。この人工口蓋床は山本一郎氏によって調音への負担が最小限になるように開発されたもので、異なる話者の調音を比較できるように標準化され、前後方向に歯茎2列、後部歯茎2列、硬口蓋3列、軟口蓋境界部1列の電極配置を持つ。また、左右方向は歯茎最前列のみ電極が6点、他の列は電極が8点配置され、歯茎から軟口蓋境界部まで計62点の計測が可能となっている。なお、EPG データの収録には WinEPG (Articulate Instruments Ltd.) およびタブレット端末を用いて行った。

3.2 分析結果

分析例として、「あんい・いんあ」「あんう・うんあ」のスペクトログラムと EPG 遷移の結果を図2に示す。こうした広母音・狭母音の組み合わせにおいては、先行/後続環境という違いに関わらず、撥音部はほぼ狭母音の鼻音化になっていることが分かる。

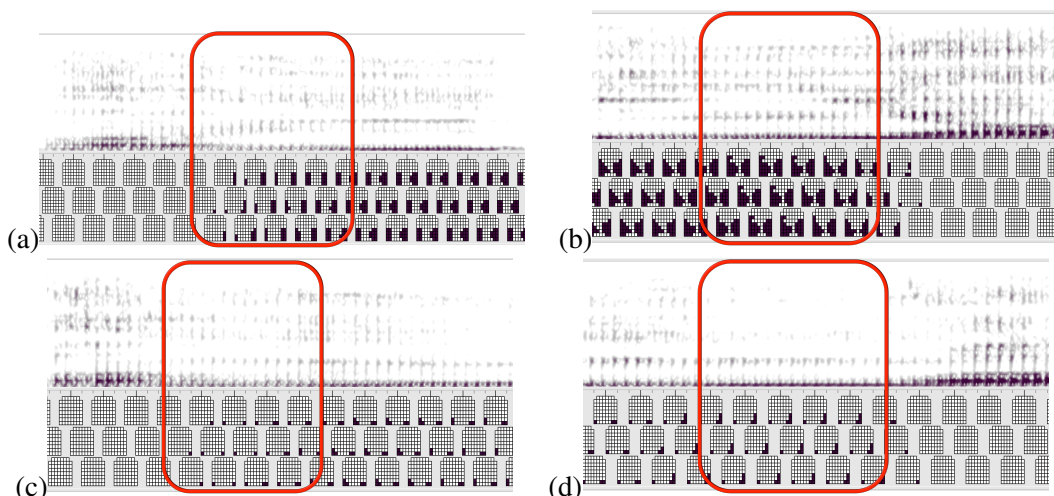


図2 EPG 遷移の結果 (a) あんい (b) いんあ (c) あんう (d) うんあ

撥音に先行／後続する母音環境と撥音部の鼻母音変異の結果を、表1にまとめておく。基本的に、撥音の前後にある母音のうち、口腔空間のより狭い母音が鼻母音化していることが分かる。また、/i/, /u/ という狭母音に撥音が挟まれた場合は、基本的に [i] に変異する傾向が強い。これは、/i/ 音が palatal という性質を明確に持つのに対し、/u/ 音は強い素性を持たないことに起因しているように思われる。これに対し、/e/, /o/ という半狭母音に撥音が挟まれた場合は、先行母音に引きずられた鼻音に変異しやすい。この原因は明確ではないが、音節境界を越える逆行同化よりも、音節内で起こる順行同化を好む可能性が考えられる。この点については、今後の課題としたい。なお表中の“?” は、EPG の反応が取れないために明確なことがいえないことを示す。また、「母音-撥音-ハ行子音-母音」の連鎖についても、表1とほぼ同一の結果が得られた。これはハ行子音の調音動態から言っても当然の結果であろう。

表1 撥音部における鼻母音の音価
後続母音

		[a]	[i]	[u]	[e]	[o]
先 行 母 音	[a]	?	[i]	[u]	[e]	?
	[i]	[i]	[i]	[i]	[i]	[i]
	[u]	[u]	[i]	[u]	[u]	[u]
	[e]	[e]	[i]	[u]	[e]	[e]
	[o]	NC	[i]	[u]	[o]	?

ハ行子音に対し、「母音-撥音-/s/-母音」「/k/-母音-撥音-/s/-母音」の連鎖では Kochetov (2014) の結果とほぼ同一の結果が得られ、服部 (1951) が指摘するように撥音部は [ʃ], ʃ に変異していた。その一例を図3に示す。このことは、破裂音と同じく /s/ 音が後続する場合にも調音位置の逆行同化が起こると共に、[cont] 素性についても同化を起こしていることを示す。この性質は、基底における撥音が [+nasal] 素性のみ指定されているという Itô & Mester (1995) の指摘とも整合的である。

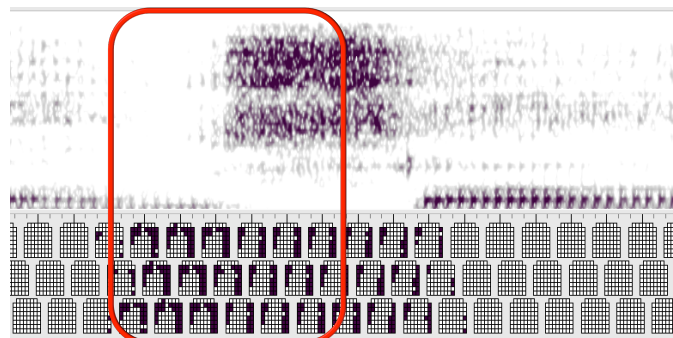


図3 「あんさ」における EPG 遷移の結果

4. 総合論議

EPG を用いた撥音変異の実験から、撥音の変異は前後に存在する音韻環境のうちで口腔空間のより狭いものに強く影響されるということが明らかとなった。撥音に子音が後続する時に逆行同化を起こすのも、日本語の phonotactics として撥音に先行する音は必ず母音であり、撥音に後続する子音のほうが必然的に口腔空間がより狭いことが原因であると考えられるだろう。すなわち、撥音の逆行同化は 2.5 節で議論したような CC-slot 連鎖が引き起こす

ものではない。このことは、促音の逆行同化の再検討にもつながるだろう。すなわち、促音における逆行同化も CC-slot 連鎖が引き起こすものではない可能性がある。促音も撥音と同じく「口腔空間のより狭い環境に影響される」とすれば、促音の場合は（語末促音を除き）必ず後続環境が子音であるため、必ず逆行同化が起こることも不思議ではない。

この議論は、日本語の音韻表示において撥音や促音に C-slot が必要かという問題につながる。「口腔空間のより狭い環境に影響される」という性質は、音節末という位置が要求する sonority の低さの反映と考えることも可能であろう。この場合、C-slot は不要となり、撥音が「鼻母音」に変異することとも自然に説明がつく。一方、C-slot を仮定する場合、鼻母音への変異は‘C’を「プロトタイプカテゴリー」と考えない限り、矛盾が起こる。この点については、促音が基底で featureless なのかといった問題とも絡むため、また稿を改めて議論を行いたい。

謝辞：本研究を実施するにあたり、人工口蓋床を作成してくださった山本一郎先生に深く感謝いたします。なお、本研究は科学研究費・基盤研究 (C)「音声知覚における摩擦性極周波数特性の影響に関する総合的研究」の援助を受けました。

参考文献

- 藤村 靖・大泉 充郎 (1972) 『音声科学』 東京大学出版会. 東京.
- Itô, J., & Mester, A. (1995) “Japanese phonology.” In J. A. Goldsmith (Ed.), *The handbook of phonological theory*. (pp. 817–838). Blackwell. Oxford.
- Kochetov, A. (2014). *Japanese in the typology of nasal place assimilation: Electropalatographic evidence*. (Formal Approaches to Japanese Linguistics (FAJL) 7, Handout).
- 益子 幸江 (2009) 「音声記号」 今泉 敏 (編) 『言語聴覚士のための音声学・言語学』 20–36, 医学書院. 東京.
- McCarthy, J. J. (1986) “OCP effects: Gemination and antigemination.” *Linguistic Inquiry*, 17, 207–263.
- 村木 正武・中岡 典子 (1989) 「撥音と促音—英語・中国語話者の発音—」 杉藤 美代子 (編) 『講座 日本語と日本語教育第 3 巻』 139–177, 明治書院. 東京.
- 斎藤 純男 (2006) 『日本語音声学入門』 三省堂. 東京.
- 服部 四郎 (1951) 『音声学』 岩波書店. 東京.
- 吐師 道子・小玉 明菜・三浦 貴生・大門 正太郎・高倉祐樹・林 良子 (2016) 「日本語語尾撥音の調音実態：X線マイクロビーム日本語発話データベースを用いて」 『音声研究』 18: 2, 95–105.
- Vance, T. J. (2008) *The sound of Japanese*. Cambridge University Press. New York.
- Yamane, N. (2013). ‘Placeless’ Consonants in Japanese: An Ultrasound Investigation. Unpublished doctoral dissertation, The University of British Columbia, type = Ph. D. dissertation,.

日本語を母語とする英語学習者におけるシュワーの削除に関する研究

坂本 洋子 (獨協医科大学)

y-saka@dokkyomed.ac.jp

1. はじめに

英語母語話者の発話において、シュワー (ə) の削除が起こることが報告されている。先行研究によると、シュワーの削除を引き起こす条件として、強勢の位置 (強勢の前か、後か)、語彙頻度、形態的複雑性、発話速度、スピーチスタイル、話者の性別、方言などが提案されている (Dalby, 1986; Patterson et al, 2003)。これらの条件の内、発話速度に関しては、通常速度よりも速い速度の方がシュワーの削除の頻度が高くなり、速い速度では 35% から 45% の削除率であったと報告されている。また、強勢の位置は強弱弱の 3 音節から成る語がシュワーの削除率が最も高いことが示されている。それでは日本語母語話者は英語を発話する際に、シュワーの削除に関して、英語母語話者と同様の傾向を示すのであろうか。本研究では、日本語を母語とする英語学習者を対象として、単語の読み上げと文の読み上げを行う 2 つの実験を行った。

2. 実験 1

実験 1 では日本語を母語とする英語学習者における単語内のシュワーの削除を調べることを目的とする。分析対象はシュワーの削除が生じる可能性のある語を 2 種類の速度で発話した音声である。

2.1. 実験方法

2.1.1. 言語材料

本研究ではシュワーの削除が起こる可能性のある単語 53 語をテスト語とした。シュワーの削除が起こる可能性のある単語は Patterson et al. (2003) で用いられた単語のうち、日本語を母語とする大学生に馴染み度の高い単語を用いた。単語の種類は音節数と強勢の位置に基づく 3 種類である。強弱のリズムを持つ 2 音節語 (2PRE)、弱強弱のリズムを持つ 3 音節語 (3PRE)、強弱弱のリズムを持つ 3 音節語 (3POST) である。単語の例はそれぞれ、select /sələkt/ (2PRE), solution /səlu:ʃən/ (3PRE), avenue /évnə(j)ù:/ (3POST) である。

2.1.2. 被験者

被験者は日本語を母語とする大学生 16 名であった。男性は 6 名、女性は 10 名であった。学年は 1 年生から 4 年生までであり、専攻は英語 13 名、交流文化学科 2 名、フランス語 1 名であった。学年は 1 年生から 4 年生までであり、平均年齢は 20 才であった。出身は日本であり、英語圏の海外滞在経験がない被験者は 6 名、英語圏の海外滞在経験がある被験者

は 10 名であり、期間は 2 週間から 1 年であった。TOEIC の点数は 400 点から 810 点であり、平均は 738 点であった。

2.1.3. 実験手順

実験手順は Dalby (1986) と氏平・窪菌(2000)の方法を参考に、以下の手順で行った。被験者は 1 人ずつ静かな部屋で実験に参加した。被験者はまずテスト語 53 語が書かれた紙を渡され、単語に慣れるように指示された。その際制限時間は設けなかった。ランダムに配置した 53 語のテスト語を一まとめにして 1 回とし、テスト語の読み上げを各被験者について 5 回行い、IC レコーダーを使って録音した。そしてこの 5 回の録音のうち第 2 回目から第 5 回目までの 4 回のテスト語の録音を測定の対象とした。読み上げの速度は普通と意味が取れる範囲でできるだけ速い速度とを 1 回ごとに交互に行うように指示された。

シュワーの削除が起こっているかどうかの判断は、Praat を用いて行った。シュワーの削除は音声的な現象であり、段階的(*gradient*)であるが、今回は削除が完全に起こっている場合のみを扱うことにした。シュワーの削除の判断基準は、聴覚による確認と波形の視覚による確認を行い、シュワーが完全に消えているものとした。これは Dalby (1986) と同じ基準である。

2.2. 実験結果

16 名の被験者の音声データを分析し、シュワーが削除されているテスト語の数を数えた。回数ごとのテスト語数は 53 語×16 名で合計 848 語であった。アクセントや発音を間違えたものは誤答とした。16 名の被験者の録音回数ごとのシュワーの削除数を表 1 に示す。

表 1：テスト語読み上げにおけるシュワーの削除数と誤答数

	シュワーの削除数	誤答数
練習	5	22
普通 1	6	17
速い 2	6	17
普通 3	4	19
速い 4	9	16

2.3. 議論

実験 1 の結果より、テスト語の読み上げにおいて、日本語母語話者はシュワーの削除率が最も高い速い 4 で 1%、普通 3 で 0.4%であり、ほとんどシュワーの削除を行っていなかった。また発話速度による違いや、音節数と強勢の位置による単語の種類による違いは見られなかった。このことは英語母語話者と日本語母語話者では異なるメカニズムでシュワーの削除を行っている可能性を示唆するものであると考えられる。

3. 実験 2

実験 2 では、日本語母語話者における文中のシュワーの削除を調べることを目的とする。分析対象は、シュワーの削除が生じる可能性のある語を組み込んだキャリア文を 2 種類の速度で発話した音声である。

3.1. 実験方法

3.1.1. 言語材料

実験 1 でテスト語として用いられた単語 53 語を埋め込んだキャリア文 53 文を言語材料とする。シュワーの削除が起こる可能性のある単語は Patterson et al. (2003) で用いられた単語のうち、日本語を母語とする大学生に馴染み度の高い単語を用いた。単語の種類は音節数と強勢の位置に基づく 3 種類である。強弱のリズムを持つ 2 音節語 (2PRE)、弱強弱のリズムを持つ 3 音節語 (3PRE)、強弱弱のリズムを持つ 3 音節語 (3POST) である。単語の例はそれぞれ、select /sələkt/ (2PRE), solution /səlu:ʃən/ (3PRE), avenue /ævən(j)ù:/ (3POST) である。キャリア文は『SCN 小学館コーパスネットワーク』の British National Corpus から検索した文を、英語母語話者が確認した上で使用した。キャリア文は日本語を母語とする大学生に馴染み度の高い語で構成されている。

3.1.2. 被験者

被験者は実験 1 に参加した日本語を母語とする大学生 16 名である。実験 1 が終わった後すぐに実験 2 に参加した。

3.1.3. 実験手順

実験手順は基本的には実験 1 と同じである。被験者はまずキャリア文 53 文が書かれた紙を渡され、文に慣れるように指示された。その際制限時間は設けなかった。ランダムに配置した 53 文のキャリア文を一まとめにして 1 回とし、キャリア文の読み上げを各被験者について 5 回行い、IC レコーダーを使って録音した。そしてこの 5 回の録音のうち第 2 回目から第 5 回目までの 4 回のキャリア文の録音を測定の対象とした。読み上げの速度は普通速度と意味が取れる範囲でできるだけ速い速度とを 1 回ごとに交互に行うように指示された。

シュワーの削除が起こっているかどうかの判断は、Praat を用いて行った。シュワーの削除は音声的な現象であり、段階的 (gradient) であるが、今回は削除が完全に起こっている場合のみを扱うことにした。シュワーの削除の判断基準は、聴覚による確認と波形の視覚による確認を行い、シュワーが完全に消えているものとした。これは Dalby (1986) と同じ基準である。

3.2. 実験結果

16 名の被験者の音声データを分析し、シュワーが削除されているテスト語の数を数えた。回数ごとのテスト語数は 53 語×16 名で合計 848 語であった。アクセントや発音を間違えた

ものは誤答とした。16名の被験者の録音回数ごとのシュワの削除数を表2に示す。

表2：キャリア文読み上げにおけるシュワの削除数と誤答数

	シュワの削除数	誤答数
練習	11	18
普通1	15	20
速い2	13	22
普通3	13	16
速い4	13	20

3.3. 議論

実験1と比較するとキャリア文読み上げにおいて、やや削除数の増加が見られた。しかしながら、最も削除数が多い普通1の15でも削除率としては1%であり、Dalby (1986)による英語母語話者の文の読み上げにおける普通の速度の削除率6%と比べて少ない傾向が見られた。さらに先行研究では、英語母語話者は速い速度での文の読み上げでは削除率がさらに43%と高くなる傾向が見られたが、今回の結果は異なる傾向を示している。このことは、英語母語話者は、文の読み上げにおいて、速い速度ではよりリラックスした、カジュアルな話し方になるのに対し、日本語母語話者は速い速度であっても、模範的な発音を行おうと注意深く発音していた可能性があると考えられる。また、日本語母語話者は速い速度の時であっても、英語母語話者の場合ほど、速度の差が出ていなかった可能性が示唆される。

4. 結論

本研究の結果から、日本語を母語とする英語学習者は、英単語と英文の読み上げ時にシュワの削除をほとんど行っておらず、発話速度や強勢の位置による差は見られないことが明らかになった。ゆえに、シュワの削除に関して、日本語母語話者は英語母語話者とは異なるメカニズムを持つ可能性が示唆されると考えられる。

参考文献

- Dalby, J. (1986). "Phonetic structure of fast speech in American English," Indiana University Linsuist Club, Bloomington.
- Patterson, D., LoCasto, P. C and Connine, C. M. (2003). "Corpora analyses of frequency of schwa deletion in conversational American English," *Phonetica* 60, 45-69.
- 氏平明・窪菌晴夫(2000)「日本語における母音長の中和について」『第14回日本音声学会全国大会予稿集』133 - 138.
- SCN 小学館コーパスネットワーク <http://www.corpora.jp/>

神奈川県三浦市方言における老年層話者のアクセントについて

坂本 薫（國學院大學大学院生）

lukun84@icloud.com

1. はじめに

本発表では、対面調査によって得られた神奈川県三浦市方言の老年層話者のアクセントについて述べる。神奈川県の方言は、俚言の蒐集と方言区画の設定に関しては豊富な先行研究があるが、音声、音韻の面ではまだその実態を明らかにする余地が残されている。坂本（2017）では、三浦市方言の老年層話者の発音の観察から、伝統的な音声・音韻的特徴について言及した。その結果を受けて、アクセントにおいても、当該方言の古相の保持が期待されたため、同一の話者を対象に調査を続行している。本発表は調査の結果から、アクセントの体系と名詞と動詞、形容詞の活用形のアクセントについて述べる。

2. 三浦市の位置

2. 1 地理的概況と産業、交通

神奈川県東南部の三浦半島の最南端に位置する、人口約4万5千人の市である。北部を横須賀市と接している。北部は農村地帯で畑作が主で、南部沿岸は江戸時代より漁港として発達。江戸時代は天領。1955（昭和30）年、二町一村が合併、三浦市として市制施行。今回の調査地点である三浦市三崎は、三浦半島の南西部、相模湾沿岸に位置する。

鉄道が三浦市内まで延伸したのは1966（昭和41）年。1975（昭和50）年、市内北西部に三崎口駅が開業するが、市の中央部には到達しておらず、以南の交通はバス路線が主となる。京浜急行線で品川駅～三崎口駅間の所要時間はおよそ65分である。

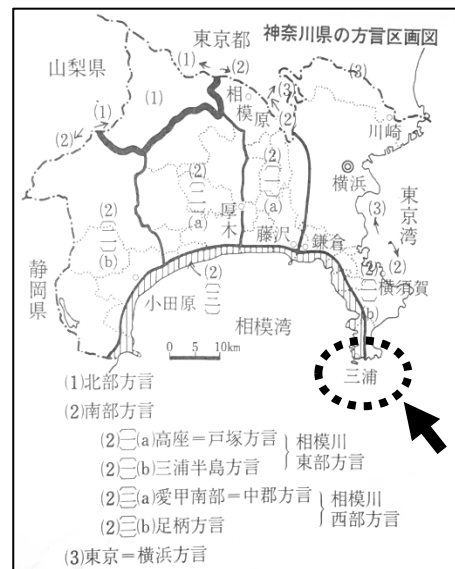


図1「神奈川県方言」『現代日本語方言大辞典』より。矢印及び点線は筆者による。

2. 2 方言区画上の位置

全国方言の区画における位置づけとして、東條操（1954）によると、神奈川県方言は東部方言に含まれる関東方言の、さらに西関東方言に含まれると述べられている。

神奈川県内の方言区画は、日野資純（1952、1964）等で述べられている（図1）。三浦市方言は、図中（2）の南部方言に含まれる相模川東部方言の下位区分である三浦半島方言に属する。ただし、西部沿岸部は相模湾沿岸方言の区画にも含まれる。

坂本（2017）では、三浦市で生育した話者を対象にした調査から、三浦半島方言の音声的特徴である[ʒe]と、相模湾沿岸方言の音声的特徴である[je]、明治期東京のアクセントと共通する「～ヤガレ」形式のアクセントが保持されていることについて言及した。

3. 調査について

調査は2017年3月～7月、三浦市内で行った。話者は三浦市南部、旧三浦郡二町谷で生まれ育った81歳男性（調査当時）。現役時代は市内で会社員として勤務。外住歴は無い。両親ともに現三浦市域で生まれ育っている。

調査は対面式で、調査項目は名詞、動詞、形容詞とその活用形。名詞語単独と「が」「を」「に」等を付けた、もしくは助詞が脱落した文の発音を観察した。動詞と形容詞については、語単独の発音と、活用形を含む文を作成してもらい、発音してもらった。調査対象は平山輝男（1960）所収の「全国アクセント比較表」に掲載されている語が中心で、これに多拍語を加えた。

4. 調査結果と考察

4. 1 方言的特徴の概観¹

調査により観察された方言的特徴を挙げる。語中語末のガ行音は規則的に鼻濁音[ŋ]が観察された。また、無声子音に挟まれた狭い母音の無声化は頻繁に生じるが、それに伴うアクセントの上がり目の移動は規則的ではなく、たとえば「吹く」[φɸɪˈkw]のように無声化した拍の後に上がり目が来る発音もしばしば観察された。連母音は「ムネータ（胸板）」や「カンゲール（考える）」、「カー（買う）」、「アエー（青い）」など、ai、ae、au、oiには融合が観察されたが、「アツイ（熱い）」「ウスイ（薄い）」などuiの融合は観察されず「カイー（痒い）」などごく少数の語にしか見られなかった。このほか、ヒ→シの交替「シッパル（引っ張る）」や、動詞語末のラ行の撥音化「トンソーダ（取るそうだ）」に加え、「コイ（これ）」「ソイ（それ）」や「デライタ（出られた）」など、一部の語にレ→イの交替が観察される。

4. 2 アクセントの体系

東京語のアクセント同様、アクセントの「上がり目」の有無、そしてその位置が弁別の特徴であり、その対立はn拍の語に対してn+1種ある。表1に3拍語までのアクセントの体系を示した。

表1：三浦市方言のアクセントの体系

拍	音韻論的解釈	語例
1	/O/	柄、蚊、毛、子、血…
	/Oɪ/	絵、尾、木、酢、田、手、荷…
2	/OO/	飴、蟻、烏賊、梅、釘、行く、着る…
	/OOɪ/	石、歌、北、寺、足、鍵…
	/OɪO/	斧、秋、汗、婿、夜、会う、見る、良い…
3	/OOO/	欠伸、錨、鎖、衣、魚、机、遊ぶ、借りる…
	/OOOɪ/	小豆、力、頭、扇、鏡、刀、瓦…
	/OOɪO/	五つ、さざえ、油、余る、起きる、赤い、黒い…
	/OɪOO/	朝日、嵐、もみじ、帰る、通る…

4. 3 名詞のアクセント

1～3拍の語のアクセントを、類別語例を中心に表2に示した。表中の1, 2拍の名詞に注目

¹ 音声の表記は片仮名と一部IPAを用いた。アクセントの上がり目は記号「ɪ」で示した。以降のページも同様。

すると、たとえば1拍Ⅰ類の「帆」や、2拍Ⅲ類の「神」、「雲」で（この2語は頭高型のアクセントも観察されたが）、それぞれ伝統的なアクセントが保持されているため、語の類で見たときに例外が少ないように見える。なお、尾高型の名詞に助詞「の」が接続する際、東京のアクセントは「ハナノイロガ…（花の色が…）」のように平板に発音されると言われるが、今回の調査では言われている通り平板に発音される場合もあったが、「ウラノハタケデ…（裏の畑で…）」のように下がり目が観察される場合もあった。これは3拍以上の名詞についても同様である。

3拍の名詞については、尾高型のアクセントと中高型のアクセントに特徴が見える。尾高型のアクセントについて、例えばⅡ類「小豆」、「毛抜き」、Ⅳ類「いたち」、「林」、「筵」などは東京のアクセントの古相と共通する。「とかげ」、「くらげ」に関しては方言独特のアクセントの可能性もある。中高型のアクセントが多く観察される点については日野資純（1984）などでも指摘されている通り、神奈川県方言の一つの特徴であると言える。Ⅲ類の「小麦」や「栄螺」、Ⅴ類の「油」や「心」など、東京のアクセントの古相と共通するものもあれば、「茸」「マミヤ（眉）」など、地域的特色のあるアクセントと考えられる語もある。ちなみに、神奈川県方言の特色として、中高型の「南瓜」のアクセントが知られているが、話者は方言としては俚言の「トーナス」を使用しており、語形「カボチャ」は後から学習した語であったためか共通語と同じ平板型のアクセントであった。

表2：三浦市方言老年層話者 1～3拍名詞のアクセント

〈 〉内の数字は類。表中*は頭高型も。**は尾高型も。***は平板型も。

拍	型	語例
1	平板型	柄、蚊、毛、子、血、帆…〈Ⅰ〉 名、葉、日、藻〈Ⅱ〉
	起伏型	世〈Ⅰ〉 矢〈Ⅱ〉 絵、尾、木、酢…〈Ⅲ〉
2	平板型	飴、蟻、烏賊、牛…〈Ⅰ〉
	尾高型	石、岩、歌、音、北…〈Ⅱ〉 足、網、泡、家、神*、雲*、坂*…〈Ⅲ〉
	頭高型	斧〈Ⅲ〉 跡、粟、息、糸…〈Ⅳ〉 秋、汗、雨、鮎…〈Ⅴ〉
3	平板型	欠伸、筏、錨、田舎、鯛…〈Ⅰ〉 間、桜、扉、百足〈Ⅱ〉 黄金、二十歳〈Ⅲ〉 畑、束**〈Ⅳ〉 兎、鰻、狐、雀、背中…〈Ⅵ〉 後ろ、鯨、葉、盥〈Ⅶ〉
	尾高型	麓***〈Ⅰ〉 小豆、毛抜き、とかげ、釣瓶、二つ、二人、タベ、力〈Ⅱ〉 頭、 いたち、瓦、拳、鏡、刀、林、筵…〈Ⅳ〉 くらげ、豆腐、ムジナ（狸）
	中高型	小麦、栄螺〈Ⅲ〉 五つ、思い、境〈Ⅳ〉 油、命*、心、姿*、涙*、柱、枕*、 まなこ*〈Ⅴ〉 茸、コケラ（鱗）、ひじき、マミヤ（眉）、若芽、山葵
	頭高型	えくぼ、緑〈Ⅱ〉 岬〈Ⅲ〉 嵐、紅葉、わらび〈Ⅳ〉 朝日、鯉***、きゅうり、 ざくろ〈Ⅴ〉 烏、高さ〈Ⅵ〉 蚕、兜、便り、病〈Ⅶ〉

次に、4拍以上の語について、目立った特徴を以下に挙げる。「足音」や「雷」など、尾高型から中二高型への変化が指摘されている、いわゆる「A型B型のアクセント」については、調査した語のほとんどで尾高型のアクセントが観察された。このほか、複合名詞について、「サケカス（酒粕）」や「ガラスマド（ガラス窓）」など、後部要素の下がり目が生かされるアクセントや、

「アカトンボ (赤とんぼ)」や「ウミボーズ (海坊主)」など、前部要素の上がり目が活かされる頭高型のアクセントなど、東京および周辺域における多拍の名詞の伝統的な姿が観察された。

4. 4 動詞のアクセント

動詞は、平板型と起伏型のアクセントの対立が明確にある。2 拍の語では「イク (行く)」と「カク (書く)」の 2 種、3 拍の語では「アガル (上がる)」、「アマル (余る)」、「トース (通す)」の 3 種の方の対立がある。本稿ではまず、3 拍の動詞の活用形のアクセントについて主に述べる。表 3 には「遊ぶ」、「入る」、「選ぶ」、「帰る」の活用形のアクセントが示してある。Ⅲ類の動詞「入る」は、東京のアクセントは頭高型で、活用形のアクセントも頭高型の動詞と同じになるが、当該方言においては一部、起伏型との揺れも見られるものの、Ⅰ類 (平板型) の動詞とほぼ同じアクセントになる。次に、各型式のアクセントに着目すると、特にⅠ類 (平板型) に特徴が見られる。連用形に「行く」が接続した形式のアクセントで、「アスピイク」のように動詞の末尾に上がり目があるのは県内の他の地点でも観察されている事象である。また、「ダロー」、「ナカッタ」が接続したアクセントは西関東方言におけるアクセントの古相を残している。また、後続形式に上がり目のある「ダンベ」や「ダトヨ」が接続したアクセントも注目すべき点である。

表 3 : 三浦市方言老年層話者 動詞の活用形のアクセント

後続形式 語例	終止形 命令形	～タ (た) ～テ (て)	～ナ (な・禁 止) ～バ (ば・仮 定)	～コト (こと・事) ～イル (れる・受身、尊敬) ～セル (せる・使役) ～テー (たい・願望) ～ネー (ない・打消)	～イク (しに行く)	～ノガ (方が、物が) ～ヤー (ば・仮定) ～ペー (よう・意思・勸 誘) ～マス (ます)
遊ぶ (3拍・Ⅰ類)	アスブ アスベ	アスンダ アスンデ	アスブナ アスベバ	アスブコト アスバイル アスバセル アスピテー アスパネー	アスピイク	アスブノガ (アスブノガとも) アスピヤー アスブペー アスピマス
※入る (3拍・Ⅲ類)	ヘール ヘーレ	ヘーッタ ヘーッテ	ヘーンナ ヘーレバ	ヘーンコト ヘーライル ヘーラセル ヘーリテー (ヘーリテーとも) ヘーラネー (ヘーラネー、	ヘーリイク	ヘーンノガ (ヘールノガとも) ヘーリヤ ヘーンペー ヘーリマス
選ぶ (3拍・Ⅱ類)	エラブ エラバ	エランダ エランデ	エラブナ エラババ	エラブコト エラバイル (エラバイルとも) エラバセル エラピテー エラバネー	エラビイク	エラブノガ エラビヤー エラブペー エラビマス
帰る (3拍・Ⅱ類)	ケール ケーレ	ケーッタ ケーッテ	ケーンナ ケーレバ	ケーンコト ケーライル (ケーライルとも) ケーラセル ケーリテー ケーンネー	(ケーリイク)	ケーンノガ ケーリヤ ケーンペー ケーリマス

後続形式 語例	～ソーダ (そうだ・様態) ～ナガラ (ながら)	～ダロー (だろう・推量) ～トヨー (そうだ・伝聞) ～ラシー (らしい)	～ダンベ (だろう) ～ダトヨ (そうだ・伝聞) ～ホーガ (方が) ～ヤガレ (やがれ)	～ナカッタ (なかった) ～ッサンナ (するな・禁止)
遊ぶ (3拍・I類)	アスピソーダ アスピナガラ	アスピダロー アスピトヨー アスピラシー アスピッセー アスピベヨー	アスピダンベ アスピダトヨ アスピホーガ アスピヤガレ (アスピヤガレとも)	アスピナカッタ (アスピナカッタとも) アスピッサンナ
※入る (3拍・III類)	ヘーリソーダ ヘーリナガラ	ヘーリダロー ヘーリトヨー ヘーリラシー ヘーリッセー ヘーリベヨー	ヘーリダンベ ヘーリダトヨ ヘーリホーガ ヘーリヤガレ (ヘーリヤガレとも)	ヘーリナカッタ (ヘーリナカッタとも) ヘーリッサンナ
選ぶ (3拍・II類)	エラビソーダ エラビナガラ	エラビダロー エラビトヨー エラビラシー エラビッセー エラビベヨー	エラビダンベ エラビダトヨ エラビホーガ エラビヤガレ (エラビヤガレとも)	エラビナカッタ (エラビナカッタとも) エラビッサンナ
帰る (3拍・II類)	ケーリソーダ ケーリナガラ	ケーリダロー ケーリトヨー ケーリラシー ケーリッセー ケーリベヨー	ケーリダンベ ケーリダトヨ ケーリホーガ ケーリヤガレ	ケーリナカッタ ケーリッサンナ

動詞のもう一つの問題点として、複合動詞のアクセントがある。坂本 (2017) では、「オキヤガレ (起きやがれ)」のような、前部要素の下がり目が生かされる例を示したが、複合動詞でも同様に、前部要素の下がり目が保たれるアクセントが観察されることが予想された。そこで、前部要素が起伏型の動詞である複合動詞を中心にアクセントを観察したところ、多くは「アミダシタ・アミダシタ (編み出した)」のように平板型か、末尾の2拍前に下がり目のあるアクセントとの間で揺れていたが、「ケッタボス (蹴倒す)」や「オモイシル (思い知る)」、「マチカメール (待ち構える)」など、より古いと考えられるアクセントも観察された。

4. 5 形容詞のアクセント

伝統相を多く残す名詞、動詞とくらべると、形容詞のアクセントの実態には変化の兆候が観察された。終止形単独のアクセントは、「アセー (浅い)」、「タケー (高い)」のように対立が保たれている語もあるが、同音異義語の「熱い・暑い」と「厚い」の区別が判然としないとか、「アケー・アケー (赤い)」と両方の方で発音される語があるなど、平板型から起伏型への統合が生じているような傾向が見える。中郡二宮町県内で、語尾の連母音が融合すると平板型の形容詞が起伏型で発音されやすくなる地域があるが、当該方言では、「オセー (遅い)」と融合しても平板型が保たれていたり、「カルイ (軽い)」と非融合形でも下がり目が観察されたりと、語尾の連母音の融合は起伏型への統合と関与していない様子であった。活用形のアクセントでは平板型と起伏型の対立は保たれているが、それぞれの形式で揺れを生じている部分がある。たとえば、起伏型の頭高型の形式は、下がり目が1拍後退した中高型との間で揺れている。また、「～カッタ」と「～ケレバ」の形式のアクセントの揺れは同一のアクセントの型への変化の方向と言える。

表4：三浦市方言老年層話者 形容詞の活用形のアクセント

後続形式 語例	終止形	～ヨ (よ) ～カ (か・疑問) ～ダ (だ・断定) ～テ (て) ～タ (た) ～バ (ば・仮定) ～キャ (ば・仮定)	～コト (事) ～デス (です) ～カラ (から) ～ジェン (ではないか) ～ノガ (方が、物が) ～ッテ (て) ～ネー (ない・打消) ～カネー (～はない) ～ナル (なる)	～ベー (べー・推量)	～ソーダ (そうだ・様態) ～トヨー (そうだ・伝聞)	ラシー (らしい)
浅い (3拍・I類)	アセー	アセーヨ アセーカ アセーダ アサクテ アサカッタ (アサカッタと揺れ) アサケレバ (アサケレバと揺れ) アサキャ	アセーコト アサイデス アセーカラ アセーゲン アセーノガ アサクッテ アサクネー アサカネー アサクナル	アサカンベー	アサソーダ アセートヨー	アセーラシー
高い (3拍・II類)	タケー	タケーヨ タケーカ タケーダ タカクテ※ タカカッタ※ タカケレバ※ タカキャ※	タケーコト タカイデス タケーカラ タケーゲン タケーノガ タカクッテ※ タカクネー※ タカカネー※ タカナル※	タカカンベー	タカソーダ タケートヨー	タケーラシー

なお、表中の※は「タカ〜」との揺れを表す。

5. 結論と今後の課題

三浦市方言のアクセントは東京式アクセントの体系を持つ。名詞と動詞のアクセントからは、当該地域やその周辺の地域のアクセントの伝統相が観察される一方で、形容詞は起伏型への統合や、新しいアクセントの型との間での揺れが見られるなど、変化が目立つ結果となった。

課題は、話者1名に対して調査を行っている点である。観察されたアクセントがどのような性質であるかは、先行研究と共通していればわかるが、合致しないデータについては複数の話者のデータをみて確認する必要がある。調査を継続し、データを充実させたい。

参考文献

秋永一枝 (2001) 「東京アクセント習得法則」 金田一春彦・監 『新明解日本語アクセント辞典』 三省堂

坂本薫 (2017) 「神奈川県三浦市方言にみられる音声、音韻の特徴」 第335回日本音声学会研究例会発表資料

東條操 (1954) 『日本方言学』 吉川弘文館

日野資純 (1952) 「相模方言の素描 (その方言区画)」 『国語学』 9集 武蔵野書院

日野資純 (1964) 「神奈川県の方言区画」 『日本の方言区画』 日本方言研究会編 東京堂

日野資純 (1984) 「アクセント研究に対する一つの提言 (アクセント類別語彙未登載語のアクセントと方言研究)」 『音声学会会報』

日野資純 (1992) 「神奈川県方言」 平山輝男ほか編 『現代日本語方言大辞典』 1巻 明治書院

平山輝男・編 (1960) 『全国アクセント辞典』 東京堂

心的態度を示すドイツ語心態詞と日本語終助詞の発話における 韻律的特徴について¹

生駒 美喜 (早稲田大学)

ikoma@waseda.jp

1. はじめに

ドイツ語の日常会話には、心態詞と言われる様々な心的態度を示す語が用いられる。心態詞は、状況によって様々な話者の意図やニュアンスを表す。例えばドイツ語の *schon* は文脈によって、1) 「きっと」という意図の「確信」、2) 相手の言ったことを肯定しつつ一部留保する「留保付肯定」、3) 相手が否定したことに対する「反論」という3つの意図を表すことができる(Ikoma 2007) :

- (1) a. Peter kommt schon. ペーターは(きっと)来るよ。「確信」
 b. Peter kommt schon. ペーターは来るよ。(けど・・・)「留保付肯定」
 c. Peter kommt schon. (いや,) ペーターは来るよ。「反論」

Ikoma (2007)では、心態詞 *doch*, *ja*, *schon*, *denn* を含む発話において、話者の意図やニュアンスによって特有の韻律的特徴が見られることが明らかになった。また、生駒(2014)は、「反論」の意図を示す心態詞 *schon*, *doch* を含む発話を分析した。その結果、ピッチアクセントの置かれる動詞部分のピッチピークがより遅く生じることで、「反論」として知覚されることが示唆された。さらに、Ikoma (印刷中)では、上記(1)c.の短文を用いた「反論」の発話の韻律的特徴を、同じ位置にアクセントが置かれる「確信」(1)a.の発話と比較した結果、「反論」の発話は「確信」の発話と比較して発話全体と各音節の母音の持続時間がより長く、発話全体の平均基本周波数(F0)および F0 最大値、最小値が「確信」と比べてより低くなっていた。また、アクセントの置かれる音節開始部分の F0 の値が「確信」と比較してより低く、F0 ピーク値がより高くなることで、アクセントのある音節の F0 の上昇幅がより大きくなっており、アクセント部分の F0 のピークがより遅く生じていることが明らかになった。

一方、ドイツ語心態詞と似た機能を持つとされる日本語終助詞 (Kawamori 1997) の発話にも、その意味・機能に韻律的特徴が関わっていることが次第に明らかにされている (片桐 1997, 大島 2013, 郡 2015)。さらには、「感心」や「疑い」などのパラ言語情報において、持続時間、ピッチなどの韻律的特徴が深く関わっていることが以前から指摘されている (森他 2014)。では、共通する心的態度を示すドイツ語心態詞を含む発話と日本語終助詞を含む発話には、それぞれどのような韻律的特徴が見られるのか。同じ心的態度を示す発話に共通した韻律的特徴は見られるだろうか。

本研究では、Ikoma (印刷中)における心態詞 *schon* を含む発話と状況に対応する日本語の終助詞「よ」を含む発話と状況文を用いて発話実験を実施し、音響分析を行い、日本語

¹ 本研究は、JSPS 科研費 24520441 および早稲田大学特定課題(B) 2017B-004 の助成を受けています。

終助詞を含む発話の音響特徴を明らかにし、Ikoma（印刷中）におけるドイツ語心態詞の発話の音響特徴と比較する。

2. 発話実験

2.1. 実験の資料と被験者

Ikoma（印刷中）の発話実験に用いたドイツ語心態詞 *schon* を含む短文 "Peter kommt schon" の日本語訳となる「ペーターは来るよ」の文を発話文として用い、Ikoma (ebd.) で作成した「確信」「留保付肯定」「反論」それぞれの状況文および対話文を日本語に訳した資料を用意した。実験の被験者として計 10 名（男性 3 名，女性 7 名）の日本語を母語とする大学生に協力してもらった。このうち 8 名は関東地方出身者で，残る 2 名のうち 1 名は北海道出身，もう 1 名は長野県出身である。

2.2. 発話実験の手順

実験は互いに知っている者同士 2 名 1 組の被験者で行った。2 名はそれぞれの録音ブースに入り，互いの声のみが聞こえる状況で録音を行った。被験者は実験資料をよく読んで理解した後，2 名で練習を行った。その後，被験者 2 名は各々の役を演じ，その状況になったつもりで，対話部分のみを 3 回繰り返して朗読²した。対話文の役割は途中で交代し，各々が両方の話者の役を演じた。

2.3. 音響分析

録音した音声データ（1 名につき 3 回×3 つの発話意図＝計 9 回の発話文，合計 89 発話³）をコンピュータ上に取り込み，Praat (Boersma & Weenink 2017) を用いて音響分析を行い，以下の音響特徴を測定した⁴：

- 1) 持続時間：発話全体，各音節の持続時間
- 2) 基本周波数 (F0)：発話全体の最大値，最小値，平均値，ピッチレンジ，2 つのアクセント句 (AP1: 「ペーターは」，AP2: 「来るよ」) における最大値，最小値，平均値，ピッチレンジ，AP1，AP2 の句頭の F0 の上昇幅および F0 ピークのタイミング
- 3) 振幅：発話全体，各音節における振幅の最大値，最小値，平均値，アクセント句 (AP1，AP2) における最大値，最小値，平均値，振幅の変動幅

さらに，発話末の「よ」の部分の句末音調 (Boundary Pitch Movement, BPM) を調べた⁵。

2.4. 分析結果と考察

2.4.1. 持続時間

図 1 に，「確信」「留保付肯定」「反論」の発話時間における「ペー」「ター」「は」「く」「る」

² この発話のバリエーションとして，A, B の対話文の前後に間投詞などを自由に入れて読んでもらう発話を同時に録音したが，本研究の分析対象にはしていない。

³ 1 名の被験者において 2 回しか繰り返されていない発話があったため，89 発話となった。

⁴ 音響分析に際し，小西隆之氏（早稲田大学国際教養学部助手）に Praat スクリプト作成にご協力いただいた。また，小西氏とユン・ジヒョン氏（上智大学大学院理工学研究科博士後期課程）にアノテーション作業にご協力いただいた。ここに感謝申し上げる。

⁵ 小西隆之氏に聴取とアノテーションを行っていただいた。

「よ」それぞれの音節の持続時間および発話全体の持続時間を示す。

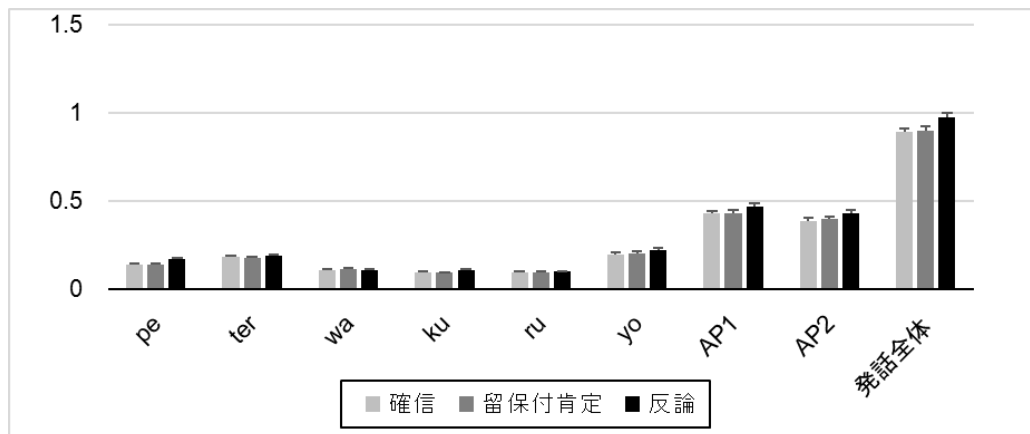


図1:各音節, AP1, AP2, 発話全体の持続時間(単位:sec)

発話の最初の音節「ペー」と最後の音節「よ」、AP1, AP2, 発話全体の持続時間において、「反論」がその他の発話意図である「確信」および「留保付肯定」よりも長くなっていた⁶。Ikoma (印刷中)におけるドイツ語心態詞を含む発話においては、文全体およびすべての母音部分の持続時間が「反論」において「確信」よりも長くなっていた⁷。発話全体の時間が「反論」の意図を含む発話において長くなっていることは、日本語とドイツ語の発話に共通する特徴として捉えることができるであろう。

2.4.2. 基本周波数(F0)

発話全体およびアクセント句 (AP1, AP2) の基本周波数の値を図2に示す。F0の最大値は、発話全体および2つのアクセント句において、いずれも「反論」が他の2つの発話意図より高いことが明らかになった⁸。また、F0変動幅(ピッチレンジ)に関しては、「留保付肯定」における発話全体のピッチレンジが他の2つの発話意図と比較して小さかった⁹。この結果は、「留保付肯定」を示すドイツ語心態詞の発話の特徴とも共通している¹⁰。AP1, AP2のピッチレンジは「反論」が他の2つの発話意図よりも大きいことが分かった¹¹。

⁶ 一元配置分散分析の結果、有意差が見られた: "pe": $F(2, 56) = 12.286, MSe=0.001, p < .01$, "yo": $F(2, 56) = 7.055, MSe=0.001, p < .01$, AP1: $F(2, 56) = 5.462, MSe=0.003, p < .01$, AP2: $F(2, 56) = 8.992, MSe=0.002, p < .01$, sentence: $F(2, 56) = 8.858, MSe=0.007, p < .01$

⁷ Ikoma においては、アクセントの位置が同じ「確信」と「反論」の発話を主として分析対象としているため、「留保付肯定」に関しては詳しい分析結果が出ていない。

⁸ 一元配置分散分析の結果以下の通り有意差が見られた。sentence: $F(2, 56) = 9.597, MSe=6.418, p < .01$, AP1: $F(2, 56) = 6.792, MSe= 8.344, p < .01$, AP2: $F(2, 56) = 13.258, MSe= 6.953, p < .01$

⁹ 一元配置分散分析の結果以下の通り有意差が見られた。 $F(2, 56) = 7.06, MSe=16.47, p < .01$

¹⁰ Ikoma (印刷中)では「反論」の発話に着目して分析を行っているため、「留保付肯定」に関するデータは示されていないが、本研究との関連で、Ikoma (ibd.)における音声データを再分析したところ、発話全体におけるピッチレンジが「留保付肯定」において他の発話意図と比べて小さいことが明らかになった。

¹¹ 一元配置分散分析の結果以下の通り有意差が見られた。 AP1: $F(2, 56) = 7.086, MSe= 8.305,$

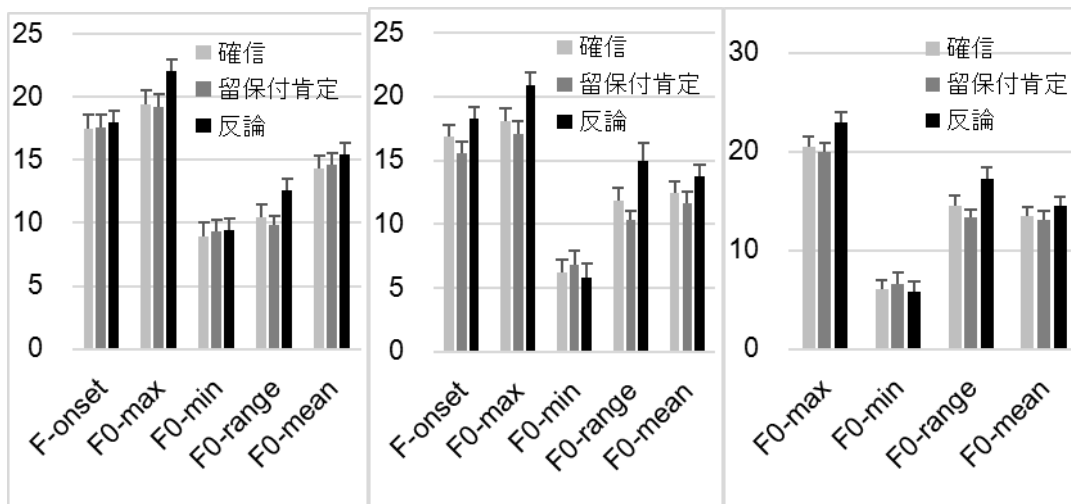


図 2: 発話全体(左)における F0 の最大値, 最小値, ピッチレンジ, 平均値, およびアクセント句 AP1 (中央), アクセント句 AP2(右)における F0 の開始値, 最大値, 最小値, ピッチレンジ, 平均値 (単位: semitone)

日本語では, アクセント句の冒頭でのピッチ上昇(「句頭の上昇」)が, 発話意図と関わっていることが知られている(森他 2014:82). そこで, 2つのアクセント句 AP1, AP2 における句頭のピッチの上昇幅および上昇のタイミングを測定した. その結果を図 3 に示す.

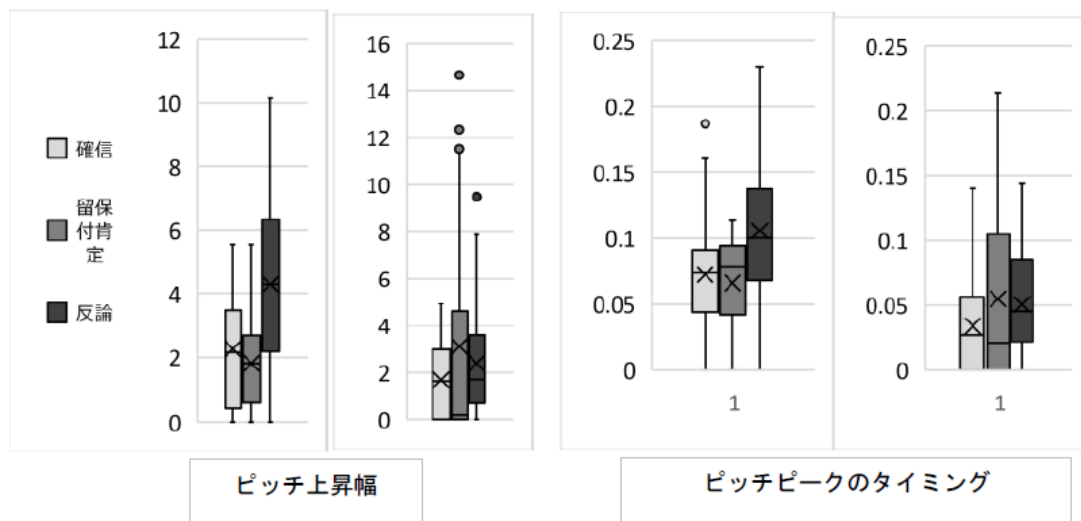


図 3 アクセント句 AP1(左 1 番目), アクセント句 AP2(2 番目)における冒頭のピッチ上昇幅(単位: semitone), アクセント句 AP1(3 番目), アクセント句 AP2(4 番目)におけるピッチピークのタイミング(単位: sec)

AP1「ペーターは」(図 3, 一番左)のアクセント句の冒頭のピッチの上昇幅が, 「反論」の発話において他の 2つの発話意図と比較して大きいことが分かる¹². 一方, AP2「来るよ」

$p < .01$, AP2: $F(2, 56) = 9.397$, $MSe = 18.172$, $p < .01$

¹² 一元配置分散分析の結果有意差が見られた: $F(2, 56) = 14.58$, $MSe = 3.29$, $p < .01$

(図3, 左から2番目)のピッチの上昇幅においては, 3つの発話意図に有意差は見られなかった¹³. ピッチのピークが生じるタイミングはAP1(図3, 左から3番目)の「反論」の発話において他の2つの発話意図と比較して遅いことが分かる¹⁴. 一方AP2(図3, 4番目)においては, 「反論」の発話におけるピッチのピークのタイミングには差異が見られなかった¹⁵が, AP2のどの音節にF0ピークが生じているかを調べた結果, 「留保付肯定」では「来るよ」の「く」の90%にF0ピークが生じているのに対し, 「確信」「反論」の発話の約60%にF0ピークが生じ, 残る発話では後続の音節「る」にF0ピークが生じていた. この結果から, 発話意図がピッチの下降タイミングにも影響している可能性が示唆される.

2.4.3. 振幅

発話全体およびAP1, AP2における振幅の最大値は, 「反論」の意図を示す発話において, 他の発話意図と比較して大きいことが明らかになった¹⁶.

2.4.4. 句末音調

終助詞「よ」の部分にどのような句末音調が現れるか, 「確信」「留保付肯定」「反論」それぞれの発話を調べた¹⁷結果, 「留保付肯定」の約60%にHの音調が現れることが分かった. 但し, Hの音調には上昇調と平板調とが存在するため, 今後の研究にて詳しく調べたい.

3. まとめ

本研究においては, 「確信」「留保付肯定」「反論」の意図を示す日本語の終助詞を含む発話「ペーターは来るよ」の音響分析を行い, 韻律的特徴を調べ, ドイツ語心態詞の発話の特徴と比較した. その結果, 以下の点が明らかになった:

- 1) 持続時間: 発話全体の持続時間が「反論」において長くなっていた. この結果はドイツ語の心態詞を含む「反論」の発話とも共通している. 一方ドイツ語心態詞の発話では「反論」の持続時間が全ての音節で伸びているのに対し, 日本語終助詞を含む発話では最初と最後の音節のみ持続時間が伸びていた.
- 2) 基本周波数: 発話全体の基本周波数(F0)は「反論」の発話の最大値が他の2つの発話意図よりも高かった. これに対し, ドイツ語心態詞の発話では「反論」においてF0最大値が低くなっていた. ドイツ語の発話では全体のピッチを下げることによって「反論」の意図を示すのに対し, 日本語の発話では, 全体のピッチを上げることで「反論」の意図を示すことが以上の結果から示唆される. 発話全体のピッチレンジは, 「留保付肯定」が他の2つの発話意図と比較して小さかった. ドイツ語の発話においても同様の結果が見られており, 今後詳しく分析をする必要がある. さらに日本語の「反論」の発話では, 最初のアクセント句の句頭上昇幅が大きく, ピークタイミングが遅く生

¹³ 一元配置分散分析の結果有意差は見られなかった: $F(2, 56) = 1.102, MSe = 16.757, n.s.$

¹⁴ 一元配置分散分析の結果有意差が見られた: $F(2, 56) = 8.813, MSe = 0.001, p < .01$

¹⁵ 一元配置分散分析の結果有意差は見られなかった: $F(2, 56) = 2.077, MSe = 0.002, n.s.$

¹⁶ 一元配置分散分析の結果有意差が見られた: AP1: $F(2, 56) = 6.408, MSe = 6.981, p < .01$, AP2: $F(2, 56) = 14.637, MSe = 7.564, p < .01$, sentence: $F(2, 56) = 10.968, MSe = 6.481, p < .01$.

¹⁷ 小西氏に句末音調の聴取とアノテーションをしていただいた.

じることが明らかになった。ドイツ語の「反論」の発話においても、アクセントのある音節冒頭の上昇幅が大きく、ピークタイミングがより遅くなるという類似の結果が見られており、今後詳しく分析する必要がある。

- 3) 振幅：日本語の「反論」を示す発話において、発話全体およびアクセント句の振幅の最大値がより大きいという結果が見られた。ドイツ語心態詞の発話に関しては今後分析を行い、共通点、相違点を明らかにしていきたい。

本研究では、日本語終助詞を含む発話の韻律的特徴を見てきたが、発話意図には韻律的特徴の他に母音の音質や発声様式（森他 2014）などの音響特徴が関わっている。今後分析を行い、ドイツ語心態詞を含む発話と日本語終助詞を含む発話に共通にみられる特徴、言語固有にみられる特徴を明らかにしていきたい。

参考文献

- Boersma, Paul / Weenink, David (2017) *Praat: doing phonetics by computer* [Computer program]. Version 6.0.29, retrieved (15 July 2017) from <http://www.praat.org/>
- Ikoma, Miki (2007) *Prosodische Eigenschaften der deutschen Modalpartikeln*. (Schriftenreihe PHONOLOGIA, Band 103). Hamburg, Dr. Kovač.
- Ikoma, Miki (印刷中) "Prosodie und Bedeutung der unbetonten und betonten Modalpartikel *schon*", *Akten des 44. Linguisten-Seminars, Tokyo 2016*. München, Iudicium.
- 生駒美喜 (2014) 「反論の意図を表すドイツ語心態詞を含む発話の韻律的特徴について」. 『第28回日本音声学全国大会予稿集』日本音声学, 123-128.
- 片桐恭弘(1997) 「12. 終助詞とイントネーション」音声文法研究会編『文法と音声』(pp.235-256), くろしお出版.
- Kawamori, Masahito (1997) "Epistemic Functions of Japanese Sentence Final Particles" In: Kajita, Masaru et al. (eds.) *Studies in English Linguistics: A Festschrift for Akira Ota*. Tokyo, Taishukan. 大修館, 1002-1015.
- 郡史郎(2015) 「終助詞『ね』のイントネーション」『音声言語の研究』10 (大阪大学大学院言語文化研究科) : 61-76. <http://doi.org/10.18910/57286> (2017.7.30.閲覧)
- 森大毅・前川喜久雄・粕谷英樹 (2014) 『音声は何を伝えているか：感情・パラ言語情報・個人性の音声科学』(日本音響学会編音響サイエンスシリーズ 12) コロナ社.
- 大島ディヴィッド義和(2013) 「日本語におけるイントネーション型と終助詞機能の相関について」『国際開発研究フォーラム』43: 47-63. <http://www.gsid.nagoya-u.ac.jp/bpub/research/public/forum/43/04.pdf> (2017.7.30.閲覧)

Assessing the Pronunciation of Japanese Learners of English

Chika Fujiyuki(1), Sayoko Eguchi(2,1), Reiko Akahane-Yamada(2,1)

(1) Graduate School of Intercultural Studies, Kobe University

(2) Advanced Telecommunications Research Institute International (ATR)

141c306c@stu.kobe-u.ac.jp, eguchi@atr-lt.jp, yamada@atr-lt.jp,

1. Introduction

There is a trend that native-like pronunciation is neither a realistic nor a suitable goal for learners of English as a second language (Isaacs, 2013). The absence of pronunciation in rating scales of frequently used speaking tests reflects this fact. Even if pronunciation is included in a scale, it is evaluated mainly in terms of intelligibility or comprehensibility (cf. Appendix 1). Having a good command on segmental and prosodic features, however, will encourage learners to speak out. The present study attempted to provide detailed pronunciation assessment using multiple metrics. Utterances produced by native speakers of Japanese on an English word reading-aloud task were evaluated by a phonetically trained Japanese-English bilingual. The correlation among the metrics was also examined to explore what property contributes to good pronunciation.

2. Method

2.1. Utterances

Utterances obtained in Eguchi and Yamada (manuscript under preparation) were used for the study. Among them, twenty English words produced by ten native speakers of Japanese on a word reading-aloud task were used. The age of the speakers, including both males and females, ranged from 19 to 40. Their English ability in terms of grade or score on English tests varied from EIKEN Grade 3 to TOEIC score 940, and none of them has been lived or studied abroad more than 12 months. The words used were following twenty words of low and high familiarity:

ahead	allegory	amass	aperitif	badminton
belfry	believe	bring	calcium	caul
Christmas	coate	conciliatory	curriculum	delicious
derelict	diode	disavow	discriminatory	doting

2.2. Assessment

Assessment was done by a phonetically trained Japanese-English bilingual. Each utterance was first transcribed with IPA and then evaluated using six metrics: 1. phoneme substitution; 2. elision; 3. epenthesis; 4. primary stress; 5. rhythm; 6. holistic goodness. One-syllable words were excluded in the evaluation with metric 4 and 5. Occurrence of phoneme substitution, elision and epenthesis was counted respectively for each speaker. Primary stress was evaluated by being classified into three

groups : no stress placement, stress placement on a wrong syllable, and stress placement on a right syllable. The ratio of the count in the third category to the entire count was obtained as accuracy in primary stress, and accuracy was obtained for each speaker and for each word. Metric 5, rhythm, was rated on a scale of 1 to 3 (1=not good, 3=good) and metric 6, holistic goodness, of 1 to 5 (1=not good, 5=good). Average scores for metric 5 and 6 were calculated by speaker and by word.

3. Result

3.1. Phoneme substitution, elision, epenthesis

Frequent occurrences were:

- Substitution of weak vowel /ə/ with /a/, /ʌ/ or /eɪ/
- Substitution of /ɪ/ with /i/
- Substitution of /aʊ/ with /oʊ/
- Substitution of /ɹ/ with /r/ (alveolar flap)
- Elision of /oʊ/ (as in -tory)

A strong correlation was found between holistic goodness and substitution ($r = 0.95$, Fig.1), but not between holistic goodness and epenthesis ($r = 0.63$, Fig.2) or elision ($r = 0.26$, Fig.3).

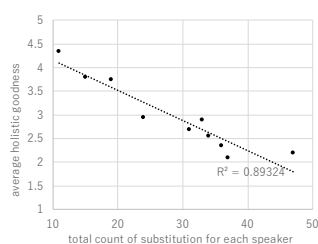


Fig.1 Correlation between substitution and holistic goodness

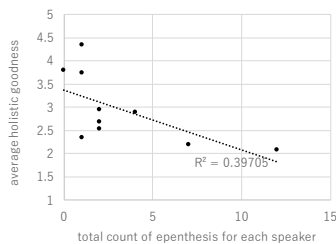


Fig.2 Correlation between epenthesis and holistic goodness

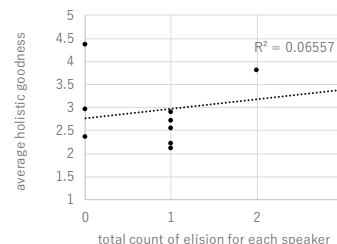


Fig.3 Correlation between elision and holistic goodness

3.2. Primary stress

Primary stress was perceived in all the bi- and polysyllabic words, except *coate*, which was pronounced as monosyllabic by most of the speakers. The accuracy by speaker varied widely, from 41% to 94 %. A strong correlation was evident between holistic goodness and the average by speaker ($r = 0.86$, Fig.4) as well as the average by word ($r = 0.79$, Fig.5).

3.3. Rhythm

The average scores by speaker varied from 1.41 to 2.76 and the average scores by word varied from 1.0 to 2.8. A very strong correlation was found between holistic goodness and the average rhythm by speaker ($r = 0.97$, Fig. 6) as well as the average rhythm by word ($r = 0.95$, Fig.7).

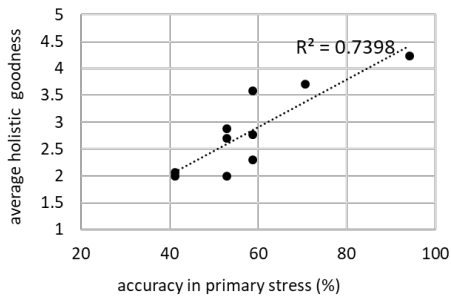


Fig.4 Correlation between holistic goodness and primary stress (by speaker)

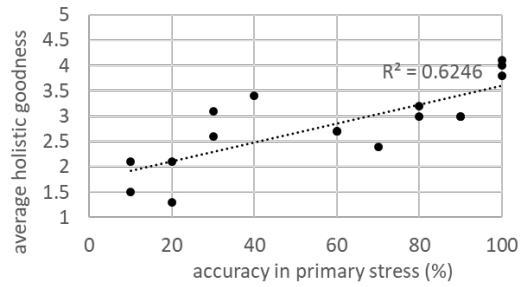


Fig.5 Correlation between holistic goodness and primary stress (by word)

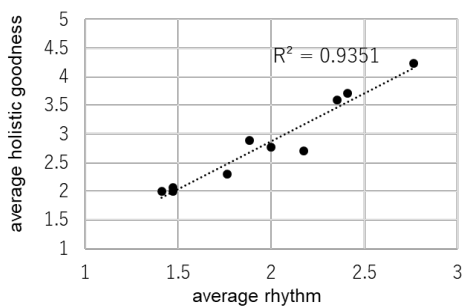


Fig.6 Correlation between holistic goodness and rhythm (by speaker)

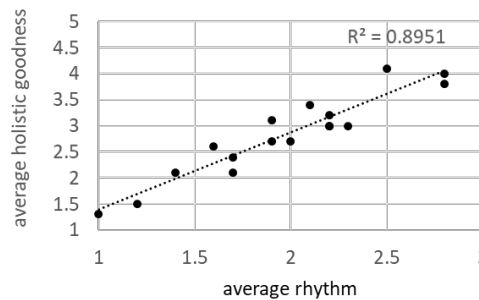


Fig.7 Correlation between holistic goodness and rhythm (by word)

3.4. The number of syllable and holistic goodness

The result of one-way ANOVA [$F(4, 45) = 2.578, p < 0.01$] suggested that the average holistic goodness varied significantly by the number of syllable. A post hoc pairwise comparisons showed that the average holistic goodness in the one-syllable group was significantly higher than the other syllable groups while the average in the six-syllable group was significantly lower than the other groups. The average of two-, three- and four-syllable groups were mutually comparable (Fig. 8).

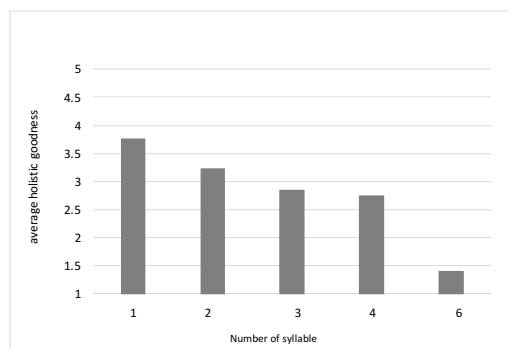


Fig.8 Correlation between the number of syllable and holistic goodness

4. Discussion

Apart from a well known substitution of /ɪ/ with /i/ by Japanese speakers, the substitution of a weak vowel /ə/ with /a/ or /ʌ/ as well as /ɪ/ with /i/ were highlighted in this study. This suggests producing weak/lax vowels is a hurdle for many Japanese learners. Since weak vowels are an essential component of English rhythm, learners must overcome this barrier to improve their pronunciation. The correlation analysis indicates a strong association between holistic goodness and phoneme substitution, primary stress and rhythm. This result indicates improving both segmental and suprasegmental features is requisite for good pronunciation. Evaluating more utterances by multiple raters is needed to enhance the reliability of the findings. Some of the metrics used in this study requires phonological expertise, yet providing detailed pronunciation assessment is manageable: it took for the rater about three to four minutes to evaluate a test word. Accumulating detailed pronunciation assessment is beneficial not only to deeper and broader understanding of the Japanese accent but also to the development of pronunciation instruction material. Such assessment data can also be applied to computer-based pronunciation assessment.

References

- Eguchi, S. and Akahane-Yamada, R. (unpublished)
- Isaacs, T. (2014). "Assessing pronunciation." In A. J. Kunnan (Ed.), *The companion to language assessment*. (pp. 140-155). Hoboken, NJ: Wiley- Blackwell

Appendix.1: Extracts from speaking skills rating in frequently used English tests

TEST	Criteria	Rating	DESCRIPTION
TOEIC	• Pronunciation (when reading aloud, pronunciation is:)	High	Highly intelligible
		Middle	Generally intelligible
		Low	Not intelligible
	• Intonation and stress levels (refer to the ability to use emphases, pauses, falling-rising pitch to convey meaning)	High	Highly effective
Middle		Generally effective	
Low		Generally not effective	
TOEFL	• Delivery difficulties with pronunciation or intonation patterns that effect intelligibility • Difficulties that requires listeners effort	4	Minor difficulties which do not affect overall intelligibility
		3	Minor difficulties which may require effort at times but do not affect intelligibility significantly
		2	Basically intelligible, though listener effort is needed because of unclear articulation, awkward intonation, or choppy rhythm and pace
		1	Consistent difficulties that cause considerable listener effort
CEFR (Common European	Pronunciation and intonation are omitted		

Framework of Reference for Languages)			
IELTS (International English Language Testing System) *public version	Pronunciation	Band.9 (highest)	<ul style="list-style-type: none"> • Use a full range of pronunciation with precision and subtlety • Sustains flexible use of features throughout • Is effortless to understand
		Band 6	<ul style="list-style-type: none"> • Use a range of pronunciation features with mixed control • Shows some effective use of features but this is not sustained • Can generally be understood throughout, though mispronunciation of individual words or sounds reduced clarity at times
		Band 4	<ul style="list-style-type: none"> • Use a limited range of pronunciation features • Attempts to control features but lapses are frequent • Mispronunciation are frequent and cause some difficulty for the listeners
		Band 3	Shows some of the features of Band 2 and some, but not all, of the positive features of Band 4
		Band 2 (lowest)	Speech is often unintelligible

英語母語話者と日本人英語学習者の冠詞の音声特徴に関する一考察

米山 聖子 (大東文化大学)・山根 典子 (広島大学)・藤森 敦之 (静岡大学)・
吉村 紀子 (静岡県立大学)

yoneyama@ic.daito.ac.jp, yamanen@hiroshima-u.ac.jp,
fujimori.atsushi@shizuoka.ac.jp, yoshimun@u-shizuoka-ken.ac.jp

1. はじめに

英語はストレス言語であるために、モーラ言語を母語とする日本人英語学習者にとっては獲得が難しいことは広く知られているところである。ストレスは英語の音韻的特徴に深く関与しており、外国語学習者にとってはとても悩ましいものである(北原 2016)。英語の強音節と弱音節が交互に起こる独特のストレスパターンは、英語の語彙とも深く関わりを持っている。英語の語彙は名詞、動詞、形容詞などの意味理解に重要な役割を果たすと考えられる内容語と冠詞、代名詞、前置詞などの文法機能に関連のある機能語とに分けられる。前者には語強勢を受ける強音節が存在し、強音節は一般的にピッチを高く、発話長を長く、そして強く発話される。これに対し、後者は通常ニュートラルなコンテキストでは弱音節として現れ、しばしば弱化の対象となる。ピッチは低く、発話長は短く、そして弱く発話される。日本人英語学習者の英語発話においては、母語である日本語の音韻特性が影響しており、語彙アクセント、音節をまたいだ時間的な構造や等時性に関わっていることが報告されている(e.g., Kubozono, 1989; Ueyama, 2000; Yamane, Yoshimura, and Fujimori, 2016a, 2016b)。

Yamane, Yoshimura and Fujimori (2016b)では、機能語である代名詞 *her* の音響的特性と文法構造について検討した。代名詞 *her* が動詞の目的語として再現される条件(e.g., *saw her*)と所有格として再現される条件(e.g., *her class*)において、*her* の発話長が句全体の発話長に占める割合を算出し、北米英語母語話者と日本語を母語とする英語学習者のデータを比較した。その結果、*her* が現れる統語的位置に関わらず、北米英語母語話者の発話における *her* の割合が日本人英語学習者のものよりも有意に低いことが明らかになっている。また、代名詞 *her* は同時に高いピッチで発話されていることも明らかになっている。

本研究では、Yamane, Yoshimura and Fujimori (2016b)の結果に基づき、機能語の一つである冠詞を日本人英語学習者がどのように英語名詞句の中で発話するのかについて検討するものである。もし、Yamane らの主張が正しいとするのであれば、冠詞(*a, the*)は機能語であるために、Yamane らの研究の名詞 *her* が動詞の目的語として再現される条件(e.g., *saw her*)と所有格として再現される条件(e.g., *her class*)と同様に、北米英語母語話者の発話における *her* の割合が日本人英語学習者のものよりも有意に低く、そして高いピッチで再現される可能性が高いと考えられる。本研究では、冠詞(*a, the*)の発話長とピッチに関する分析を行うことで、仮説の検証を行うものである。

2. 実験

2.1. 被験者

英語初級レベルの日本語を母語とする大学生男女 11 名（日本人英語学習者）と北米英語を母語とする成人男女 10 名（北米英語母語話者）である。

2.2. 録音文

録音されたトークンは、中学校 2 年生用の英語検定教科書（板垣, 2011）から抜粋した 7 文 58 語から構成されるパラグラフ(物語)である。以下にそのパラグラフを示す。

Some years ago, Mr. Sato had a very kind student in his class. She had a pretty name, Aika. Her classmates liked her very much. Sometimes Mr. Sato saw her at school early in the morning. In her hands, she always had very pretty flowers. She picked them from her garden. Everyone in her class loved the colorful flowers.

録音は静かな教室で一人ずつ行った。A4 1 ページに書かれたパラグラフを一週間前に渡し、練習するように伝えた。録音は普段通りの音読をするように伝えた。録音中に発話間違いが生じ場合には、自分のペースで言い直すように指示した。

2.3. 分析

今回は a very kind student (Condition 1), a pretty name (Condition 2), the morning (Condition 3), the colorful flowers (Condition 4) の 4 つの名詞句に焦点を当てた。測定については、発話長とピッチについて行った。発話長分析については、冠詞の発話長と句全体の発話長を測定し、冠詞の発話長の句全体の発話長に占める割合を算出した。ピッチの分析については、冠詞のピッチの平均値と後続する強音節のピッチの平均値を測定し、後者から前者を引いた値を算出した。

3. 分析結果

3.1. 冠詞の発話長に関する分析

冠詞の発話長の句全体の発話長に占める割合について、一被験者内要因一被験者間要因の混合検定を行った。被験者内要因は Condition の 4 条件（Condition 1: a very kind student, Condition 2: a pretty name, Condition 3: the morning, Condition 4: the colorful flowers）であり、被験者間要因は Group の 2 条件（北米英語母語話者, 日本人英語学習者）である。従属変数は冠詞の発話長が句全体の発話長に占める割合である。図 1 は 4 条件における冠詞の発話長が句全体の発話長に占める割合をグループごとにプロットしたものである。

Condition に主効果が認められた (Condition: $F(3, 57) = 3.665, p = 0.017$)。Condition についての多重比較 (Bonferroni 検定) から、Condition 1 (a very kind student) の条件 (0.078, SE = 0.007) よりも Condition 4 (the colorful flowers) の条件 (0.107, SE = 0.008) のほうが、有意に冠詞の発話長の句全体の発話長に対する割合が高いことが明らかになった ($p = 0.05$)。Group にも主効果が見られた ($F(1, 19) = 16.859, p < 0.001$)。冠詞の発話長の句全体の発話長に対する割合は北米英語母語話者 (0.074, SE = 0.007) のほうが日本人英語学習者 (0.116, SE = 0.007)

よりも有意に低かった。Condition と Group の交互作用は認められなかった。

Condition ごとの Group の比較では、Condition 4 (the colorful flowers) 以外の 3 条件において北米英語母語話者の句全体の冠詞の割合が日本人英語学習者のものよりも有意に低いことが明らかになった (Condition 1: $F(1,19) = 5.671, p = 0.028$; Condition 2: $F(1,19) = 14.980, p = 0.001$; Condition 3: $F(1,19) = 13.057, p = 0.002$)。また、Group ごとの Condition の多重比較 (Bonferroni 検定) では、有意差が存在したのは日本人英語学習者の Condition 1 (a very kind student) と Condition 2 (a pretty name) の比較であった。Condition 2 (a pretty name) のほうが Condition 1 (a very kind student) よりも有意に句全体の冠詞の割合が高いことが明らかになった ($p = 0.05$)。

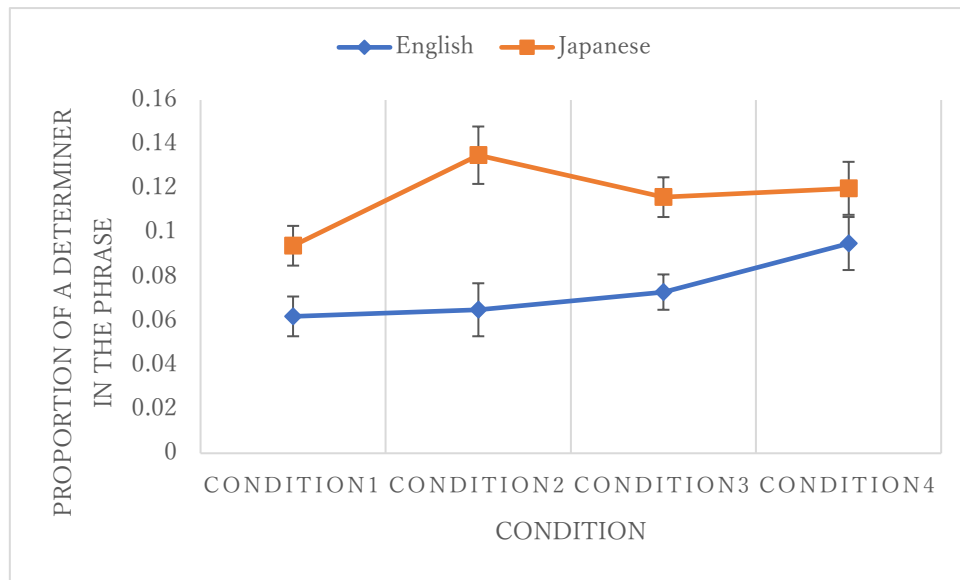


図 1: グループ別4条件における冠詞の発話長が句全体の発話長に占める割合

3.2. 冠詞と後続強音節のピッチに関する分析

冠詞のピッチの平均値と後続強音節のピッチの平均値の差 (後続強音節のピッチの平均値から冠詞のピッチの平均値を引いた値) について、一被験者内要因一被験者間要因の混合検定を行った。被験者内要因は Condition の 4 条件 (Condition 1: a very kind student, Condition 2: a pretty name, Condition 3: the morning, Condition 4: the colorful flowers) であり、被験者間要因は Group の 2 条件 (北米英語母語話者, 日本人英語学習者) である。従属変数は冠詞のピッチの平均値と後続強音節のピッチの平均値の差である。4 つの条件すべての数値が測定できた被験者は北米英語母語話者が 7 名, 日本人英語学習者が 8 名であった。これらの被験者データを今回の分析の対象とした。図 2 は 4 条件におけるにおける冠詞と後続強音節のピッチの平均値の差をグループごとにプロットしたものである。

Condition に主効果が認められなかった ($F(3,39) = 1.334, p > 0.05$)。Group には主効果が見られた ($F(1,13) = 12.508, p < 0.005$)。冠詞と後続強音節のピッチの平均値の差の数値は北

米英語母語話者のほうが (-5.819, SE = 3.184) 日本人英語学習者よりも (9.602, SE = 2.979) 有意に低かった。Condition と Group の交互作用は認められなかった。

Group ごとの Condition の多重比較 (Bonferroni 検定) では、北米英語母語話者の Condition 1 と Condition 4 の比較が有意であった (Condition 1: 7.837, SE = 4.432; Condition 4: -24.999, SE = 4.712; $p = 0.001$) Condition ごとの Group の比較では、Condition 4 (the colorful flowers)において北米英語母語話者の冠詞と後続強音節のピッチの平均値の差の数値が日本人英語学習者のものよりも有意に低かった (北米英語母語話者: -24.999, SE = 4.712; 日本人英語学習者: 9.052, SE = 4.408; $F(1,13) = 27.849, p < 0.001$) .

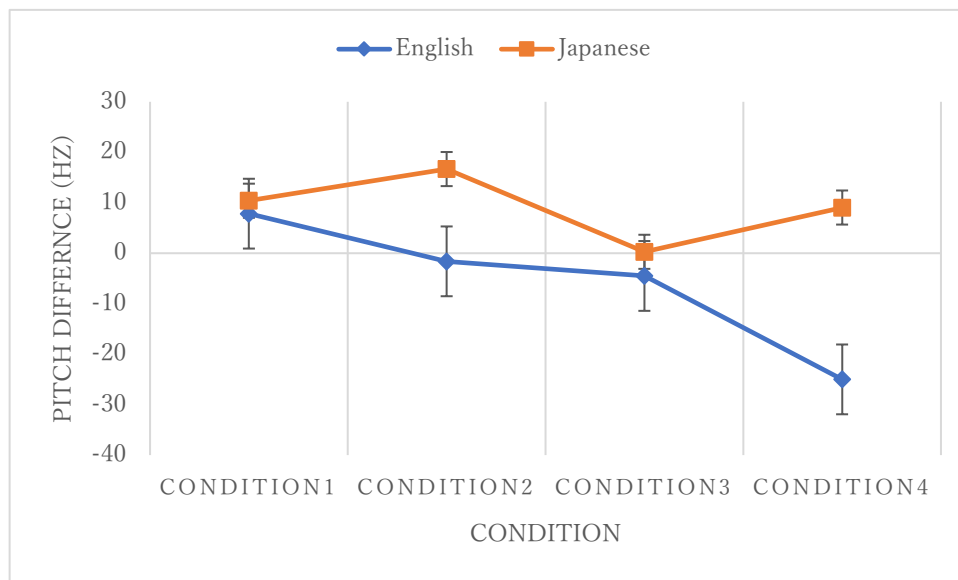


図 2: グループ別4条件における冠詞と後続強音節のピッチの平均値の差

4. 考察

まず、日本人英語学習者の冠詞の発話に関して発話長とピッチ分析の結果から検討を試みる。北米英語母語話者の発話における冠詞の時間長の句全体の時間長に対する割合が日本人英語学習者のものよりも有意に低いことが Condition 4 (the colorful flowers) 以外の Condition 1 (a very kind student), Condition 2 (a pretty name), Condition 3 (the morning)の条件について明らかになった。この結果は Yamane ら(2016b)の機能語の一分類である代名詞の場合と同様に、日本語を母語とする英語学習者は、機能語である冠詞を短く発話することができないことを示すものと考えられる。

冠詞と後続強音節のピッチの平均値の差の分析では、日本人英語学習者は全ての条件でプラスの数値を表しており、一貫して冠詞のその直後の強音節のピッチの平均値のほうが冠詞よりのピッチの平均値よりも高い。日本人英語学習者の冠詞と直後の強音節のピッチの比較については、予想通りの傾向を示すといつてもよいかもしれない。しかしながら、日本

人英語学習者は機能語である冠詞を内容語と同等の扱いをしている可能性も完全には否定できない点と北米英語母語話者のデータとは異なる点を鑑みると、日本人英語話者のピッチデータから冠詞と統語構造を明らかにする決定的な音響特徴ということは現段階では難しいと考えられる。

今回のピッチの分析結果は、日本人英語学習者は名詞句というローカルなドメインでのピッチコントロールについて検討するために行われたが、北米英語母語話者のデータはパラグラフというグローバルなドメインでの音響的なコントロールを示すものとして扱うことができ、日本人英語学習者とは異なるパターンを示している。

北米英語母語話者のデータではパラグラフの最初の文に出てくる Condition 1 (a very kind student)とパラグラフの最後の文に出てくる Condition 4 (the colorful flowers)とでは、2つの分析において有意な違いが見られた。冠詞の発話長の句全体の発話長に占める割合の分析では、Condition 1 (a very kind student)から Condition 4 (the colorful flowers)に向かって割合は増加する傾向にあり、結果的に Condition 1 (a very kind student)と Condition 4 (the colorful flowers)とで有意差が認められた。これに対し、冠詞と後続強音節のピッチの平均値の差の分析では、Condition 1 (a very kind student)から Condition 4 (the colorful flowers)に向かって差の値はプラス値からマイナス値へ変化する傾向にあり、結果的に Condition 1 (a very kind student)と Condition 4 (the colorful flowers)とでは有意差が認められた。

これら二つの現象は別々の現象と考えるのではなく、北米英語母語話者がパラグラフ全体をドメインとした情報構造伝達のプランニングに基づいた時間長とピッチのコントロールの産物であると考えられる。今回の北米英語母語話者のピッチの分析結果は、英語のイントネーション句末でみられる final lengthening と final lowering によって句の時間長の増加とピッチの下降が得られた結果である。Condition 4 (the colorful flowers)はイントネーション句末であると同時にパラグラフの最後の句でもあったため、この側面が強調され、ほかの3条件と明らかに異なる振る舞いが見られたのではないかと推測される。

本研究で明らかにできなかった点も多いが、今後の課題としてここでは3つ挙げておきたい。一つ目は冠詞の発話長に関する分析で日本人英語学習者の Condition 2 (a pretty name)のほうが Condition 1 (a very kind student)よりも有意に句全体の冠詞の割合が高いことが明らかになった点についての解釈についてである。確かに Condition 2 は She had a pretty name, Aika. というセンテンスに生起し、Aika という同格に当たる名詞を後ろに持つ点でほかの3条件と異なるが、なぜ Condition 1 (a very kind student)とだけ有意差が生じたかについては説明ができていない。今後、様々なタイプの名詞句を扱うことでこの点については検討していきたい。

2つ目は北米英語母語話者のピッチのデータについてである。ピッチのコントロールは英語においてもグローバルなドメインのイントネーションについてと、ローカルなドメインの音節強勢の両方に関わっているが、今回の北米英語母語話者のデータは前者を示すものであり、後者を示すものではなかった。なぜこのような結果になったのか、またはなぜ前

者と後者の両方を示す結果でなかったのかについて検討していきたい。

最後に、今後は被験者数を増やし、強勢の有無に関係があるとされるラウドネスについての分析も加えて、ピッチ、発話長、ラウドネスの3つの観点から総合的に冠詞の音響的特性と統語構造との関係について分析を試みたい。

参考文献

板垣信哉 (2011) *Sunshine English Course 2*. 東京：開隆堂.

北原真冬 (2016) 「英語のストレスに立ち向かう日本語話者」日本音韻論学会 (編)『現代音韻論の動向— 日本音韻論学会 20 周年記念論文集 —』(pp. 124-127) 東京：開拓社.

Kubozono, Haruo (1989) "Syntactic and rhythmic effects on downstep in Japanese", *Phonology*, 6(1), 39-67.

Ueyama, Motoko (2000) *Prosodic transfer: An acoustic study of L2 English vs. L2 Japanese*. Doctoral dissertation, University of California Los Angeles.

Yamane, Noriko, Yoshimura, Noriko, and Fujimori, Atsushi. (2016a). "Prosodic transfer from Japanese to English: Pitch in focus marking", *Phonological Studies* 19, 97-104. Tokyo: Kaitakusha.

Yamane, Noriko, Yoshimura, Noriko and Fujimori, Atsushi (2016b) "Japanese EFL Learners' Production of Pronouns in English: Evidence for L2 Prosodic Structure", a paper presented at Acoustics Week in Canada 2016, September 21-23, 2016, The Sutton Place Hotel, Vancouver, British Columbia.