

日本語学習者の心的態度と感情の強調順位に関する知覚習得

福岡昌子 (三重大学)
fukuoka@cie.mie-u.ac.jp

1. はじめに

本研究では、「聞き手」である日本語母語話者や日本語学習者は、話者が制御できる心的態度 (mental attitudes) と話者が制御できない感情 (feelings) のパラ言語的情報としてのメッセージ (話者が伝えようとしていること) について、どのように強調の順位付けをして知覚しているか、言語の普遍性や個別の言語依存性がどのように影響するか、明らかにする。

2. 先行研究

2.1 パラ言語的情報の研究

Murray, I. and Arnott, J. (1993) は、感情発話に関する聴覚印象や音響機器を使った先行研究から、怒り・幸福感・悲しみ・恐怖・不快などの感情発話の特徴に対し、Speech Rate・ピッチ平均・強さ・声質・ピッチ変化・調音における関連性について報告した。Ladd, et al (1985) は、発話意図と感情の知覚実験の結果、パラ言語的情報に対し基本周波数 (F0) レンジの違いが最も重要であることを明らかにした。日本語のパラ言語的情報は、藤崎 (1994) が音声により伝達される情報を、言語的情報 (linguistic information), パラ言語的情報 (para-linguistic information), 非言語的情報 (non-linguistic information) に分類することを提案した。パラ言語的情報の研究の解明には、まだ多くの課題が残されている。

2.2 日本語を第2言語とするパラ言語的情報の習得研究

福岡 (1998) は初級の中国人日本語学習者を対象に、勧誘と否定の言語的情報が同形の自然音声及び、それらの文末の基本周波数とピッチ曲線を操作した14個の合成音声を用いて調査を行った。その結果、中国語母語話者は文末のピッチの変動による発話音声とその発話意図とを一致させることが困難であることを報告し、中国語には文末1拍分の急上昇する文末上昇疑問文がないことが要因となる可能性を示唆した。また、前川・北川 (2002) は、米語母語話者を対象に日本語学習経験の有無別に多次元尺度法 (MDS) を用いてパラ言語的情報の知覚空間の構築を試みた。その結果、感心・落胆・無関心 (文末下降), 強調・中立・疑い (文末上昇) の中で、非学習者には感心・無関心 (下降) と疑い (上昇) の聞き分けが困難であることを報告した。これらの先行研究より、パラ言語的情報には言語の依存性があり、感情の知覚には高度の言語の普遍性があることが指摘された。本研究においても、第2言語学習者のパラ言語的情報の習得には、個別言語による言語の依存性と言語の普遍性が大きく関わると考える。

2.3 心的態度と感情のパラ言語的情報

上記の藤崎 (1994) の分類では、話者が制御できる発話意図や心的態度はパラ言語的情報、話者が意識的に制御できない感情や年齢・性別・健康状態等の個人的特徴は非言語情報に分類される。話者が制御できる心的態度として、鹿島 (2002, p. 138) 「質問・確認・納得など表現する意図は同じでも『丁寧、ぶっきらぼう、自慢げ、不安げ、自信満々、自

信なく、無関心、横柄、相手を気遣う』など、様々な発話のニュアンスがある」と指摘している。研究対象としては丁寧に関する研究が多い。一方、話者が制御できないものとされる感情では、森・前川・粕谷（2014）は、「メッセージとして意図的に生成される感情」を「不随意的に表出される感情」と区別して、藤崎（1994）の分類におけるパラ言語的情報にまとめることを提案している。本研究もこの提案と同じ立場に立つ。

本研究では、話者が制御できる心的態度と制御できない感情におけるパラ言語的情報としての話者のメッセージを、「聞き手」である日本語母語話者や日本語学習者がどう知覚するかに着目する。「聞き手」はパラ言語的情報における話者のメッセージについて、順位付けして聞く知覚能力を内在的に持っており、その強調順位（recognition order of emphasis）に従って音声から意味を識別していると考えられる。

3. 研究方法

3.1 分析方法

- 1) 本研究では、強調の知覚において、強調部分を高く長く発話した音声 > 強調部分を長く発話した音声 > 強調していない音声 > 発話速度の速い音声の順に順位付けが行われるという仮説を立て、その仮説に従う順位付けで知覚した被験者の割合（「強調順一致率」）を中心に分析する。
- 2) 声の高さ（F0）、長さ（Duration）、発話速度（Speech Rate）などが異なる自然音声を中心に（発話速度を調整した合成音声を含む）分析する。
- 3) 日本語母語話者と日本語学習者のパラ言語的情報における強調順位の違いを調べる。
- 4) 日本語学習者にとって一般的に習得が苦手とされている音声（例：特殊音など）を含むパラ言語的情報の場合など、強調順位の知覚への影響を分析する。

以下、声の高さ、長さ、発話速度を、F0, Duration, Speech Rate とする。

3.2 調査語（表 1）

表1. 心的態度と感情の調査語（日本語・中国語・英語・韓国語）

（「 」はアクセント核「 」は平板型、太字は 2.3.1 の a.で、下線は b.で強調された箇所）

	日本語（自然音声）グループ	中国語（自然音声）	英語（自然音声）	韓国語（自然音声）	
心的態度	評価や、好・悪などの心的態度	1. <u>こんな</u> に「(形動)勉強したよ。」	我学了 <u>这么多</u> 呢。(代詞)。	I studied so <u>much</u> . (形)	이렇게나 (形) 공부했어..
		2. <u>いい</u> ね(形)、これ。	挺好呀、这个。(形)。	This is <u>great</u> . (形)	좋다 (形) 이거..
		3. <u>いっぱい</u> 「(副詞)勉強したよ。」	我学了 <u>好多</u> 呢。(副詞)。	I studied a <u>lot</u> . (名)	많이 (形) 공부했어..
		4. <u>かわいい</u> い(形)。このネコ。	这个猫、 <u>真</u> 可爱。(副詞)。	This cat is <u>cute</u> . (形)。	귀엽네 (形) 이 고양이.
感情 (positive)	ポジティブ	5. <u>うれ</u> しい(形)	<u>高兴</u> (形)。	I'm <u>happy</u> (形)	기쁘 (形)。
	感 (positive)	6. <u>しあ</u> わせ(名)	<u>幸福</u> (形)	I'm <u>pleased</u> (形)	행복해 (形)
感情 (negative)	ネガティブ	7. <u>かな</u> しい(形)	<u>难过</u> (形)	I'm <u>sad</u> (形)	슬프 (形)
	情 (negative)	8. <u>こわ</u> い(形)	<u>恐怖</u> (形)。	I'm <u>scared</u> (形)。	무서워 (形)。

3.3 日本語と母語の自然音声グループの刺激音（表 2, 図 1）

表 1 の心的態度と感情について、音声提供者には録音前の指示では「4 段階の状況と場面設定」の説明（例：「安くてかなり便利な物だった」 > 「かなり便利な物だった」 > 「便利な物だった」 > 「良さそうな物だった」）と、強さの段階の説明をした（表 2）。そして、強調すると思うところを「高く長く発話してください (a.)」、「長く発話してください (b.)」、「普通に発話してください (c.)」と指示し録音した。「強調すると思うところ」は自分の

解釈で自由に強調して発話してもらった。その結果、録音された音声では表1の太字や下線で示された部分が、高く/長く発音された。母語の自然音声においても、日本語の意味や発話のニュアンスを考慮に入れ、中国語、英語、韓国語を選定し、日本語の自然音声グループと同様に、a. b. c. d.の音声を用いて調査した。

表2. 日本語の自然音声グループの刺激音(a~c:自然音声, d:合成音声)

刺激音の略称	刺激音の音声について	強調段階
a. F0&Duration 音声 a.	(強調すると思うところを), 高く長く発話した自然音声	強い
b. Duration 音声 b.	(強調すると思うところを), 長く発話した自然音声	↑
c. 無加工音声 c.	強調を意識しないで発話した自然音声	↓
d. Speech Rate 音声 d.	c.の発話全体の Speech Rate を68%に速めた合成音声	弱い

3.4 音声提供者

日本語 (40代, 50代男性東京出身), 中国語 (20代女性瀋陽出身), 韓国語 (20代女性ソウル出身2名), 英語 (50代男性ロンドン出身). 録音は Roland R-05 または SONY IC RECORDER ICD = SX77 を使用した。

3.5 調査対象者

(2016年4月~2016年6月実施)

- 1) 日本語 (母語) 話者 (出身地の異なる大学1年生) 50名
- 2) 中国語話者 (中国 [鎮江]の大学日本語学科1年生15名, 3年生15名)
- 3) 韓国語話者 (韓国 [水原]の大学日本語学科1年生15名, 3年生15名)
- 4) 英語話者 (豪州 [ホバート]の大学日本語専攻コース1年生15名, 3年生15名)

3.6 知覚調査の実施方法

まず, 被験者には静かな教室で実験方法について説明と知覚練習を行った。次に, 3.3で作成した各4種類の音声をランダムに並べ, 1つの調査文につき1セットとして3回聞かせ, その都度順位付けさせた。これを全ての調査文に対して計8回繰り返した。設問間は10秒, ランダムに並べた各設問のa~dの音声間は5秒とした。その後, 各音声の「心的態度」と「感情」について「強」から「弱」に1位から4位の順位付けをさせた。表3の回答用紙を使用し, 外国人被験者の回答用紙には全て翻訳をつけた。調査時間は, ①母語の自然音声調査 (7分): 母語が日本語以外の言語話者, ②日本語の自然音声調査 (14分) とした。

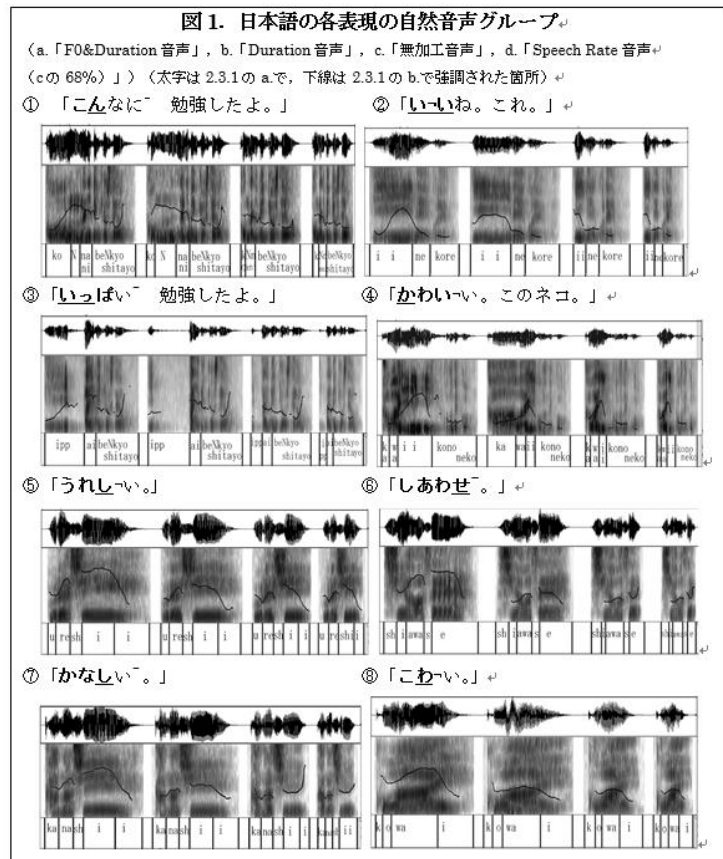


表3. 日本語の自然音声 例2. 「いいね。これ。」の知覚調査の回答用紙

4つのことばを3回聞いて, 発話者の心的態度 (mental attitude) の強い順に, ABCDを並べてください。

①A「いいね。これ。」 B「いいね。これ。」 C「いいね。これ。」 D「いいね。これ。」
() > () > () > ()

②A「いいね。これ。」 B「いいね。これ。」 C「いいね。これ。」 D「いいね。これ。」
() > () > () > ()

③A「いいね。これ。」 B「いいね。これ。」 C「いいね。これ。」 D「いいね。これ。」
() > () > () > ()

● 心的態度の強い順番は? () > () > () > () (強い → 弱い)

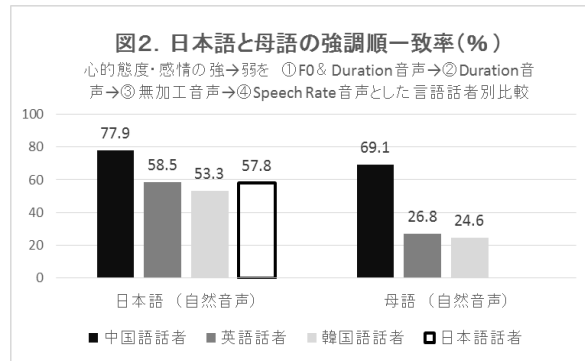
3.7 統計解析手法

統計解析ソフトは、SPSS Statistics 22 (IBM Japan, Ltd.) を使用した。調査に参加した回答者全てを解析対象とし、調査内容及び回答者属性別に群分けし、回答の頻度集計及び比較検討を行った。各設問における回答順序は、「F0&Duration 音声 (a) > Duration 音声 (b) > 無加工音声 (c) > Speech Rate 音声 (d) 」の順で一致させた被験者の割合を「強調順一致率」と定義し算出した。また、強調順一致の強さを weighted κ -value を算出し評価を行った。群間の比較には、名義尺度には Fisher's exact test, 連続尺度には Mann Whitney U test を使用し、多重比較における補正には Bonferroni 法を用いた。有意水準は 5% とした。

4. 調査の結果

4.1. 音声種別 (日本語と母語) の比較

図 2 は、日本語の自然音声、各母語の自然音声の 3 種の音声種別ごとに、心的態度と感情の設問 (1~8) 間における「強調順一致率」の平均を示したものである。回答者すべての平均 (調整平均) では自然音声が 66.1%, 各母語の自然音声が 51.5% (日本語話者を含む) だった。



4.2. 各言語話者の日本語と母語 (自然音声) の知覚結果に関する比較 (表 5)

日本語の設問 (1~8) 間における「強調順一致率」の平均が高い順は、中国語話者 77.9% > 英語話者 58.5% > 日本語母語話者 57.8% > 韓国語話者 53.3% だった。母語の自然音声の結果では、中国語話者 69.1% > 日本語母語話者 57.8% > 英語話者 26.8% > 韓国語話者 24.6% の順だった。中国語話者は、日本語の設問 (1~8) 間における「強調順一致率」が高く ($p < 0.001$), 同様に、どの言語話者も日本語の自然音声の方が母語の自然音声より「強調順一致率」が高かった ($p < 0.001$)。

表 5. 各言語話者の日本語と母語 (自然音声) の「強調順一致率」

(*全 8 問の平均「強調順一致率」, **全言語話者の平均「強調順一致率」)

	日本語の自然音声グループ				母語の自然音声グループ				
	心的態度		感情		心的態度		感情		平均*
	1.2.3.4.	5.6.	7.8.		1.2.3.4.	5.6.	7.8.		
中国語話者	70.9%	87.1%	82.9%	77.9%	64.4%	73.5%	74.1%	69.1%	
英語話者	53.6%	64.3%	62.5%	58.5%	22.3%	26.8%	35.7%	26.8%	
韓国語話者	46.7%	65.0%	55.0%	53.3%	14.2%	31.7%	38.3%	24.6%	
日本語 (母語) 話者	62.5%	65.0%	41.0%	57.8%	(62.5%)	(65.0%)	(41.0%)	(57.8%)	
調整平均**	62.4%	74.6%	64.8%	66.1%	45.6%	55.6%	59.1%	51.5%	

4.3. 日本語と母語 (自然音声) のパラメータ間比較

表 6 と表 7 は、日本語および母語の自然音声において、①F0&Duration に関する音声 (a), ②Duration に関する音声 (b), ③無加工音声 (c), ④Speech Rate に関する音声 (d) のパラメータごとに結果を示した。

表6. 日本語の自然音声の順位付け結果 表7.母語(学習者)の自然音声の順位付け結果

(■最も多い回答, □25~50%の回答(最も多い回答を除く), ▨順位が異なる回答)

自然音声	中国語話者				英語話者				韓国語話者				日本語話者			
	1位	2位	3位	4位	1位	2位	3位	4位	1位	2位	3位	4位	1位	2位	3位	4位
1. ことば(内容)																
F0&Duration音声(a)	85.5	9.6	3.6	1.2	61.5	23.1	7.7	7.7	75.0	25.0	0.0	0.0	73.5	22.4	0.0	4.1
Duration音声(b)	8.4	86.7	2.4	2.4	23.1	57.7	11.5	7.7	20.8	75.0	4.2	0.0	18.4	73.5	4.1	4.1
無加工音声(c)	1.2	2.4	83.1	13.3	0.0	19.2	50.0	30.8	0.0	0.0	66.7	33.3	2.0	4.1	77.6	16.3
Speech Rate音声(d)	4.8	1.2	10.8	83.1	15.4	0.0	30.8	53.8	4.2	0.0	29.2	66.7	6.1	0.0	18.4	75.5
2. ういゑ(内容)																
F0&Duration音声(a)	86.6	11.0	1.2	1.2	66.7	16.7	8.3	8.3	87.0	8.7	4.3	0.0	81.6	14.3	0.0	4.1
Duration音声(b)	11.0	86.6	0.0	2.4	16.7	66.7	8.3	8.3	13.0	82.6	4.3	0.0	18.4	79.6	2.0	0.0
無加工音声(c)	1.2	1.2	92.7	4.9	0.0	16.7	83.3	0.0	0.0	4.3	78.3	17.4	0.0	0.0	93.9	6.1
Speech Rate音声(d)	1.2	1.2	6.1	81.5	16.7	0.0	0.0	83.3	0.0	4.3	13.0	82.6	0.0	6.1	4.1	89.8
3. うっけい(内容)																
F0&Duration音声(a)	86.7	7.2	2.4	3.6	73.1	11.5	3.8	11.5	50.0	50.0	0.0	0.0	75.5	22.4	2.0	0.0
Duration音声(b)	12.0	85.5	2.4	0.0	11.5	73.1	11.5	3.8	50.0	50.0	0.0	0.0	20.4	77.6	0.0	2.0
無加工音声(c)	0.0	4.8	71.1	24.1	0.0	15.4	76.9	7.7	0.0	0.0	83.3	16.7	0.0	0.0	89.8	10.2
Speech Rate音声(d)	1.2	2.4	24.1	72.3	15.4	0.0	7.7	76.9	0.0	0.0	16.7	83.3	4.1	0.0	8.2	87.8
4. まわりの(内容)																
F0&Duration音声(a)	75.9	16.9	4.8	2.4	72.0	12.0	8.0	8.0	87.0	8.7	0.0	4.3	63.3	34.7	2.0	0.0
Duration音声(b)	18.1	78.3	3.6	0.0	12.0	72.0	8.0	8.0	8.7	87.0	0.0	4.3	34.7	61.2	4.1	0.0
無加工音声(c)	4.8	4.8	81.9	8.4	0.0	16.0	84.0	0.0	0.0	0.0	95.7	4.3	2.0	4.1	81.6	12.2
Speech Rate音声(d)	1.2	0.0	9.6	89.2	16.0	0.0	0.0	84.0	4.3	4.3	4.3	87.0	0.0	0.0	12.2	87.8
5. うれしい																
F0&Duration音声(a)	95.2	2.4	0.0	2.4	76.9	3.8	0.0	19.2	100.0	0.0	0.0	0.0	93.9	2.0	2.0	2.0
Duration音声(b)	1.2	92.8	4.8	1.2	3.8	80.8	15.4	0.0	0.0	72.0	20.0	8.0	2.0	73.5	24.5	0.0
無加工音声(c)	0.0	4.8	95.2	0.0	3.8	15.4	76.9	3.8	0.0	16.0	80.0	4.0	2.0	24.5	69.4	4.1
Speech Rate音声(d)	3.6	0.0	0.0	96.4	15.4	0.0	7.7	76.9	0.0	12.0	0.0	88.0	2.0	0.0	4.1	93.9
6. しあわせ																
F0&Duration音声(a)	96.4	1.2	0.0	2.4	81.5	3.7	0.0	14.8	96.0	4.0	0.0	0.0	83.7	10.2	4.1	2.0
Duration音声(b)	1.2	95.2	3.6	0.0	0.0	74.1	22.2	3.7	4.0	92.0	4.0	0.0	6.1	81.6	10.2	2.0
無加工音声(c)	1.2	3.6	86.8	8.3	3.7	18.5	66.7	11.1	0.0	4.0	88.0	8.0	8.2	6.1	71.4	14.3
Speech Rate音声(d)	1.2	0.0	9.5	89.3	14.8	3.7	11.1	70.4	0.0	0.0	8.0	92.0	2.0	2.0	14.3	81.6
7. かなしい																
F0&Duration音声(a)	92.9	7.1	0.0	0.0	80.8	7.7	3.8	7.7	84.6	11.5	0.0	3.8	44.9	6.1	18.4	30.6
Duration音声(b)	4.8	91.7	2.4	1.2	3.8	76.9	15.4	3.8	7.7	73.1	11.5	7.7	0.0	63.3	32.7	4.1
無加工音声(c)	1.2	1.2	92.9	4.8	0.0	15.4	73.1	11.5	7.7	11.5	65.4	15.4	42.9	14.3	42.9	0.0
Speech Rate音声(d)	1.2	0.0	4.8	94.0	15.4	0.0	7.7	76.9	0.0	3.8	23.1	73.1	12.2	16.3	6.1	65.3
8. こわい																
F0&Duration音声(a)	98.8	0.0	1.2	0.0	74.1	11.1	3.7	11.1	92.3	3.8	3.8	0.0	75.5	8.2	2.0	14.3
Duration音声(b)	1.2	97.6	0.0	1.2	7.4	74.1	14.8	3.7	3.8	92.3	3.8	0.0	8.2	71.4	20.4	0.0
無加工音声(c)	0.0	1.2	78.6	20.2	0.0	14.8	70.4	14.8	0.0	3.8	73.1	23.1	14.3	10.2	44.9	30.6
Speech Rate音声(d)	0.0	1.2	20.2	78.6	18.5	0.0	11.1	70.4	3.8	0.0	19.2	76.9	2.0	10.2	32.7	55.1

5. 考察および結論

●本研究の結果：日本語母語話者や日本語学習者の心的態度及び感情に関する「強調順一致率」は、日本語の自然音声 66%，母語の自然音声 52%で、①F0&Duration 音声 (a) > ②Duration 音声 (b) > ③無加工音声 (c) > ④Speech Rate 音声 (d) の順だった。即ち、被験者は母語に関わらず強調部分を、高く長く発話した音声 > 長く発話した音声 > 強調していない音声 > 発話速度の速い音声の順に知覚した。強調された発話は「高く、長く」発音される可能性が高い。日本語母語話者や日本語学習者における共通した強調順位に関する知覚認識の結果は、パラ言語的情報における言語の普遍性の可能性を示唆するものである。

●習得：日本語の「強調順一致率」の調査では、英語や韓国語を母語とする学習者は撥音が、韓国語を母語とする学習者は促音と、特殊音が含まれるパラ言語的情報において強調順位が異なる傾向が観察された。日本語の特殊音は、一般に多くの日本語学習者にとって習得が難しく、特殊音を含むパラ言語的情報としての知覚においても影響があると考えられる。本研究の結果は、日本語学習者にとって苦手な音声であった場合に、それらの音を含むパラ言語的情報の知覚においても、習得が難しくなる可能性があることを示唆する。

●日本語母語話者の興味深い結果：知覚結果から、ネガティブ感情「かなしい」「こわい」

に対する結果は仮説の強調順位と逆順位化傾向を示した。日本語母語話者は、特にネガティブ感情の「かなしい」において、F0 が低く Duration が短い無加工音声 (c) や Speech Rate 音声 (d) を、話者が伝えたい悲しみを強調した音声であると認識していることがわかった ($\kappa = 0.229$)。パラ言語的情報の強調を示す音声は、他表現で観察されたように必ずしも F0 や Duration が大きく関係するのではなく、Duration の短さや Speech Rate も悲しみの強調を高く示す音声であることが明らかになった。これは日本語のパラ言語的情報の個別性と見ることができると思われる。

●言語話者別：韓国語話者は、心的態度やポジティブ感情において Speech Rate 音声 (d) を、話者が強調した音声として判断する傾向がうかがえた。英語話者は、パラ言語的情報の音声として F0 や Duration や Speech Rate 以外の音声、例えば音の大きさやストレスの位置などが話者の強調を表している可能性がある。一方、中国語話者は、日本語のパラ言語的情報において F0 の高低差に敏感であり、日本語と母語の両言語において完全に「強調順一致率」を支持する結果だった。特に中国語話者には、ネガティブ感情「かなしい」においても高い「強調順一致率」が見られ、F0 が高い音声を最も悲しみの強い音声と認識し、他の言語話者と異なる結果が認められた。これらの点は、その言語特有のパラ言語的情報が大きく関係しており、その言語特有の文化歴史的背景等の要因による言語の依存性が存在するのではないかと考えられる。

●心的態度と感情：本研究における「うれしい」や「かなしい」の感情表現は、音声提供者によって強調度を区別して「意志的に生成された感情表現」であり、一方「聞き手」である日本語母語話者や日本語学習者も、話者の「意志的に生成された感情表現」をパラ言語的情報として認識できていた。森・前川・粕谷 (2014) の提案にあるように、「メッセージとして意図的に生成される感情」は、「不随意的に表出される感情」と区別し、非言語的情報ではなくパラ言語的情報に分類した方が、今後の感情やパラ言語的情報に関する研究がより発展するのではないかと考えられる。

参考文献

1. 鹿島央 (2002) 『日本語教育をめざす人のための基礎から学ぶ音声学』スリーエーネットワーク
2. 福岡昌子 (1998) 「イントネーションから表現意図を識別する能力の習得研究—中国方言話者を対象に自然音声・合成音声を使って—」『日本語教育』96, 37-48.
3. 福岡昌子 (2017) 「パラ言語的情報の協調順位に関する日本語学習者の知覚」『音声研究』21-3, 1-14.
4. 藤崎博也 (1994) 「韻律研究の諸側面とその課題」『日本音響学会平成6年度秋季研究発表会講演論文集I』日本音響学会, 287-290.
5. 前川喜久雄・北川智利 (2002) 「特集—言語コミュニケーションの科学へ向けて— 音声はパラ言語的情報をいかに伝えるか」*Cognitive Studies*, 9(1), 46-66.
6. 森大毅・前川喜久雄・粕谷秀樹 (2014) 『日本音響学会編音響サイエンスシリーズ12 音声は何を伝えているか—感情・パラ言語的情報・個性の音声科学』コロナ社
7. Ladd, R. K, Silverman, F. Tolkmitt, G. Bergmann and R. Scherer (1985) "Evidence for the independent function of intonation contour type, voice quality, and F0 range in signaling speaker affect." *Journal of the Acoustical Society of America* 78(2), 435-444.
8. Murray, I. and Arnott, J. (1993) "Toward the simulation of emotion in synthetic speech: A review of the literature on human vocal emotion" *Journal of the Acoustical Society of America*, 93(2), 1097-1107.

中国人上級日本語学習者の特殊拍の知覚的混同 ——意味文脈の影響——

張 林姝（神戸大学大学院） 山田 玲子（ATR/ 神戸大学）

1. はじめに

日本語の音声学習における特殊拍の習得は多くの先行研究で論じられてきた（戸田，2003）。しかし、それらの研究の多くは特殊拍の有無による区別を扱ったものであり、特殊拍間の関係性に言及した研究は少ない。特殊拍間の主たる区別は質的（スペクトルパターンとその変化）差異であるが、特殊拍の有無の区別はその質的差異に時間長、つまり量的差異も加わる。したがって、量的差異がない特殊拍間の区別にも問題がある可能性は十分にある。そもそも音韻論では特殊拍は自立拍とは一線を画すのに対し、音声学では各特殊拍は、調音方法も調音点も異なる単音とみなしている。音声学の観点からも特殊拍間の混同に着目することは妥当である。

本橋（2005）は英語母語話者を対象に、/Q/あり・/Q/なしの混同があるかどうかを検証するための知覚実験を行った。その結果、/Q/の有無による混同以外にも、/Q/と/R/の混同が観察された。Zhang, Hayashi, and Akahane-Yamada（2017）は中国人日本語上級者と日本語母語話者を対象に、特殊拍が含まれる無意味語を刺激語とし、キャリア文の中で刺激語を同定する知覚実験を行った。中国語母語話者の正答率は、特殊拍の種類や生起位置にかかわらず、日本語母語話者より低かった。また、中国語母語話者の誤答パターンでは、特殊拍同士の混同（/R/↔/Q/；/N/→/R/）が観察された。

一方、外国語の音声知覚には音響的要因のみならず、意味文脈も大きな影響を及ぼすことが報告されている。意味文脈は、単独で出現する単語と意味性が高い文脈と意味性が低い文脈に分けられる。母語話者を対象に行われた音韻知覚実験から、意味文脈が大きな影響を与えることが多数報告されている（Craig, 1988; Burton, Baum, & Blumstein, 1989; Griffin & Bock, 1998）。Rothwell and Akahane-Yamada（2003）は日本語を母語とする英語学習者を対象に、アメリカ英語の/r/-/l/の知覚同定実験を、3種類の文脈環境（①意味性の高い文脈：意味文；②意味性の低い文脈：中立文；③単語単独）で検証した結果、弁別の正答率は意味文>単語>中立文の順に並べることが分かった。Ikuma and Akahane-Yamada（2003）は同じ手法で/r/-/l/・/s/-/θ/・/b/-/v/の知覚弁別を検証した結果、/r/-/l/では同じ結果が見られた。

本研究では、上記の2点、つまり特殊拍の混同と意味文脈の影響について検討した。その際、4種類の特殊拍（/N/・/R/・/Q/・なし）の混同を調べるための単語セットを3種類の文脈環境（WD：単語単独呈示；CS：意味情報が利用できる文脈文内の呈示；NS：意味情報が利用できない中立文内の呈示）に配置し、中国人上級日本語学習者を実験群、日本語母語話者を統制群に、聴取実験を行った。

2. 方法

2.1. 刺激

2.1.1 音声コーパスの作成

拍の種類で対立する4語セットをNTTデータベースシリーズ『日本語の語彙特性』（収録語数 69,084 語）から抽出した。本研究では、見出し語の拍数、アクセント型とともに、単語親密度の情報が付与されている第1巻の単語親密度データベースと第3巻の単語アクセントデータベースを使用した。

特殊拍の違い（/R/-/N/-/Q/-X）で対立する4語セットを検索した結果（表1）。6,642セットが抽出された。このセットから表2の条件のものを除外し、31セット（124語）の刺激語を選択した。さらに、親密度の低い語を避けるため、各セット内の最低親密度が高いものから24セット（96語）に絞り、実験に使用した。

表 1: 抽出したターゲット語の例(X:特殊拍なし)

種類	R	N	Q	X
語例	封筒(ふうとう)	奮闘(ふんとう)	沸騰(ふつとう)	不当(ふとう)

表 2: 除外したセット

項目	例
(1) 同音異義語で異なるセット	「せいと（征途）-「せんと（遷都）-「セット」-「せと（瀬戸）」を除外し、「せいと（生徒）-「せんと（遷都）-「セット」-「せと（瀬戸）」を残す
(2) セット内でアクセントが異なるセット	「(せいかい) 正解」-「(せんかい) 旋回」-「(せっかい) 切開」-「(せかい) 世界」を除外する（「世界」のアクセントのみ頭高型であるため）
(3) 外来語が含まれるセット	「カーキ」-「(かんき) 換気」-「(かっき) 活気」-「(かき) 柿」を除外する（「カーキ」は外来語であるため）
(4) 読み方に揺れがある語を含むセット	「(そうこう) 走行」-「(そんこう) 損耗」-「(そっこう) 速攻」-「(そこう) 素行」を除外する（「損耗」を「そんもう」と読むのが一般的であるため）
(5) アクセント型に揺れがある語を含むセット	「(せいき) 世紀」-「(せんき) 戦記」-「(せっき) 石器」-「(せき) 籍」を除外する（「石器」のアクセントには頭高型と平板型との両方があるため）

（* (1)の同音異義語で異なるセットを除外する場合には、同データベースの「親密度」を基準とし、同音異義語の中で親密度が最も高い語を含むセットを残し、他を除外した。）

2.1.2 キャリア文の作成

選択した24セット、96語について3人の音声学に精通した日本語母語話者が意味文と中立文を作成した（表3）。日本語母語話者20人（大学生または大学院生）が意味文の自然さをチェックし、その結果に基づき意味文を修正した。

表 3: 三種の文脈環境の例

WD (単語)	CS (意味文)	NS (中立文)
封筒	明後日の朝 8 時まで、封筒をポストへ投函してください。	今から 5 分の間に、封筒から連想できる単語をすべて書いてください。
奮闘	問題が解決できたのは、奮闘し続けた日々があったからだ。	今から 5 分の間に、奮闘から連想できる単語をすべて書いてください。
沸騰	火傷を負う危険性があるため、沸騰した鍋を素手で触ってはならない。	今から 5 分の間に、沸騰から連想できる単語をすべて書いてください。
不当	従業員が会社を訴えたのは不当な扱いを受けていたからだ。	今から 5 分の間に、不当から連想できる単語をすべて書いてください。

2.1.3 音声収録

音声収録は ATR の可変残響室で残響を最小にした状態で行った。日本語を母語とするアナウンサー 1 人 (男性、40 代) は音声収録用リストを自然なスピードで読み上げ、その音声収録した。

2.2. 実験参加者

中国人日本語学習者 12 名 (20 代) は実験に参加した。全員日本滞在中の留学生であり、かつ神戸大学の大学院生である。全員聴覚的異常がない。その中で、日本語能力試験 N1 合格者は 7 名、N2 合格者は 5 名であり、全員日本語上級者であった。

2.3. 手続き

実験参加者は静かな部屋でパソコンの画面提示に従い、4 択強制選択課題を実行した。4 つの選択肢は拍の種類によってセットとされた 4 つの単語であった。例えば、「ふうとう」-「ふんとう」-「ふつとう」-「ふとう」のようなセットであった。課題は 3 つの部分に分けられた。

2.3.1 単語学習課題

単語学習課題は 6 種類あった：①実験参加者は画面上で提示された漢字表記の単語を見て、セットとなる 4 つの選択肢の中から、その漢字語に該当するふりがなを選択した (以下同様) ②ふりがなから→漢字語③漢字語→中国語の意味④中国語の意味→漢字語⑤ふりがな→中国語の意味⑥中国語の意味→ふりがな。全課題にはフィードバックがあった。

2.3.2 知覚課題

単独単語聴取課題では、実験参加者は刺激語を聞いて、4 つの選択肢から聞いた単語を選択した。刺激語がランダムに提示され、選択肢の表示順はトライアルごとにシャッフルされた。文内の単語聴取課題では、画面上では文の拍を「*」で表記され、文中にあるターゲット語の位置は「 」で表された。実験参加者は刺激文を聞いて、4 つの選択肢から文内で聞いた単語を選択した。選択肢はひらがなで表示された。意味文も中立文も同じブロックに入っていて、ランダムに呈示された。選択肢の表示順序はトライアルごとにシャ

ップルされた。フィードバックはなかった。

2.3.3 意味文における意味情報の利用に対する再認テスト

意味文は音声なしで、文字のみの形で画面上に提示された。実験参加者はひらがなで表記されたセットとなる4つの語から意味文に適切な項目を選出した。

3. 結果

実験参加者ごと、意味文脈の種類・拍の種類による正答率を求め、拍の種類（撥音：/N/・促音：/Q/・長音：/R/・特殊拍なし：I）と意味文脈の種類（単語：WD・意味文：CS・中立文：NS）、を要因に分散分析を行った。その結果、学習者の場合、拍の種類と意味文脈の種類の交互作用がなかったが、拍の種類（ $F(3,30)=4.464, P<.05$ ）の主効果（図1）があった。一方、母語話者はどの拍の語においてもほぼ100%の正答率を示した（図2）。

3.1. 拍の種類

学習者では、正答率はN/（97.8%）>R/（97%）>>Q/（89.7%）>X（88.9%）の順に低かった。多重比較（1%水準）を行った結果、XはN/とR/との差が有意であり、/Q/との差が有意ではなかった。/Q/も同様に、/N/とR/との差が有意であり、Xとの差が有意ではなかった。即ち、Xと/Q/は低正答率であり、/N/と/R/は高正答率であった。

3.2. 特殊拍間の混同

意味文脈の種類ごとに誤聴パターンを観察した結果、どの環境においても特殊拍間の混同が観察された（図3）。

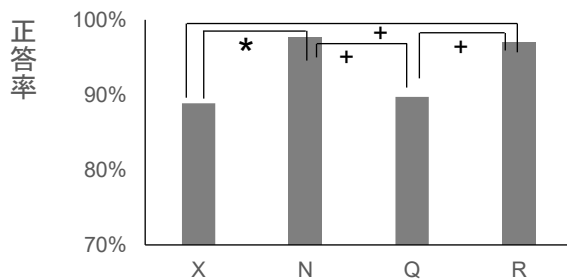


図 1: 中国人学習者の各拍における正答率

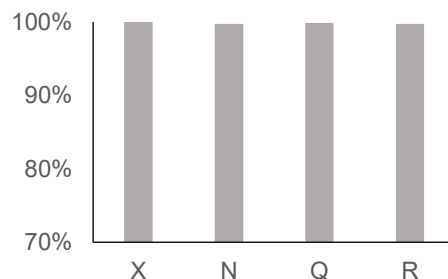
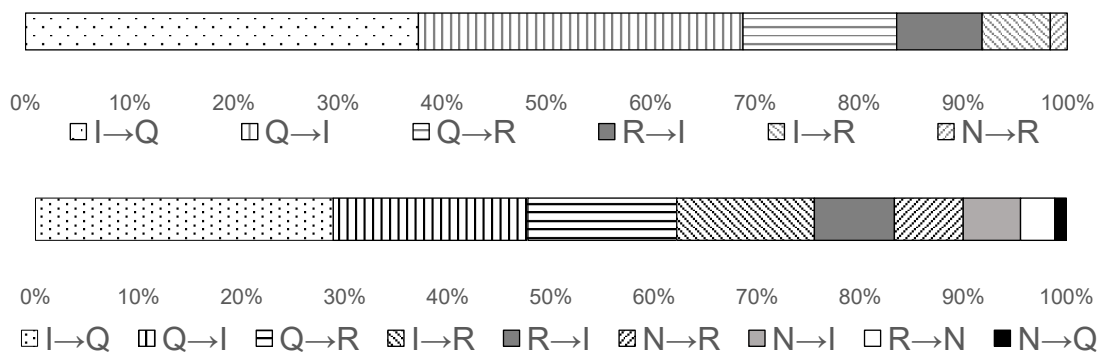


図 2: 日本語母語話者の各拍における正答率



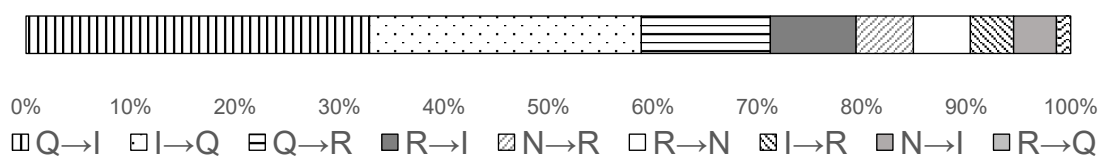


図 3: 各文脈環境における聞き間違いのパターン(上から順にWD、CS、NS)

3.3. 文脈環境

母語話者はどの文脈においてもほぼ 100%の正答率であった。それに対し、学習者全員の結果をみれば、意味文脈の主効果がなかったが、学習レベルに分ければ、違いが出た(図4)。意味文脈の種類は、N1 合格者に影響を与えなかったが、N2 合格者に影響を与えた。WD 環境においても、CS 環境においても、NS 環境においても、N1 合格者の正答率が高く、意味文脈の種類による正答率の差がなかった。CS 環境では WD 環境と NS 環境より若干正答率が高かった。一方、N2 合格者は、CS 環境における正答率は NS 環境より有意に低かった。WD 環境の正答率は CS 環境と NS 環境の中間にあり、両者との差がなかった。N1 合格者と N2 合格者の正答率の差が CS 環境でのみ現れた。意味文脈の意味情報の利用に対する再認テストでは、N1 合格者が高い確率で意味文脈の意味情報を利用できたが、N2 合格者がさほど利用できなかった。両者の利用率には有意に差があった(図5)。

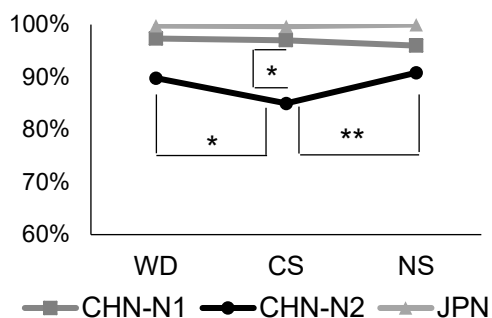


図 4: 文脈による正答率の違い

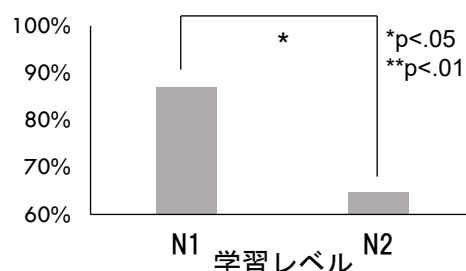


図 5: CS の情報利用に関するテストの結果

以上をまとめると、次のことが分かる。まず、特殊拍の中で/Q/の正答率が最も低かった。なお、特殊拍間においても、知覚的混同が観察された。/Q/を/R/に間違いやすい傾向、/R/と/N/が混同されやすい傾向が見られた。その結果は Zhang et al. (2017) と一致した。また、意味文脈が音声知覚に与える影響は、日本語学習者の言語背景に関わる傾向にある。

4. 考察

3.1 と 3.2 では、中国人日本語学習者の知覚において、正答率は/N/ > /R/ > /Q/の順に低下し(図1)、/N/や/R/は比較的間違いなく同定したのに対し、/Q/の同定は困難だった。また、Zhang, Hayashi, and Akahane-Yamada (2017) と同様、中国人日本語学習者は特殊拍/N/・/R/・/Q/の間でも混同が起こることが示された(図3)。このことから、特殊拍を学習する際には、特殊拍の有無による対立の区別を訓練するだけでは不十分であり、特殊拍同士の区別

についても訓練する必要があることが示された。

3.3 では、意味情報を利用して判断が可能な CS の正答率が、意味情報が使えない WD・NS より高くなることはなく、N2 レベルの中国人日本語学習者では CS のほうが正答率が低かった。Rothwell and Akahane-Yamada (2003) では、日本人英語学習者の /r/ vs /l/ の知覚は正答率が CS > WD > NS の順番に低くなることが示されたが、本実験では異なる結果となった。それについては、本実験における実験参加者の知覚レベルが高かったため、顕著な結果が出なかったことによる可能性がある。Ikuma and Akahane-Yamada (2003) では、日本人英語学習者の /s/ vs /th/ の知覚において、CS の優位性がなく、つまり日本人英語学習者は /s/ vs /th/ のような比較的聞き取りが容易な音韻対立を知覚する際、意味文脈を利用しなかったことが示された。本実験では、中国人学習者の日本語レベルが高かったことにより聞き取りが容易になり、意味文脈を利用しなかった可能性がある。また、意味文が長く複雑だったため、学習者は意味情報を使わず、音声情報に頼って回答し、つまり CS と同じストラテジーで知覚した可能性も考えられる。本実験では、日本語レベルがより高い中国人日本語学習者のほうは正答率が高かった (図 4 と図 5) ことから、学習段階によって音に頼るか意味文脈に頼るかという知覚のストラテジーが変わる可能性があるかと推測できる。今後、学習段階を考慮した検討が必要であることが示されたといえる。

参考文献

- Burton, M. W., Baum, S. R., & Blumstein, S. E. (1989). Lexical effects on the phonetic categorization of speech: The role of acoustic structure. *Journal of Experimental Joint Meeting*, 3pSC21, 953-958.
- Craig, C. H. (1988). Effect of three conditions of predictability on word-recognition performance. *Journal of Speech and Hearing Research*, 31, 588-592.
- Griffin, Z. M., & Bock, K. (1998). Constraint, word frequency, and the relationship between lexical processing levels in spoken word production. *Journal of Memory and Language*, 38, 313-338.
- Ikuma, Y., & Akahane-Yamada, R. (2004). An empirical study on the effects of acoustic and semantic contexts on perceptual learning of L2 phonemes. *The Japan Society of English Language Education*, 15, 101-108.
- Rothwell, A., & Akahane-Yamada, R. (2003). Effect of semantic context on Japanese listeners' perception of English /r/ and /l/. 『日本音響学会講演論文集』, 485-486.
- Zhang, L., Hayashi, R., & Yamada, R. (2017). Confusion between Japanese Special Morae; long vowels/R/, moraic nasals/N/, and geminates/Q/: Perception by Chinese adults learning Japanese. *Chinese journal of phonetics* (in press, written in Chinese).
- 天野成昭・近藤公久 (1999) . 『NTT データベースシリーズ：日本語の語彙特性』 東京：三省堂.
- 戸田貴子 (2003) . 外国人学習者の日本語特殊拍の習得 『音声研究』 7 (2), 70-83.
- 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.

日本人教員の英語発音:英語・日本語・中国語母語話者による 「ふさわしさ」の評価

内田 洋子 (東京海洋大学) 杉本 淳子 (聖心女子大学)
uchidayo@kaiyodai.ac.jp, sugimoto@u-sacred-heart.ac.jp

1. はじめに

中学校・高等学校では英語授業を原則として英語で行うことが求められ、2020年度からは小学校でも教科としての英語教育が行われる動きが進む中、教員の英語運用力およびそれに付随する英語発音の習得が不可欠となっている。それでは、“acceptable pronunciation (教員としてふさわしい英語発音)”とはどのようなものであろうか。

第二言語習得に関する近年の研究では非母語話者が母語話者並みの発音能力を獲得するのは極めて難しいとしており(Derwing & Munro, 2015), 第二言語・外国語音声教育においては“intelligible pronunciation (理解可能な発音)”の習得を目指すので十分という考え方が主流となっている(Celce-Murcia et al., 2010; Derwing & Munro, 2015). しかし、教員自身は国内外を問わず、英語母語話者のような発音を目指すべきであり、なまりのある英語は教員として望ましくないと考える傾向がある(Jenkins, 2007; Uchida & Sugimoto, 2017). 外国語なまりと教員としてのふさわしさ(“teacher suitability”といった表現が使われている)に焦点を当てた研究も、少数ではあるが行われ始めている(Boyd, 2003; Moran, 2016).

学習者や教員といった異なるグループ間で「なまり」と「ふさわしさ」がどのように捉えられているかを知るために、Sugimoto and Uchida (in press) では、日本語母語話者の学生・日本語母語話者の教員・英語母語話者の教員の3グループに、教員免許状取得を希望する日本語母語話者の英語発音の“accentedness (なまり)”と“acceptability (ふさわしさ)”を判定してもらった。その結果、3グループいずれも「なまり」と「ふさわしさ」の判定に高い相関を示した。また、グループにかかわらず、教員としてふさわしくないと判定された発音には共通性があった一方、教員としてふさわしいとされた発音にはばらつきが観察された。

この結果をふまえて、本研究では、英語・日本語・中国語を母語とする聴取者の判定を比較することにより、母語が異なる聴取者が「ふさわしさ」をどのように捉えているか知るための実験を行った。手順はSugimoto and Uchida (in press) を踏襲し、次の2点について検討した：(1) 各グループ内で「なまり」と「ふさわしさ」の判定の仕方に関係性があるか；(2) グループ間で「なまり」と「ふさわしさ」の判定の仕方に違いがみられるか。

2. 実験の方法

2.1. 音声資料

中学校の英語教科書を実験協力者(発話者)に発音してもらい、音声ファイルを作成した。内容は1992年にセヴァン・スズキが国連の環境サミットで行ったスピーチを題材にしたものである(Togo et al., 2013, p. 80)：

- I am here to speak for starving children around the world.
- I am afraid to breathe the air because I don't know what chemicals are in it.
- Did you have to worry about these things when you were my age?

計 40 語から成る 3 つの文は、日本語母語話者にとって難しいとされる単音を含み(例:/ɑ:, αə, l, r, θ, ð/), 適切なリズムやイントネーションの使用ができているかを確認できる文構造を持っている。録音後、著者 2 名が各話者の発音の音声的特徴について記述を行った。

2.2. 実験協力者

発話者：上記音声資料を発音したのは 20 名の東京都内の 2 つの大学いずれかに通う大学生 (男 9 名・女 11 名) である。うち 13 名は教員免許状取得志望者向けの音声学を履修中の学生、残りの 7 名は教員免許状とは関係のない別の授業履修者の録音データ (57 名分) の一部から補充された。Cambridge English Placement Test による CEFR レベルは A2 = 6 名, B1 = 7 名, B2 = 6 名, C1 = 1 名であった。発話者には数日前にスクリプトを渡し、予め意味を確認し練習する機会を与えた。

聴取グループ：録音音声を聴取したのは英語母語話者の教員 10 名 (English Listeners; EL), 日本語母語話者の学生 10 名 (Japanese Listeners; JL), 中国語母語話者の学生 10 名 (Chinese Listeners; CL) の 3 グループ、計 30 名である。EL (年齢幅：37～56 歳、平均年齢：44.7 歳、男 7 名・女 3 名) のうち、4 名は米国出身、2 名はオーストラリア出身、残りの 4 名は英国、カナダ、アイルランド、スペイン出身であったが、母語はいずれも英語であった。全員、日本での生活経験を有し、日本人の話す英語に慣れていて、日本の平均滞在年数は 13.4 年 (年数幅：4.8～20 年)、日本での平均教歴は 12.4 年 (年数幅：4.8～20 年) であった。JL (年齢幅：20～21 歳、平均年齢：20.4 歳、男 0 名・女 10 名) の CEFR レベルは A2 = 7 名, B1 = 1 名, B2 = 2 名, CL (年齢幅：21～26 歳、平均年齢：24 歳、男 5 名・女 5 名) の CEFR レベルは A2 = 3 名, B1 = 4 名, B2 = 2 名, C1 = 1 名であった。CL の日本平均滞在年数は 1.6 年 (年数幅：0.1～3 年)、日本語の平均学習歴は 2.1 年 (年数幅：0.3～5 年)、英語の平均学習歴は 13.6 年 (年数幅：9～19 年) であった。

2.3. 聴取実験の手順

聴取実験に先立ち、聴取者には全員、音声資料のスクリプトを提示した。JL と CL は、教科書の録音音声もあわせて聞いた。

実験は二部構成とした。第一部では、発話者 20 名の音声を聴取者にランダムに提示し、聴取者はそれぞれの発話サンプルが「なまりがある発音か、ない発音か」について、1～9 のスケールで評価した (1 = 非常に強いなまりがある ; 9 = 全くなまりがない)。

第二部では、同じ発話者 20 名の音声を改めてランダムに提示し、聴取者は「英語の先生にふさわしい発音か、ふさわしくない発音か」について、1～9 のスケールで評価した (1 = 全くふさわしくない ; 9 = 非常にふさわしい)。音声の提示は 1 回限りとした。手順に慣れてもらうために、それぞれのパートの最初に 3 問の練習を行った。

第二部終了後、アンケートに回答してもらった。自身の言語背景について記入した後に、

次の2つの質問に答えた：(1)どのような基準で「なまりのある発音」を判断しましたか？；
(2)どのような基準で「英語の教員としてふさわしい発音」を判断しましたか？

上記すべての手続きは、1人あたり30分以内で完了した。なお、アンケートで得られた回答のうち、英語による回答内容(13名分)は著者が、中国語による回答内容(4名分)は言語学を専攻する中国語母語話者の研究協力者が日本語訳を行った。

3. 結果

3.1. 各グループ内の「なまり」と「ふさわしさ」評価の相関

各グループ内では高い評価者間信頼性(Cronbach's $\alpha > .80$)が確認された。各グループの「なまり」と「ふさわしさ」の評価の平均値を算出したところ、「なまり」の評価はEL ($M = 5.3$, $SD = 1.75$), JL ($M = 5.5$, $SD = 2.56$), CL ($M = 5.4$, $SD = 2.34$), 「ふさわしさ」の評価はEL ($M = 6.2$, $SD = 1.64$), JL ($M = 5.3$, $SD = 2.73$), CL ($M = 5.6$, $SD = 2.33$)であった。また、3グループ全てにおいて、「なまり」と「ふさわしさ」の間に強い正の相関が確認できた (EL: *Pearson* $r = .93$; JL: $r = .93$; CL: $r = .94$)。このことから、3グループ全ての聴取者にとって「なまり」と「ふさわしさ」の評価には類似点があることがわかった。

3.2. グループ間の「なまり」と「ふさわしさ」評価の比較

「なまり」と「ふさわしさ」の評価について、各グループの発話者平均を表1にまとめた。「なまり」と「ふさわしさ」に対する聴取グループと発話者の主効果、および聴取グループと発話者の交互作用について、従属変数を「なまり」「ふさわしさ」、独立変数を発話者、聴取グループ、発話者と聴取グループの交互作用とする二元配置分散分析を適用した。

「なまり」は発話者の主効果が有意であった： $F(19, 513) = 29.4, p < .05$ 。全体の平均値はS09に対する評価が最も高かった ($M = 7.2, SD = 1.76$; 表1)。一方、聴取グループの主効果は有意ではなかった： $F(2, 27) = .08, p > .05$ 。聴取グループと発話者の交互作用は有意であった： $F(38, 513) = 3.17, p < .05$ 。

「ふさわしさ」についても発話者の主効果が有意であった： $F(19, 513) = 40.95, p < .05$ 。また、全体の平均値はS08が最も高かった ($M = 7.5, SD = 1.31$; 表1)。聴取グループの主効果については有意ではなかった： $F(2, 27) = 3.03, p > .05$ 。聴取グループと発話者の交互作用は有意であった： $F(38, 513) = 5.55, p < .05$ 。

「なまり」と「ふさわしさ」ともに、聴取グループと発話者の交互作用が有意であったことから、聴取グループごとにそれぞれ評価の高い発話者と低い発話者の検証を、評価の順位(上位3名と下位3名)を用いて行った。その結果、「なまり」については、グループ間で評価の高い発話者に差がみられた。S02 (EL, JLともに2位)とS09 (JL, CLともに1位)の2名を除けば、上位3名はグループ間で一致しているとはいえない。反対に、評価の低い発話者については、グループ間で一致がみられた。ELとCLは下位3名が一致しており、S18, S19, S20は3グループ全てにおいて下位4位内であった。

「ふさわしさ」についても同じ傾向がみられた。各グループ内の上位3名を比較すると、S02 (EL=2位; JL=3位)とS09 (JL=1位; CL=2位)のように部分的な一致はみられる

ものの、グループ間でばらつきがある。反対に評価の低い発話者はグループ間でかなりの一致がみられる。S18, S19, S20 は、3 グループ全てにおいて下位 4 位以内である。唯一の例外は S15 で、EL では 18 位の評価であったものの、JL (10 位) と CL (14 位) では評価が低いわけではなかった。

表 1 : 「なまり」と「ふさわしさ」の評価結果

発話者コード	CEFR	なまり (accentedness) <i>M(SD)</i>				ふさわしさ (acceptability) <i>M(SD)</i>			
		EL	JL	CL	全体	EL	JL	CL	全体
S01	B1	6.7 (1.57)	5.9 (2.42)	6.4 (1.78)	6.3 (1.92)	7.1 (1.29)	6.6 (2.01)	6.8 (1.03)	6.8 (1.46)
S02	B2	6.3 (1.34)	8.4 (0.70)	6.5 (1.90)	7.1 (1.66)	7.1 (1.52)	7.7 (1.83)	6.1 (1.66)	7.0 (1.75)
S03	C1	6.2 (1.69)	6.2 (1.87)	6.3 (2.00)	6.2 (1.79)	7.0 (1.16)	7.2 (0.79)	7.8 (0.79)	7.3 (0.96)
S04	B2	6.1 (1.52)	6.0 (1.76)	5.2 (1.99)	5.8 (1.76)	6.8 (1.14)	6.3 (1.89)	5.2 (1.62)	6.1 (1.67)
S05	B2	6.0 (1.83)	6.5 (2.64)	6.6 (2.01)	6.4 (2.13)	6.8 (1.69)	7.8 (1.87)	6.3 (1.70)	7.0 (1.81)
S06	B1	6.0 (1.41)	6.1 (1.66)	6.4 (1.84)	6.2 (1.60)	6.8 (1.23)	4.9 (1.79)	6.5 (1.27)	6.1 (1.64)
S07	B1	5.8 (1.81)	7.8 (0.92)	6.9 (1.10)	6.8 (1.53)	7.3 (0.82)	6.8 (1.69)	7.7 (0.68)	7.3 (1.17)
S08	B1	5.8 (1.14)	6.6 (2.22)	7.2 (1.62)	6.5 (1.76)	6.7 (1.64)	7.6 (1.08)	8.1 (0.74)	7.5 (1.31)
S09	B2	5.6 (1.58)	8.5 (0.71)	7.4 (1.51)	7.2 (1.76)	6.3 (1.49)	7.9 (1.60)	7.9 (1.45)	7.4 (1.65)
S10	A2	5.5 (1.90)	5.1 (2.42)	7.2 (1.48)	5.9 (2.12)	6.8 (1.32)	4.2 (2.44)	6.6 (2.41)	5.9 (2.37)
S11	A2	5.5 (1.43)	5.2 (2.04)	4.0 (2.11)	4.9 (1.94)	6.8 (0.63)	5.1 (1.52)	4.5 (1.78)	5.5 (1.68)
S12	B1	5.4 (1.84)	5.2 (1.62)	5.9 (1.97)	5.5 (1.78)	6.5 (1.27)	5.8 (2.30)	6.0 (1.41)	6.1 (1.69)
S13	B2	5.4 (1.35)	5.8 (1.48)	4.7 (1.57)	5.3 (1.49)	6.5 (1.08)	5.0 (2.11)	5.1 (1.91)	5.5 (1.83)
S14	B2	5.1 (1.37)	7.1 (0.88)	5.9 (2.85)	6.0 (2.01)	5.6 (1.35)	7.6 (0.84)	6.4 (2.55)	6.5 (1.87)
S15	B1	4.7 (1.42)	5.4 (1.78)	3.7 (1.77)	4.6 (1.75)	4.9 (1.79)	6.2 (1.81)	4.9 (2.38)	5.3 (2.04)
S16	A2	4.4 (1.58)	2.7 (2.06)	3.9 (1.20)	3.7 (1.75)	5.8 (1.14)	2.2 (1.55)	4.4 (1.78)	4.1 (2.10)
S17	B1	4.4 (1.08)	4.2 (2.15)	5.3 (1.77)	4.6 (1.73)	5.9 (1.52)	2.8 (1.23)	4.5 (1.27)	4.4 (1.83)
S18	A2	4.2 (1.62)	2.5 (1.18)	3.1 (1.73)	3.3 (1.64)	5.0 (1.70)	1.5 (0.97)	2.5 (1.35)	3.0 (2.00)
S19	A2	4.2 (1.62)	2.9 (2.13)	2.5 (1.84)	3.2 (1.96)	4.9 (1.60)	2.1 (1.85)	3.0 (1.16)	3.3 (1.92)
S20	A2	2.4 (1.08)	1.1 (0.32)	2.3 (1.64)	1.9 (1.26)	3.5 (1.84)	1.1 (0.32)	1.6 (0.84)	2.1 (1.55)

¹発話者コードはELのなまり評価の高い順 ²青字はグループ内評価上位3名、赤字はグループ内評価下位3名

3.3. 「なまり」と「ふさわしさ」に対するコメント

実験終了後、すべての聴取者に何を基準に「なまり」と「ふさわしさ」を判断したか、自由記述をしてもらった。「なまり」について、言及が多かった項目は次の通りである：イントネーション(13名)、リンキング(11名)、子音(9名)、母音(7名)。「ふさわしさ」については4名が「なまり」と同じ基準(EL=2名、JL=1名、CL=1名)、4名が類似の基準(EL=3名、JL=1名)を用いて評価したと答えた。「ふさわしさ」について、言及の多かった項目は次の通りである：わかりやすさ(7名)、発話速度(7名)、はっきりした発音(6名)、流暢さ(6名)、正確さ(5名)、明瞭度(4名)。

4. 考察

4.1. 各グループ内の「なまり」と「ふさわしさ」評価の相関

3つの聴取グループ全てにおいて「なまり」と「ふさわしさ」には高い正の相関がみられた。つまり、なまりの度合いがより低い発音が英語教員にふさわしく、なまりの度合いがより強い発音は英語教員にふさわしくないと判断される傾向があった。「なまり」が「ふさ

わしさ」を判定する一つの基準であることは確かなようだ。しかし、本実験では「なまり」の後に「ふさわしさ」の判定を行うという手順を採用したため、この順番が結果に影響を与えた可能性も否めない。手順を反対にするなどした追加実験も行う必要がある。

4.2. グループ間の「なまり」と「ふさわしさ」評価の比較

二元配置分散分析により「なまり」と「ふさわしさ」ともに評価に聴取グループ間の差はないことが判明した。これは英語母語話者教員・日本語母語話者教員・日本語母語話者学生の3グループを比較した実験(Sugimoto & Uchida, in press)とは異なる結果であった。他方で「なまり」と「ふさわしさ」両方について、発話者の評価の差を順位で確認したところ、評価が上位の話者にはグループ間の違いが観察されたのに対し、評価が下位の話者はグループ間で比較的一致していることがわかった。この結果は Sugimoto & Uchida (in press) と一致する。今回の研究結果より、英語・日本語・中国語という、母語が異なる聴取グループを対象とした場合も、「なまりの強い・教員にふさわしくない」英語発音のイメージは共有されていることがわかる。さらに、著者2名による音声記述の結果によると、評価の低かった英語発音は、分節音と超分節的要素両方に問題があり、また比較的発話速度が遅いという特徴を共通してもっている傾向があることがわかった。また評価の低かった3名の発話者(S18, S19, S20)はいずれも CEFR レベル A2 であった。しかし全ての A2 レベルの発話者の英語発音が必ずしも低い評価ではなかった。一方、評価の高い英語発話についてはグループ間で差がみられた。

4.3. 評価に差がみられた発話者

グループ間で評価の差が大きかった英語発音の記述結果を3件紹介する。一件目は JL, CL ともに評価が高かったが、EL の評価は低かった S09 である（なまり：JL, CL = 1 位, EL = 9 位；ふさわしさ：JL = 1 位, CL = 2 位, EL = 13 位）。S09 の発音の特徴は、やや誇張しすぎと感じるほどピッチ幅を大きく使ったイントネーションであった。日本語母語話者の英語はしばしば平坦なイントネーションが短所として知られているため、このピッチ幅の大きいイントネーションが非母語話者である中国語・日本語母語話者からは高評価を受けたと推察できる。しかし S09 の発話には /l-/r/ の区別や母音挿入といった分節音の問題点が含まれており、英語母語話者の評価が低かった原因であると推測される。これは過去の研究における報告とも一致する (Riney, Takagi, & Inutsuka, 2005)。二件目は S15 である。この話者は EL の「ふさわしさ」の評価は 18 位と低かったものの、JL は 10 位, CL は 14 位であった。S15 の特徴は、発話中、一貫してそり舌の調音の構えが観察されたことである。これは日本人英語学習者に時々みられる傾向であり、EL の評価が低かったことは興味深い。三件目は S10 である。「なまり」の評価は CL が JL や EL と比べると高く (JL = 15 位, CL = 2 位, EL = 10 位)、「ふさわしさ」の評価では CL/EL と JL に差がみられた (JL = 15 位, CL = 6 位, EL = 5 位)。この話者は /θ, ð/ の発音に多少問題があったものの、それ以外の点では分節音・超分節的要素ともに大きな問題はなかった。他には、ややテンポが速い発音であること、そして少々言い淀みがあったことが特徴としてあげられる。文章も短いため断定的

なことは言えないが、中国語母語話者にとってはこれらの要因が「なまり」と「ふさわしさ」の評価にそれほど影響がなかったと考えられる。

5. おわりに

本研究の結果、「なまり」が「ふさわしさ」の判定に一定の役割を果たすことがわかった。さらに「なまりの強い・教員にふさわしくない」発音のイメージは母語にかかわらず共有されているものの、「なまりの少ない・教員にふさわしい」発音については、聴取グループによって捉え方に差があることがわかった。グループ間の差の中には、母語の異なる非母語話者に共通している点と、母語によって異なる点が存在するようである。

また自由記述のコメントから、「ふさわしさ」の判定には「なまり」以外の要因（例：発話速度、わかりやすさ、明瞭度、流暢さ）が関わることもわかった。実験手続きに改良を加えながら、これらの「なまり」以外の指標と「ふさわしさ」の関係性についても今後研究をする必要がある。英語の非母語話者である日本人教員にとって「ふさわしい発音」の特徴を明らかにすることは、彼らが目指すべき発音を設定する上で極めて重要である。

参考文献

- Boyd, S. (2003). Foreign-born teachers in the multilingual classroom in Sweden: The role of attitudes to foreign accent. *International Journal of Bilingual Education and Bilingualism*, 6(3&4), 283-295. doi:10.1080/13670050308667786
- Celce-Murcia, M., Brinton, D. M., Goodwin, J. M. (with Griner, B.) (2010). *Teaching pronunciation: A course book and reference guide*. New York: Cambridge University Press.
- Derwing, T. M., & Munro, M. J. (2015). *Pronunciation fundamentals*. Amsterdam: John Benjamins Publishing Company. doi:10.1075/llt.42
- Jenkins, J. (2007). *English as a lingua franca: Attitude and identity*. Oxford: Oxford University Press.
- Moran, M. (2016). *Arizona teachers' speech: Phonological features and listener perceptions* (Unpublished doctoral dissertation). Northern Arizona University, Flagstaff, Arizona, US.
- Munro, M. J., Derwing, T. M., & Morton, S. L. (2006). The mutual intelligibility of L2 speech. *Studies in Second Language Acquisition*, 28, 111-13. doi:10.1017/S0272263106060049
- Riney, T. J., Takagi, N., & Inutsuka, K. (2005). Phonetic parameters and perceptual judgments of accent in English by American and Japanese listeners. *TESOL Quarterly*, 39, 441-466. doi:10.2307/3588489
- Sugimoto, J., & Uchida, Y. (in press). Accentedness and acceptability ratings of Japanese English teachers' pronunciation. *Proceedings of 9th PSLLT Conference*.
- Togo, K. et al. (2013). *Columbus 21*. Vol. 3. Tokyo: Mitsumura Tosho.
- Uchida, Y., & Sugimoto, J. (2017, June). *Towards the implementation of ELF-oriented pronunciation teaching in Japan*. Paper presented at the meeting of ELF 10 and Changing English: 10th Anniversary Conference of English as a Lingua Franca, Helsinki, Finland.

日本語母語話者による英語音声の韻律生成と知覚

— 音節を中心に —

江口 小夜子（神戸大学大学院，国際電気通信基礎技術研究所(ATR)）

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を学習する際，英語と日本語の音節構造の違いから韻律の習得が困難な場合がある。例えば，語内の音節数を数える課題では，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報告されている(Tajima and Akahane-Yamada, 2004a)。また，音節知覚に影響を及ぼす諸要因として，子音構成，重子音の位置，音節数，母音構成，黙字，音節主音的子音を対象とし，各要因の影響の強さを比較した結果，音節数の影響が最も顕著であることが示された(江口・山田，2017)。生成に関する研究では，日本語母語話者が発音した英単語を音響的に分析した結果，子音連続間や子音の後に母音が挿入されやすいことが示されている(Tajima and Akahane-Yamada, 2004b)。一方，母音挿入以外の観点から発音を評価した報告は少なく，日本語母語話者が適切な音節数で発音できるかという点については十分な検証がなされていない。通常発話の音声のシラブル数を定量化することは難しいが，本研究では，英単語を音節毎に区切って発音する分節発音課題を考案し，生成面のシラブルカウントを評価した。分節発音課題は，音声呈示および綴り呈示の2つの条件で行った。また，発音課題と同じ刺激語を使用し，音節数をカウントする知覚課題を行い比較するとともに，分節発音課題の妥当性について検討した。

2. 方法

2.1. 実験参加者

日本語母語話者 14 名(男性 7 名，女性 7 名；21 歳～51 歳，平均 24 歳)が参加した。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

発音課題として，英単語を音節毎に区切って発音する分節発音課題を音声呈示および綴り呈示の2つの条件で行った。また，知覚課題として，発音課題と同じ刺激語を使用し，聞こえた英単語の音節数のカウントを行った。以下，音声呈示条件での分節発音課題を ASP 課題 (Auditory Segmented Production 課題)，綴り呈示条件での分節発音課題を OSP 課題 (Orthography Segmented Production 課題)，知覚課題を SC 課題 (Syllable Count 課題) とする (表 1)。

表 1: 課題の種類

略記	課題名	内容
ASP 課題	Auditory Segmented Production 課題	音声呈示条件での分節発音課題
OSP 課題	Orthography Segmented Production 課題	綴り呈示条件での分節発音課題 (音声は再生されない)
SC 課題	Syllable Count 課題	音声呈示条件で 音節数をカウントする知覚課題

分節発音課題では、実験参加者は英単語を音節毎に区切って発音した。その際、音節と音節の間には少し間を空けて発音することが求められた。ASP 課題では、英単語音声再生された後、実験参加者は録音開始/停止ボタンを操作して音声の録音を行った。OSP 課題では綴りが画面上に表示された。音声は再生されなかった。実験参加者は ASP 課題と同様に録音開始/停止ボタンを操作して音声の録音を行った。

SC 課題では、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。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。

2.3. 刺激

親密度の低い英単語を選定し、1~6 音節語の 6 種類(1 σ ~6 σ ; 本稿では、刺激語を構成する音節数を σ で示す)で構成された 48 語を刺激語とした。これらの語は、BNC (British National Corpus)を基準に、出現頻度 250 以下から選んだ。これら語を男性のアメリカ英語母語話者 1 名が発音したものを刺激音とした。

2.4. 手続き

音声または綴り呈示条件の分節発音課題が 2 ブロック、知覚課題が 1 ブロックの合計 3 つのブロックを表 1 の順序で実施した。各ブロック内ではランダムな順に出題した。1 つのブロックは休憩なく実験したが、ブロック間では自由に休憩をとった。

3. 結果

3.1. 分節発音課題における呈示条件の影響

ASP 課題、OSP 課題に対する反応を音節数で比較した。各実験参加者の刺激条件

毎の正答率を従属変数とし、課題（ASP 課題、OSP 課題）と音節数（1σ～6σ）を要因とした 2 要因の分散分析を行った。その結果、課題の主効果[F(1,13)=22.809, p<.001] (図 1)、音節数の主効果[F(5,65)=9.271, p<.001] (図 2)、交互作用[F(5,65)=6.214, p<.001] (図 3) が有意だった。課題については、ASP 課題の正答率が 55.7%と OSP 課題の正答率 71.4%より有意に低く、日本語母語話者は音声呈示条件での発音が困難であった。次に、音節数については、1σ, 2σ, 3σ, 4σ > 5σ > 6σ の順に正答率が低く、音節数が多くなると発音が困難になる傾向が示された。また、課題と音節数の交互作用に関して、ASP 課題では、2～6 音節語と音節数が多くなるほど正答率は低くなったが、OSP 課題では 1～5 音節語の正答率の間に差がなく、音節数の影響が小さいことが示された。

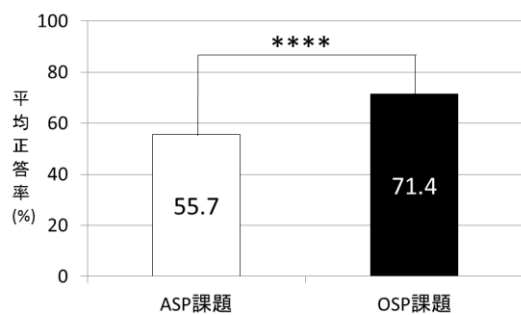


図 1: 課題毎の正答率(%)

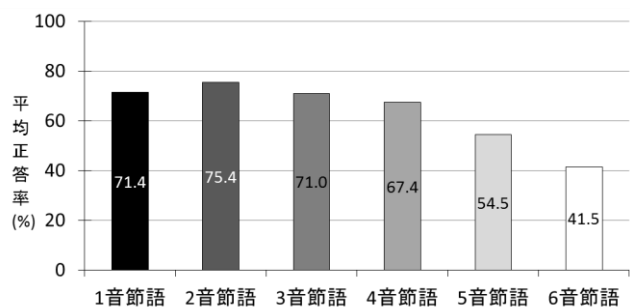


図 2: 音節数毎の正答率(%)

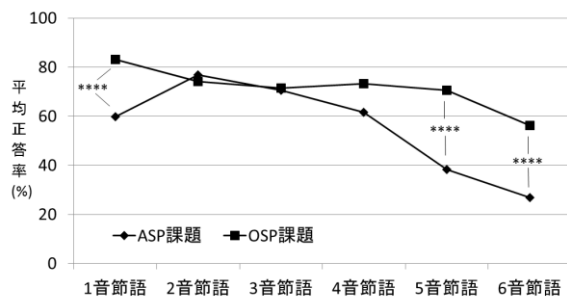


図 3: 刺激条件毎の正答率(%)

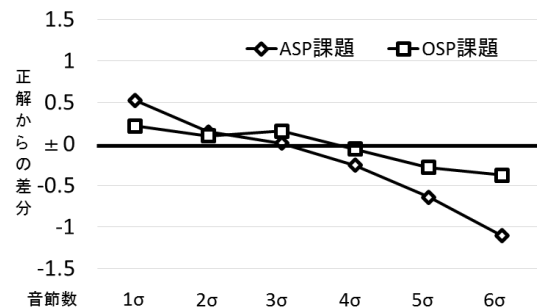


図 4: 正解からの差分

次に、実験参加者が回答した音節数と正解の音節数との差分を求め、各刺激条件に対する反応を比較した。「実験参加者が回答した音節数」－「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、課題の主効果[F(1,13)=8.176, p<.05], 音節数の主効果[F(5,65)=32.400, p<.001], および交互作用[F(5,65)=14.676, p<.001]が有意だった(図 4)。正解からの差分は OSP 課題よりも ASP 課題の方が有意に大きく、音声呈示条件

での発音が困難であることが示された。また、 $6\sigma < 5\sigma < 4\sigma < 3\sigma < 2\sigma < 1\sigma$ の順でプラスの方向に大きく、音節数が多いほど少なめの音節で発音すること、この傾向はASP課題の方が大きいことが示された。

3.2. 知覚課題と音声呈示条件の分節発音課題の比較

SC課題、ASP課題に対する反応を音節数で比較した。各実験参加者の刺激条件毎の正答率を従属変数とし、課題(SC課題、ASP課題)と音節数(1 σ ~6 σ)を要因とした2要因の分散分析を行った。その結果、課題の主効果[F(1,13)=31.024, p<.001](図5)、音節数の主効果[F(5,65)=14.750, p<.001](図6)が有意であり、交互作用はなかった(図7)。課題については、ASP課題の正答率がSC課題より有意に低く、知覚するより発音する方が困難であった。音節数については、両課題ともに音節数が多くなると正答率が低下する傾向がみられたが、1音節語の正答率は2音節語より低かった(図7)。

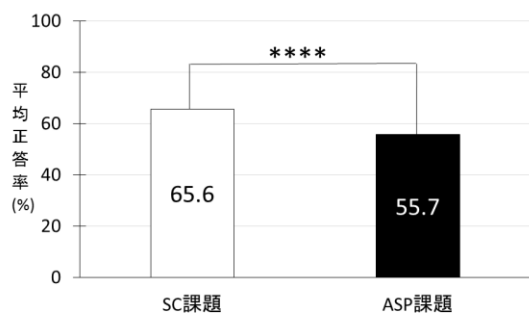


図 5: 課題毎の正答率(%)

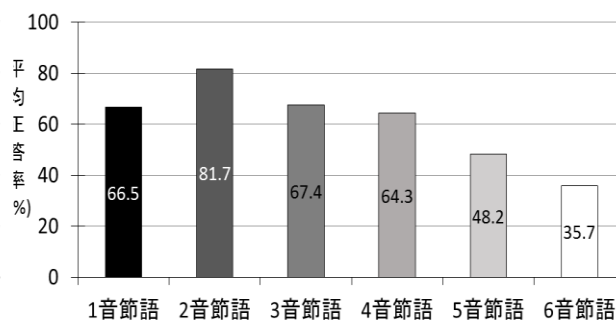


図 6: 音節数毎の正答率(%)

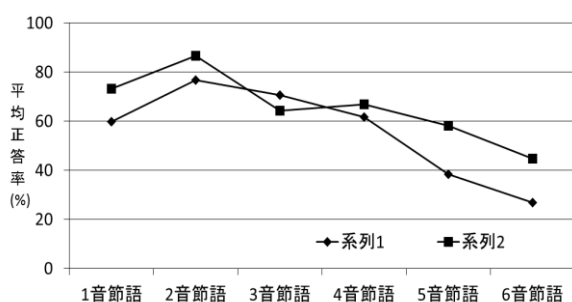


図 7: 刺激条件毎の正答率(%)

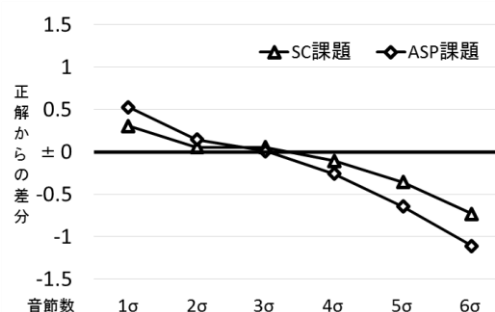


図 8: 正解からの差分

回答と正解の差分(「回答した音節数」-「正解音節数」)を従属変数とした分散分析を行った結果、課題の主効果はなく、音節数の主効果[F(5,65)=37.666, p<.001]、および交互作用[F(5,65)=7.049, p<.001]が有意だった(図8)。音節数が多いほど、少なめの音

節で発音すること、この傾向は ASP 課題の方が大きいことが示された。

3.3. 知覚と生成の関係

知覚の正答率と生成の正答率の関係を分析するため、実験参加者毎の SC 課題と ASP 課題の正答率について、ピアソンの積率相関係数を求めたところ、 $r = 0.93$. ($p < .001$) と高い相関が認められ、SC 課題の正答率が高いほど ASP 課題の正答率が高かった (図 9)。

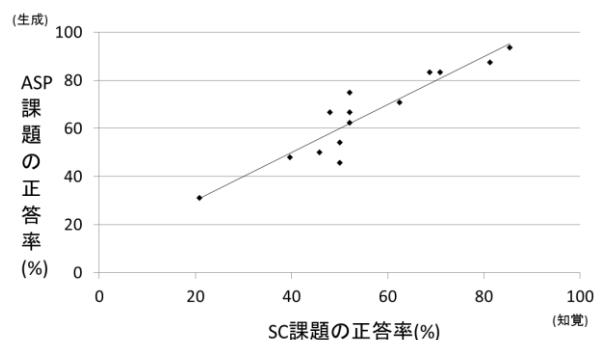


図 9: 刺激条件毎の ASP 課題と SP 課題の正答率 (%)

4. 考察

3.1.では、分節発音課題を行い、音声呈示条件と綴り呈示条件の違いによる影響を調べた。その結果、日本語母語話者は、音声のみを呈示されても適切な音節数で発音することは困難であった。一方、綴りが呈示された場合には正答率が 70%以上と高く、綴りの情報から分節する位置や音節数を推測できた可能性が考えられる。また両条件ともに、生成における音節数の影響が示された。知覚課題を行った Tajima and Akahane-Yamada (2004a)や江口・山田 (2017)の結果においても音節数の影響が報告されており、生成面と知覚面で傾向が一致したことは、生成と知覚のリンクの可能性を示唆する。

3.2.では、発音課題と知覚課題の結果を比較した。その結果、日本語母語話者は、音節数を正しくカウントすることができても、その音節数どおりに発音することは困難であった。また、両課題ともに、1音節語のシラブルカウントが困難であり、知覚課題の結果については、Tajima and Akahane-Yamada (2004a) や江口・山田 (2017) と同様の結果となった。

3.3 では、知覚の正答率と発音の正答率の関係を調べた結果、知覚において音節数を正しくカウントできる人ほど、正しい音節数で発音できることが示された。音韻習得に関する先行研究では、知覚と生成の間で、訓練の効果が互いに転移することが明らかにされている (Bradlow et al., 1997; Akahane-Yamada et al., 1998 など)。今後、韻律習得における知覚と生成の関係についても検討していきたい。

本研究では、分節発音課題を用いることで生成面のシラブルカウントを評価した。通常発話の音声のシラブル数を定量化することは難しいが、分節発音課題を用いるこ

とにより，比較的容易に定量化することができた。一定かつ知覚課題の結果と同様な傾向が得られたことから，その妥当性が示されたと考えられる。

参考文献

- Akahane-Yamada, R., E. McDermott, T. Adachi, H. Kawahara and J. S. Pruitt (1998) “Computer-based second language production training by using spectrographic representation and HMM-based speech recognition scores.” *Proceedings of the 5th International Conference on Spoken Language Processing*, 5, 1747-1750.
- Bradlow, A. R., D. B. Pisoni, R. Akahane-Yamada and Y. Tohkura (1997) “Training Japanese listeners to identify English /r/ and /l/: IV. Some effects of perceptual learning on speech production.” *The Journal of Acoustical Society of America*, 101, No.4, pp.2299-2310.
- 江口小夜子・山田玲子 (2017) 「日本語母語話者による英語音声の音声とストレス位置知覚に及ぼす諸要因」『第30回日本音声学会全国大会予稿集』160-165.
- Tajima, K. and R. Akahane-Yamada (2004a) “Perception of syllables in second-language speech: A comparison of phonetic and phonological factors.” 日本音響学会講演論文集, 467-468.
- Tajima, K. and R. Akahane-Yamada (2004b) “Production and perception of syllable structure in second-language speech.” *Proceedings of the 18th International Congress on Acoustics*, 4, 3321-3324.

中国人日本語学習者による日本語の拗音/Cju/と/Cjo/の母音の知覚

劉奕櫟・任宏昊（早稲田大学大学院） 近藤真理子（早稲田大学）
 {liuyili, haoer21262}@{toki, fuji}.waseda.jp, mkondo@waseda.jp

1. はじめに

第二言語の音声習得には、単音レベルと韻律レベルの二つの側面があると言われている。近藤（2012）は韻律の習得が外国語の音声習得において重要であると主張しているが、個々の音の誤りが、コミュニケーション上の支障をきたすことも報告されている（鮎沢, 2001）。

日本語の単音習得に関する多くの先行研究では、中国人日本語学習者による日本語発音において、母音の混同という問題点が指摘されている（坂本, 2003; 朱, 2005）。中でも、「とうきょう（東京）」を「とうきゅう」と発音するなど、拗音における/u/と/o/の混同が多く報告されている（劉, 1983; 助川, 1993）。この問題は生成の面にとどまらず、知覚の面でも起きている（林, 1981）。先行研究（杉山, 1985; 北村, 1992）によると、これは母語干渉の結果であると考えられているが、それ以外の要因はほとんど言及されていない。

本研究では、中国語を母語とする日本語学習者（以下「学習者」）を対象とし、無意味語の中の拗音の位置や頭子音が、学習者の拗音知覚における/u/と/o/の弁別にどんな影響を与えるのか、日本語の習熟度が異なるグループ別に、静音環境と雑音環境で、調査した。さらに、日本語母語話者（以下「母語話者」）との相違点も分析・考察した。

2. 知覚調査

2.1. 実験参加者

知覚調査の対象は、以下の4つの被験者群で、合計61名の若年健聴者（20代～30代）である。学習者は日本語レベル判定テスト（J-cat）の得点によって、初・中・上級グループに分けられた。

初級学習者群(Beginner learners) : 0–150点（20名）

中級学習者群(Intermediate learners) : 150–250点（19名）

上級学習者群(Advanced learners) : 250点以上（13名）

日本語母語話者群(Native speakers) : 関東圏出身の東京方言話者（9名）

学習者は中国北方方言話者で、中国の大学及び大学院で日本語を専攻する学生である。学習者は全員、留学など長期にわたる日本滞在歴がない。

2.2. 刺激語

拗音の位置と拗音拍における頭子音の音声特徴という二つの要因を調べるために刺激語を選定し、刺激語セットを作成した。参加者に本実験の研究目的を認識されないように、/Cju/と/Cjo/を含まない単語も加えた。全てのテスト語は早稲田大学の音声実験室で関東出身

の東京方言話者 2 名に発話してもらい、デジタルサウンドレコーダー(MarantzPMD 561) およびマイクロホン(SONY F-780)を用いて、サンプリング周波数 44.1 kHz で録音した。

さらに、背景雑音の有無が拗音における/u/と/o/の弁別に与える影響の検証も本研究の目的である。そのため、静音環境で録音した上述の音声データを Praat (Version 6.0.29) のプラグイン Vocal Toolkit で正規化し、背景雑音と合成して雑音を入れた音声データを作成した。本実験では、非母語話者にとって最も聴き取りにくい背景雑音 (Lecumberri & Cooke, 2006) と言われている複数の話者によるバブルノイズを用いた。また、本実験では、先行研究 (Simpson and Cooke, 2005; van Dommelen and Hazan, 2010) を踏まえ、信号雑音比(Signal-to-Noise Ratios, SNR)を-6dB に設定した。

静音下と雑音下での聴取実験を行うことから、意味的・文脈的な手がかりを排除するために、刺激音として 2 音節の無意味語/CjVCa/または/CaCjV/を用いた。C は日本語の[p], [b], [tɕ], [dz], [ɕ], [k], [g], [ç]のいずれかで、それぞれを/Cju/または/Cjo/の中に埋め込み、背景雑音のありとなしの環境で知覚実験を行った。

2.3. 手順

静かな部屋で参加者がヘッドホンを通して刺激語を聴取し、パソコン画面に表示された二つの選択肢から聞こえた音声と最も近い単語を選択した。また、各刺激が 2 回ランダムに呈示された。テスト語の再生と選択肢の呈示には Praat を用いた。

3. 結果

3.1. 背景雑音と母語別

二つの聴取環境での学習者全体及び母語話者の平均正答率を図 1 に示す。

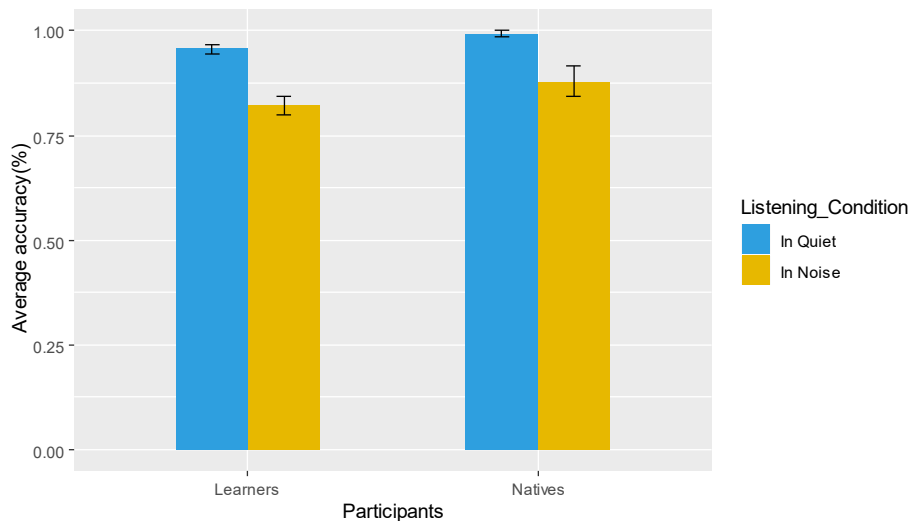


図 1: 母語話者および学習者の平均正答率

静音環境での正答率に比べ、背景雑音を伴う試行の正答率は両グループの間でともに顕著に低下する傾向が見られた。ウェルチの F 検定 (Welsh's F test)を行った結果、聴取環境の

主効果($F(1,89.68) = 140.97, p < .001$)および参加者グループの主効果($F(1,28.85) = 6.47, p = .017 < .05$)が有意であった。参加者グループと聴取環境の交互作用には有意差が認められなかった。従って、先行研究 (Florentine, 1985; 増田, 荒井, 川原, 2014) の結果と同様に非母語話者のみならず母語話者も雑音の影響を受けることが分かった。一方、Takata & Nábělek (1990) の研究において、雑音環境が非母語話者に与える影響は母語話者に与える影響より大きいと指摘されているが、本実験では正聴率の下降率は聴取グループ間で有意差が認められなかった。学習者は雑音なしの環境での正聴率が 96%であり、背景雑音の条件での正聴率は 82%である。この場合の下降率は 14%である。同様の方法で計算した母語話者の下降率は 12%であった。

3.2. 日本語習熟度

異なる聴取環境における学習者の日本語習熟度別の/Cju/と/Cjo/の弁別の平均正答率を図 2-A に示す。

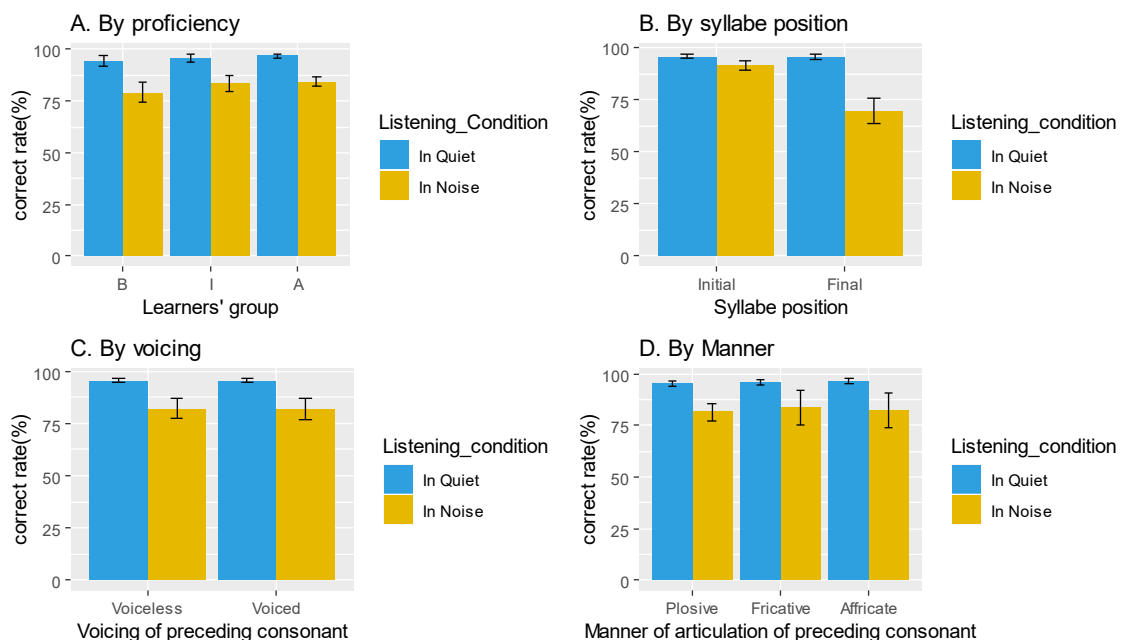


図 2: 日本語習熟度、音節位置及び頭子音の音声特徴別での学習者の平均正答率

次に、二つの聴取環境で聴取者の母語が正答率に与える影響を分析するため、それぞれウェルチ F の検定を行った。その結果、聴取環境に関わらず、学習者グループ間の日本語レベルによる有意差は認められなかった。

3.3. 拗音拍位置

各聴取環境における拗音拍位置別の正答率を図 2-B に示す。雑音を入れない環境では、語頭と語尾の正答率の差が観察されなかった。統計分析の結果から拗音拍位置の主効果は有意ではなかった。

それに対して、雑音がある場合、拗音拍が語尾に来る場合の正答率が大幅に減少した。拗音拍の位置を独立変数とし、平均正答率を従属変数にして異分散 t 検定を行った結果、位置による正答率が有意であった($t(50.7) = 6.67, p < .01$)。

3.4. 頭子音の音声特徴

頭子音の音声特徴を声帯振動の有無と調音方法の二つに分けてそれぞれ雑音あり・雑音なしの条件下分析を行った。その正答率をそれぞれ図 2-C と 2-D に示す。静音環境においても、背景雑音でも、声帯振動の有無と調音方法の主効果が認められなかった。

3.5. 知覚上の類似性 (nativelikeness)

各聴取環境での学習者及び母語話者の平均正答率を図 3 に示す。

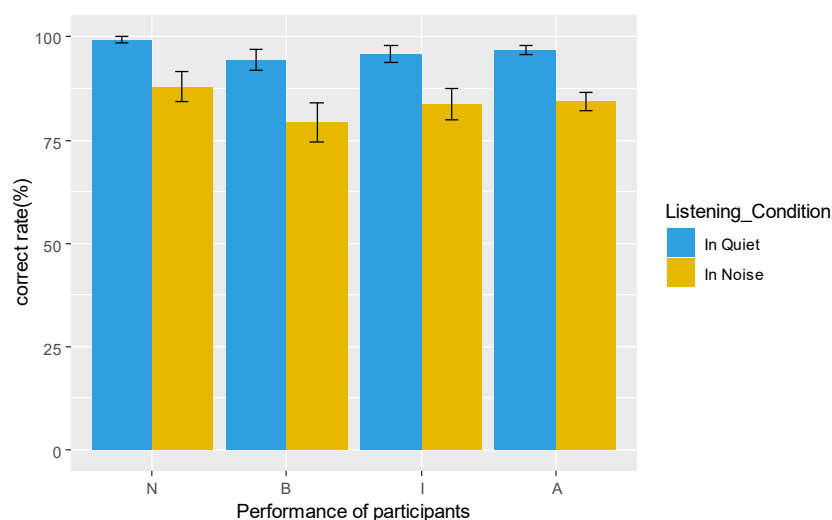


図 3: 学習者習熟度別及び母語話者の平均正答率

3.1 で母語の主効果が有意であったと述べたが、各日本語レベルの学習者と母語話者との類似性を明らかにする必要がある。Games-Howell 法による多重比較検定を行い、どの群同士の平均正答率に有意差が見られるかを調べた。結果、静音環境では母語話者と初・中・上級全てのグループの学習者の間に有意差が認められた。

一方、背景雑音の下での多重比較検定の結果により、母語話者と初級学習者との間で平均正答率に有意差があったが、中級と上級学習者との間には有意差が認められなかった。

4. 考察

この知覚実験の結果では、静音の聴取環境で学習者全体の弁別テストの正答率が 95% を超え、学習者聴取グループの間で有意差が認められなかった。つまり、中国人日本語学習者の知覚において、拗音拍の中の /u/ と /o/ の混同はほぼ発生してないことになる。これは先行研究における記述と一致しておらず、/Cju/ と /Cjo/ の混同は中国人日本語学習者の間に共通する問題点であるとは言えないことになる。全体的に高い正答率の理由として、拗音「ユ」の頻度が考えられる。以下に述べるように、今回の実験では雑音環境での知覚実験を行った

め、無意味語を使ったが、日本語の語彙では/CjuuCV/または/CjuQCV/は一般的だが、/CjuCV/環境を持つ単語は非常に限られており、刺激語を録音した日本語母語話者が発音の難しさを指摘していた。つまり、過剰に調音されていたため、学習者にとっては/CjoCV/と聞き分けられた可能性がある。

一方静音環境では、すべての習熟度群の学習者の正答率と母語話者の正答率の差が有意であった。臨界期仮説によると、ある年齢を超えると言語の習得が困難になる。今回募集した学習者の日本語学習開始平均年齢は18歳で、ほとんどの学習者は大学入学してから日本語を学習しているため、臨界期以降に日本語を学習しているため、母語話者並みの聴き取り能力を獲得している可能性が低い。それに対して、雑音のある環境での弁別実験の結果として、母語話者と初級学習者との間にのみ有意差が認められたが、中・上級学習者の正答率は母語話者に近似している。背景雑音の影響に加え、刺激音としての無意味語は母語話者にとって意味的・文脈的な手がかりにならないため、母語話者の正答率が中級以上の学習者の水準まで下がった可能性がある。

また、雑音を入れた聴取環境で、拗音拍が語頭にきた場合と語尾の場合の正答率に有意差が見られた。語尾に位置にする拗音拍における/u/と/o/の弁別の正答率が低かった原因は、/a/に比べて、/u/と/o/の聞こえ度が小さいため、刺激音の最後にあった場合、学習者の聴き取りが妨げられたためと考えられる。

5. まとめ

本稿では、雑音あり・なしの条件で中国語母語話者52名を対象に、どんな要因が学習者の拗音における/u/と/o/の弁別に影響を与えるかについて調査した。加えて、異なる聴取環境で、学習者と母語話者の聞き分けの差異も考察した。

調査の結果は以下の通りである。(1) 静音環境では、日本語能力にかかわらず、全体的に聞き取りの正答率が高かった。異なる日本語レベル間には有意差が認められなかったが、母語話者と初・中・上級全ての学習者グループの間に有意差が見られた。(2) 雑音の環境では、雑音が強い影響を与えたようで、学習者も母語話者も正答率が顕著に減ったが、学習者グループ間の正答率は類似しているものの、初級レベル学習者と母語話者の間に有意差が見られた。(3) 全学習者グループでは、雑音下で拗音が語頭にcoming環境と語尾にcoming環境では、正答率に有意差が見られた。(4) 頭子音の音声特徴は母音の弁別に影響を及ぼさなかった。

そして、雑音なしの聴取環境で、学習者の平均正答率が100%に近かったが、すべての学習者にとって雑音のある環境で母音の弁別が困難になったと検証された。日本語の習熟度に関わらず、背景雑音下における聴取訓練が必要であることが本研究の結果から示唆される。

また、海外在住経験の長さ(LOR: length of residence)も第二言語音声知覚に影響を与えるため (Flege, 1995)、半年以上の日本滞在歴がある学習者を対象とした、拗音/Cju/と/Cjo/の母音の知覚実態についての調査も必要である。また、今回の調査では知覚の面のみ考察した

が、今後は学習者産出上の拗音における/u/と/o/の混同の有無について検討し、産出と知覚の関連性を明らかにしたい。

参考文献

- 鮎澤孝子 (2001) 「日本語教育のための音声研究」『音声研究』, 5(1), 71-74.
- 北村よう(1995) 「中国語話者に対する発音指導」『東海大学紀要』, 15, 77-79.
- 近藤真理子 (2012) 「日本語学習者の音声習得における第一言語特有の干渉と普遍言語的干渉」『早稲田大学大学院文学研究科紀要』, 3, 21-34.
- 坂本恵 (2003) 「中国人学習者のための発音指導について」『東京外国語大学留学生日本語教育センター論集』, 29, 171-181.
- 朱春躍 (2011) 「中国語話者の日本語「ユ」はなぜ「ヨ」に聞こえるのか」『音声文法』東京：くろしお, 103-122.
- 助川泰彦 (1993) 「母語別に見た発音の傾向－アンケート調査の結果から－」『日本語音声と杉山太郎 (1985) 「日本語の発音－中国語の発音の学習から」『日本語教育』, 55, 97-110.
- 林佐平 (1981) 「初学段階における日本語の音声教育－中国人初心者の聴音問題点と母国語の干渉について」『日本語教育』, 45, 133-144.
- 増田斐那子, 荒井隆行, 川原繁人 (2014) 「日本語母語話者によるバブルノイズ下の母音間英語子音知覚: 英語習熟度と子音知覚の相関関係」『日本音響学会誌』, 70, 284-287.
- 劉淑媛 (1983) 「中国人学習者によく見られる発音上の誤りとその矯正方法」『日本語教育』, 53, 99-101.
- Flege, J. E. (1995). Second-language speech learning: Theory, findings, and problems. *Speech perception and linguistic experience*.
- Lecumberri, M. G., & Cooke, M. (2006). Effect of masker type on native and non-native consonant perception in noise. *The Journal of the Acoustical Society of America*, 119(4), 2445-2454.
- Simpson, S. A., & Cooke, M. (2005). Consonant identification in N-talker babble is a nonmonotonic function of N. *The Journal of the Acoustical Society of America*, 118(5), 2775-2778.
- Takata, Y., & Nábělek, A. K. (1990). English consonant recognition in noise and in reverberation by Japanese and American listeners. *The Journal of the Acoustical Society of America*, 88(2), 663-666.
- Van Dommelen, W. A., & Hazan, V. (2010). Perception of English consonants in noise by native and Norwegian listeners, *Speech Communication*, 52(11-12), 968-979.

Voiced geminates in Indo-Aryan languages: Acoustic evidence from Punjabi

Qandeel Hussain (North Carolina State University, Raleigh)
qhussai@ncsu.edu

1. Introduction

Geminates are described as long consonants that differ from their singleton counterparts in terms of duration (Davis, 2011). Cross-linguistic studies have shown that voiceless geminates are more frequent than voiced geminates (Blevins, 2004). In some languages (e.g., Tokyo Japanese), voiced geminates are partially devoiced (Kawahara, 2015). This devoicing is generally attributed to the aerodynamic constraints. From an articulatory perspective, maintaining voicing and long closure duration in voiced geminates is challenging (Ohala, 1983). Punjabi, an Indo-Aryan language, has been reported to contrast voiceless/voiced singletons and geminates (Bhatia, 1993). However, it is not known if voiced geminates in Punjabi are completely voiced and how their duration differs from voiceless geminates.

Geminates in Punjabi are restricted to word-medial position and are always preceded by short vowels (Bhatia, 1993). Singletons, on the other hand, can occur freely with both long (/i e ε a o ɔ u/) and short (/ɪ ə ʊ/) vowels. The most important acoustic correlate of the word-medial geminate stops of Punjabi is closure duration (Hussain, 2015). This is evident from other languages as well where closure duration of geminates has been shown to differ significantly from singletons (Hindi: Ohala, 2007; Bengali and Turkish: Lahiri & Hankamer, 1988). There are also durational differences in singletons and geminates in terms of voicing. Voiced singletons and geminates are generally shorter than voiceless singletons and geminates (Siene Italian: Stevens & Hajek, 2004).

Although closure duration is a primary acoustic correlate to the distinction between singletons and geminates, speakers may use the duration of preceding vowel as an additional cue to gemination. In Makasar, preceding vowels are consistently shorter before geminates than before singletons (Tabain & Jukes, 2016). Other languages that shorten the preceding vowels include Italian (Esposito & Di Benedetto, 1999), Hindi (Ohala, 2007) and Bengali (Lahiri & Hankamer, 1988). There are very few languages that lengthen the preceding vowels when the following consonant is a geminate. Japanese is usually reported to show this kind of durational pattern. Idemaru and Guion (2008) demonstrated that the duration of preceding vowels in Japanese is longer before geminates (75 ms) than before singletons (59 ms). The duration of the

preceding vowels can also vary depending on the voicing of the following consonant. Vowels are longer before voiced than voiceless consonants (Chen, 1970).

The duration of the following vowels has also been reported to serve as a cue to quantity distinction between singletons and geminates. The mean duration of the following vowels in Japanese is 63 ms after geminates but 76 ms after singletons (Idemaru & Guion, 2008). Similarly, in Punjabi, the duration of following vowels is longer in voiceless singletons than in voiceless geminates (Hussain, 2015). Italian, unlike Japanese and Punjabi, does not show any effect of gemination on the following vowels (Esposito & Di Benedetto, 1999).

The aim of the current study is to investigate the voicing (voiceless vs. voiced) and length (singleton vs. geminate) contrasts in word-medial stops of Punjabi. In particular, we investigate how voicing affects the durational properties of the word-medial stops and the flanking vowels.

2. Methods

2.1 Participants

Four male Punjabi speakers participated in the experiment (24 to 26 years of age, mean 24.7 years). All the participants spoke the Lyallpuri dialect of Punjabi that is widely used in Faisalabad and surrounding areas. In addition to Punjabi, the participants were also fluent in Urdu and had some knowledge of English.

2.2 Stimuli and experimental procedure

The stimuli consisted of six pairs of word-medial voiceless/voiced singleton and geminate stops, with C1V1C2V2 template (C1: /k^h/; V1: /ə/; C2: singleton /p b t̪ d̪ k g/ or geminate /p: b: t̪: d̪: k: g:/; V2: /a/).¹ All the words were elicited in citation form. Each word was repeated five times, resulting in 240 tokens (12 words × 5 repetitions × 4 speakers = 240). The participants were invited into a quiet room at the Corpus Linguistics Lab, Government College University, Faisalabad. Before the experiment, a practice session was conducted to familiarize participants with the task. Target words were written in Punjabi orthography (Shahmukhi script) and presented on a computer screen. The recordings were made using a Zoom H6 digital voice recorder with a built-in microphone at a sampling rate of 44.1 KHz.

2.3 Acoustic and statistical analyses

A total of 240 tokens were segmented in Praat (Boersma & Weenink, 2014). One token was excluded due to mispronunciation. The preceding (V1) and following (V2) vowels

¹Punjabi also contrasts retroflex and palatal stops that are not reported here.

of the remaining 239 tokens were segmented as clear F2 energy in the spectrogram. Closure durations and VOTs of C2 were collapsed to get the total C2 duration (Hussain, 2015). A repeated-measures ANOVA was conducted using the `aov()` function in R (R Core Team, 2012). V1, C2, and V2 durations were used as dependent variables, place (labial vs. dental vs. velar), voicing (voiceless vs. voiced), and length (singleton vs. geminate) were used as independent variables (p value was set at 0.05).

3. Results

Figure 1 shows the durations of preceding vowel (V1), word-medial consonant (C2) and following vowel (V2) across voicing (voiceless vs. voiced), length (singleton vs. geminate), and place (labial vs. dental vs. velar). Statistical results are presented in Table 1. The output of the repeated-measures ANOVA indicated that there was an effect of voicing, length, and place on the duration of preceding vowel (V1). There were no significant interactions in any of the factors for V1. The duration of C2 varied across voicing and length, and there was a significant interaction between voicing \times place, length \times place, and voicing \times length \times place. The effect of place on C2 duration was not significant. The duration of the following vowel (V2) was significantly different across consonantal length and places of articulation. However, no effect of voicing was observed on V2 duration.

Table 1: Results of repeated-measures ANOVA with V1, C2, and V2 durations as dependent variables, voicing (voiceless vs. voiced), length (singletons vs. geminates), and place (labial vs. dental vs. velar) as independent variables. Bold values indicate significant results with the p value of 0.05.

Factors	V1			C2			V2		
	df	F	p	df	F	p	df	F	p
Voicing (V)	1	9.19	=0.002	1	172.33	<0.001	1	0.01	=0.918
Length (L)	1	18.22	<0.001	1	1226.84	<0.001	1	67.49	<0.001
Place (P)	2	43.83	<0.001	2	1.21	=0.297	2	9.49	<0.001
V \times L	1	0.01	=0.915	1	3.17	=0.076	1	0.24	=0.620
V \times P	2	0.22	=0.798	2	5.19	=0.006	2	0.92	=0.399
L \times P	2	0.85	=0.424	2	10.31	<0.001	2	0.59	=0.552
V \times L \times P	2	1.76	=0.174	2	4.96	=0.007	2	0.19	=0.819

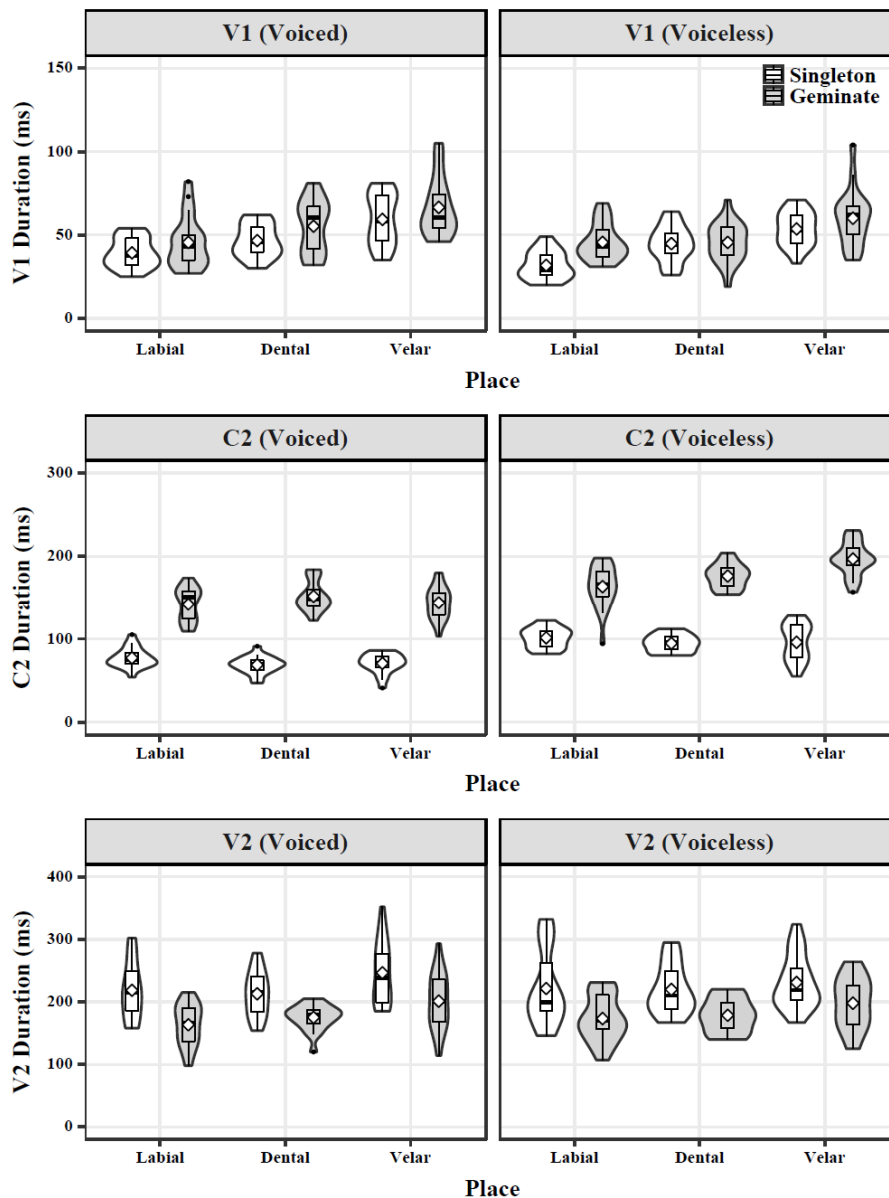


Figure 1: Violin (box) plots of preceding vowel (V1), word-medial consonant (C2), and following vowel (V2), across voicing (voiceless vs. voiced), length (singleton vs. geminate), and place (labial vs. dental vs. velar). White diamonds indicate the mean.

4. Discussion

The current study examined the durational correlates of voicing and length in word-medial stops of Punjabi. The results indicate that geminates are longer than their singleton counterparts. In terms of voicing, voiced geminates are shorter in duration compared to voiceless geminates. This is in line with other studies where voiced

geminate were also reported to have short duration (Sienese Italian: Stevens & Hajek, 2004).

The duration of the preceding vowel (V1) varied depending on the voicing (voiced > voiceless) of the word-medial stops (C2). This is consistent with the previous studies that showed lengthening of preceding vowels when the following consonant is voiced (Chen, 1970). It was also noted that V1 duration before geminates was slightly longer than before singletons, as found in Japanese (Idemaru & Guion, 2008). The duration of the following vowel (V2) was longer when C2 was a singleton. These findings indicate that Punjabi speakers differentiate the consonantal length and voicing contrasts by adjusting the duration of the consonants as well as preceding and following vowels.

References

- Bhatia, T. K. (1993) *Punjabi: A cognitive descriptive grammar*. London and New York: Routledge.
- Blevins, J. (2004) *Evolutionary phonology: The emergence of sound patterns*. Cambridge: Cambridge University Press.
- Boersma, P., and Weenink, D. (2014) *Praat: Doing phonetics by computer*. <http://www.fon.hum.uva.nl/praat/>
- Chen, M. (1970) "Vowel length variation as a function of the voicing of the consonant environment." *Phonetica* 22, 129–159.
- Davis, S. (2011) "Geminates." In M. van Oostendorp, C. J. Ewen, E. Hume and K. Rice (eds.) *The Blackwell Companion to Phonology*. (pp. 837–859). Malden, MA & Oxford: Wiley-Blackwell.
- Esposito, A., and Di Benedetto, M. G. (1999) "Acoustical and perceptual study of gemination in Italian stops." *Journal of the Acoustical Society of America* 106(4), 2051–2062.
- Hussain, Q. (2015) "Temporal characteristics of Punjabi word-medial singletons and geminates." *Journal of the Acoustical Society of America* 138(4), EL388–EL392.
- Idemaru, K., and Guion, S. (2008) "Acoustic covariants of length contrast in Japanese stops." *Journal of the International Phonetic Association* 38(2), 167–186.
- Kawahara, S. (2015) "The phonetics of obstruent geminates, sokuon." In H. Kubozono (ed.) *The Handbook of Japanese Language and Linguistics: Phonetics and Phonology*. (pp. 43–73). Berlin: De Gruyter.
- Lahiri, A., and Hankamer, G. (1988) "The timing of geminate consonants." *Journal of Phonetics* 16, 327–338.

- Ohala, J. J. (1983) “The origin of sound patterns in vocal tract constraints.” In P. F. MacNeilage (ed.) *The Production of Speech*. (pp. 189–216). New York: Springer-Verlag.
- Ohala, M. (2007) “Experimental methods in the study of Hindi geminate consonants.” In M. J. Sole, P. Beddor, and M. Ohala (eds.) *Experimental Approaches to Phonology*. (pp. 351–368). Oxford: Oxford University Press.
- R Core Team (2012) *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. Version 3.4.2: <https://cran.r-project.org/bin/windows/base/old/3.4.2/>
- Stevens, M., and Hajek, J. (2004) “Comparing voiced and voiceless geminates in Sienese Italian: What role does preaspiration play?” In *Proceedings of the 10th Australian International Conference on Speech Science & Technology*. (pp. 340–345). Macquarie University, Sydney, Australia.
- Tabain, M., and Jukes, A. (2016) “Makasar (illustrations of the IPA).” *Journal of the International Phonetic Association* 46(1), 99–111.

中国語母語話者による母音長の知覚と産出の関係

張 格格 (横浜国立大学大学院)
foamofthedream@gmail.com

1. はじめに

日本語学習者の特殊拍の問題に関する研究は昔から、多くの研究者たちによって様々な視点から研究がなされている。これらの研究が長く続けられている原因としては、特殊拍は独立した拍を形成しながらも、独立した音節を形成しないということが挙げられる(窪菌 1998)。つまり、長さとしての自立性を持っているが、単独では音節にならないと考えられる。また、学習者の習得状況について、知覚判断に曖昧な領域が広くて、長音と短音の知覚範疇化が進んでいない学習者は上級になっても存在する(戸田 1998b)。そのため、学習者の日本語の特殊拍を把握しにくいと考えられる。

日本語学習者としての筆者の経験でも、また周りの学習者から聞いた話でも、まず長音・短音の長さの問題について、先生からの指導がほとんどなかった。そして、長音については、聞くときに、弁別できるが、話すときには長かったり、短かったりしてしまうことがある。逆も同じく、言うときには正しく言えるが、聞くときに、長音を脱落して聞く。また、長音がないときに、長音に聞こえてしまうこともある。そのため、母語話者とコミュニケーションするとき、それが意図伝達の障害となり、誤解を招く可能性があると考えられる。このような問題があるときに、産出や知覚をそれぞれ練習すれば良いのか、それとも聞く話すを一緒に練習できるシャドーイング練習をすれば良いのか、練習方法がわからない。

本研究では、このような原因を踏まえ、長音を中心に、中国語を母語とする日本語学習者の自然発話の音声データと聴取テストの結果をもとに、聞き分ける能力と言い分ける能力の状況を分析し、その関連性を考察することを目的とする。

2. 先行研究

小熊(2008)は、初級、中級、上級、超級の英語、中国語、韓国語を母語とする日本語学習者を対象に、インタビュー形式の自然発話を分析し、発話リズムの不自然さの全体的な傾向を検討した。その結果、全体的な傾向として、拍の増加、拍の減少のうち、短音が長音に聞こえる、長音が短音に聞こえることの数突出して多いことが観察された。その中、中級・上級学習者の不自然さの数が初級・超級学習者より多いことが見られた。さらに、日本語能力が上級レベルの学習者や上級レベル以上の学習者の発話に対しても指摘が多く、母音を短く聞こえるように産出することは習得が難しいことが推測される。

山田(1999)では、日本語を母語とする英語学習者成人 10 人を対象として、英語/r/と/l/のペアの聞き分ける練習をして、訓練前に約 7 割程度だった正解率が、訓練後のテスト

では約9割まで上昇し、訓練で使われなかった単語や、訓練で用いなかった人の声でも9割程度聞き取れるようになった結果が見られた。さらに、リスニングの訓練しか行わなかったにもかかわらず、発音にも進歩があることが観察された。つまり、英語の場合には、知覚能力と産出能力の間には一方の上昇につれ、もう一方も上昇するという相関関係があると考えられる。

3. 調査方法

3.1. 実験参加者

中国語を母語とする、日本語中級学習者5名、上級学習者5名（20～30代）を対象者とした。中級学習者は全員日本語能力試験N2の合格者で、日本の日本語学校に在籍している。上級学習者全員が日本語能力試験N1に合格し、中国の大学で日本語を専攻とし、来日前に4年以上の日本語学習歴を有している。現在、日本の大学院に在籍している大学院生である。対象者の音声評価に参加する日本語母語話者4名（20～40代）は全員日本語教師である。

3.2. 知覚調査

学習者の知覚能力を測定するために、学習者の既習知識に影響される可能性があると考えられる、無意味語の長短音の弁別聴取テストを行った。そして、母音の無声化が学習者にも影響を与えるので、それを考慮した上で、弁別テストの内容として、無意味語のミニマルペア、4種（AB・AB―/AB・A―B/AB―・A―B―/A―B・A―B―）、各12ペア、計48ペアを用いた（表1に弁別テストに用いた無意味語ペアの例を示す）。単語内にアクセント型が学習者の弁別に影響を与えられると考えられる（皆川 1995a; 1997）。そのため、無意味語をそれぞれの語を平板型で読み上げることにした。

音声モデルを作成にあたって、東京語母語話者（男性、20代）が無意味語リストにある96語をランダムで、それぞれ3回読み上げ、録音し、音声モデルを作成した。

作成した音声モデルを対象者に聞かせ、4つの選択肢から強制選択させた。繰り返し3回、計288語を解答用紙に記入してもらった。

表 1: 弁別テストに用いた無意味語ペアの例

	AB・AB―	AB・A―B	AB―・A―B―	A―B・A―B―
1	のま・のま―	さそ・さ―そ	せて―・せ―て―	れ―ま・れ―ま―
2	はよ・はよ―	さね・さ―ね	まも―・ま―も―	へ―け・へ―け―
3	ぬも・ぬも―	らも・ら―も	もね―・も―ね―	な―と・な―と―
...

3.3. 産出調査

自然状況の発話モデル中の産出能力を測定するために、対象者に1分間以内に自由発話をさせ、録音した。その際、対象者自身の関心や興味を持っている内容をテーマにした。発話を中断しないように、メモやキーワードなどを書いておいた。

自由発話の録音データについて、日本語母語話者が話者の意図の理解と意味不明な単語があるかどうか、もう一つは発話の中に、母音の長・短の問題があるのかという二つの側面から評価し、1～5の範囲の点数をつける。(1：全然できていない；5：母語話者レベル)

4. 結果と考察

4.1. 知覚能力

本研究の調査結果について、まず、学習者の知覚能力を測定する母音長の弁別テストの結果を図1、図2に示す。上級学習者全体の平均誤答率は12.7%である一方、中級学習者は39.9%であり、中級学習者と上級学習者の間には顕著な差が見られた。上級学習者になって、誤りの減少が見られることから、学習者の日本語の習熟度の向上に従って、知覚能力も上昇すると考えられる。

4.2. 産出能力

対象者の自由発話について、アクセントやイントネーションなどの音声要素を除き、母音の長短だけを分析対象とした。その結果を図3に示す。中級学習者と上級学習者の間に、知覚能力のような大きな差がなかったと言える。しかし、日本語母語話者により、発話の全般的なイメージとして、上級学習者の方は自然である一方、中級学習者の発話は全体的に不自然な部分が多いと考えられる。

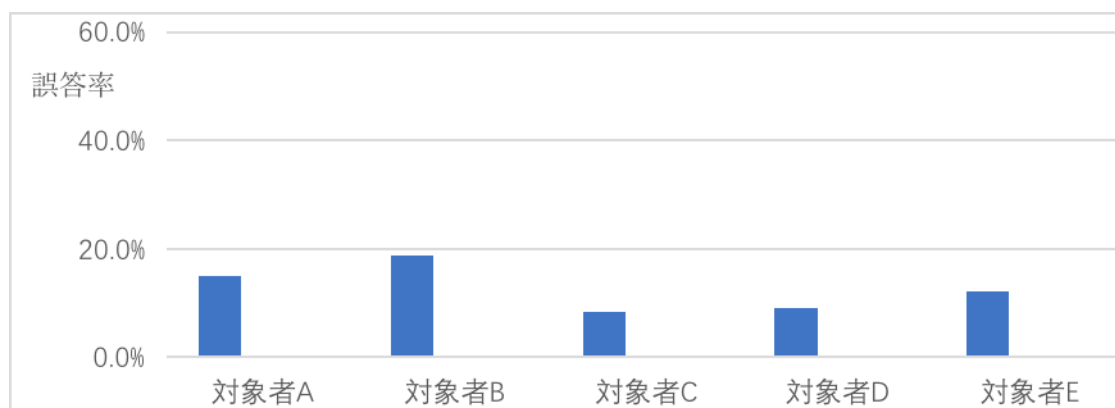


図 1:上級者の弁別テストの誤答率

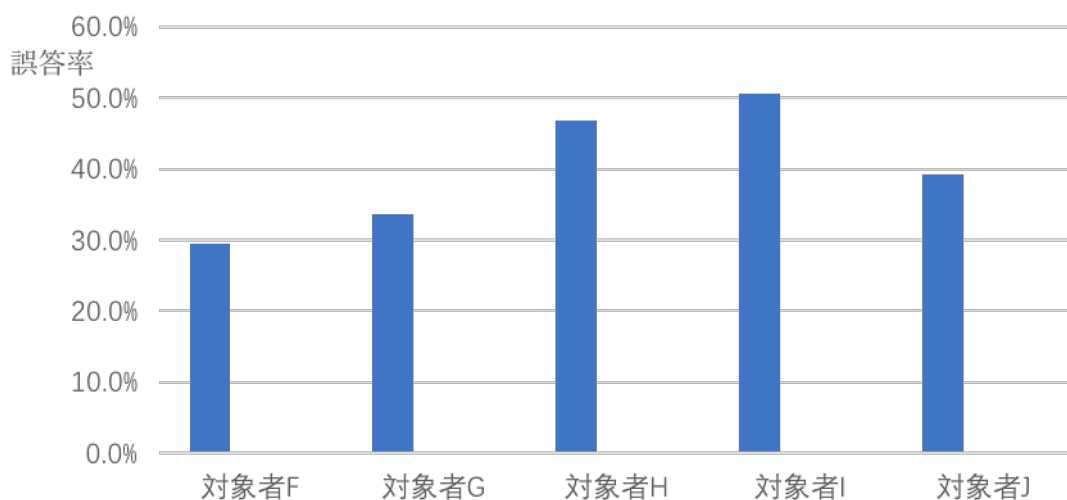


図 2:中級者の弁別テストの誤答率

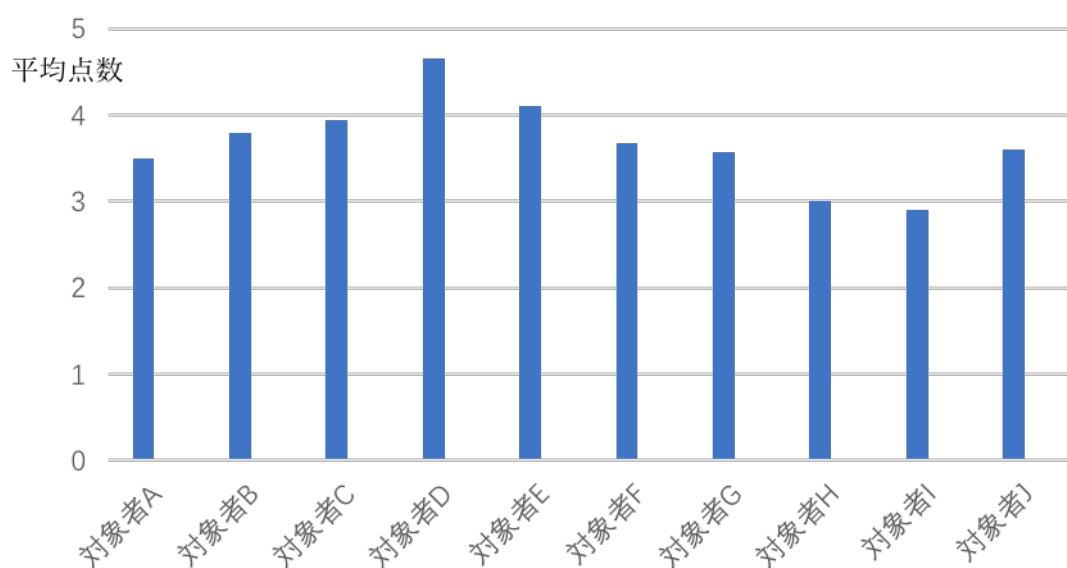


図 3:産出能力の平均点数

4.3. 知覚能力と産出能力の相関

図 4 に知覚能力と産出能力の相関を示す。学習者の日本語習熟度の上昇に従い、産出能力と知覚能力の間に正の相関関係があると考えられる。しかし、個人差によって、正の相関関係があると判断できない状況もあると考えられる。

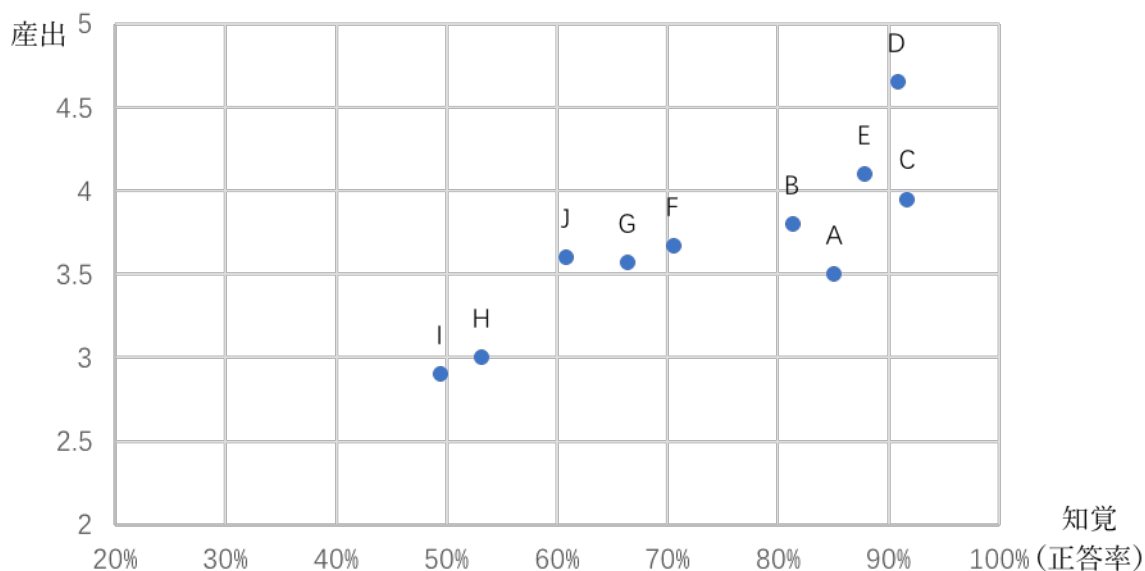


図 4：知覚能力と産出能力の相関

5. 終わりに

本研究では、中国語を母語とする中級・上級日本語学習者の母音長短の知覚能力と産出能力を測定し、二つの能力の相関関係について検討した。知覚に関しては、中級学習者と上級学習者の間に激しい差が存在しているのに対して、産出に関しては、今回対象者となる中級・上級学習者の発話は、長音・短音の誤用があるにもかかわらず、意思疎通上の問題になっていない。つまり、長音を短くしてしまい、または、短音を長く発音してしまう場合でも、母語話者が理解できる範囲の中にあると判断できる。上級者を分析したところ、個人差により、知覚の問題になるパターンは、正しく産出できることが観察された。現段階では、多少ばらつきがあるものの、概観すると、言い分ける能力と聞き分ける能力の間に正の相関関係があると考えられる。しかし、個人差によって、正の相関関係が観察できない状況もあると判断したため、今後、調査人数を増やし、知覚能力と産出能力の相関を明らかにしていきたい。

参考文献

- 皆川泰代(1995a)「日本語学習者における長音知覚の諸要因：英語・韓国語話者の場合」『平成7年度日本音声学会全国大会予稿集』52-57.
- 皆川泰代(1997)「長音・促音の識別におけるアクセント型と音節位置の要因：韓国・タイ・中国・英・西語母語話者の場合」『平成9年度日本語教育学会春季大会予稿集』123-128.

- 窪菌晴夫(1998)「モーラと音節の普遍性」『音声研究』2:1.
- 戸田貴子(1998b)「日本語学習者による促音・長音・撥音の知覚範疇化」『文藝言語研究言語篇』33.
- ATR 人間情報通信研究所(1999)「完全版 英語リスニング科学的上達法音韻編」講談社
- 小熊利江(2008)『発話リズムと日本語教育』風間書房.

中国語を母語とする日本語学習者による感情音声の 知覚と生成の関係

李 歆玥 (神戸大学大学院) Aaron ALBIN・林 良子 (神戸大学)

lixinyue12200@163.com・albin@people.kobe-u.ac.jp・

rhayashi@kobe-u.ac.jp

1. 背景と目的

音声による感情伝達は文化を超えた普遍性があることが知られている (Scherer et al. 2001)。一方、第二言語の音声における感情の表出は、母語話者によるそれと異なることも報告されている (甲斐他 2003)。また、声の高さ・長さ・強さ・声質などの複合的な音響特徴が感情の知覚に関与していることも指摘されている (エリクソン他 2006)。

音声による感情伝達は不可欠なコミュニケーションスキルであり (Erickson 2005)、音声コミュニケーションによる意思疎通の誤解を避けるためには、外国人学習者にとっても同様に重要であると思われる。例えば、中国語を母語とする日本語学習者が喜んでいるのに、日本人には怒っているように聞こえたといった事象をよく耳にするが、これも感情伝達の失敗による誤解であると考えられる。

桜庭他 (2001, 2004) では、日本児と米国児が幸福、悲しみ、怒りおよび平静の感情を込めた発話/pikachuu/を、日本語母語話者とアメリカ英語母語話者に聞かせ、音声を示す感情を選択させた。その結果、米国児の発話よりも、日本児が発話した感情音声の同定率が高かった。また、甲斐他 (2003) では日本語母語話者と韓国人日本語学習者による感情音声をを用い、日本語母語話者と韓国人日本語学習者に聞かせる知覚実験を行なった。その結果、感情同定率が 80%以上の音声の数を見ると、日本語母語話者の発話のほうが韓国人日本語学習者より多かった。これらの先行研究は、ある言語の母語話者による音声のほうが非母語話者のそれよりも意図した感情を正しく聞き取られることを示唆している。

これらの研究を踏まえ、筆者らは 2 名の日本語母語話者 (以下 JN) および 4 名の中国語を母語とする日本語学習者 (CJ) の 7 つの感情 (「中立(Neutral)」「喜び(Happy)」「怒り(Angry)」「悲しみ(Sad)」「驚き(Surprised)」「恐れ(Afraid)」「嫌悪(Disgusted)」) を込めた発話の日本語音声データを収録し、それらを用いて発話者とは別の JN12 名、日本語能力試験 N1 合格者の CJ12 名、および日本語学習歴のない中国人非日本語学習者 (CN) 12 名の 3 群に聴取させ、音声を示す感情を選択してもらった (李他, 2018)。その結果、どの群でも CJ による音声の方が、発話者の意図した感情との同定率が高くなった ($p < .001$)。さらに、JN による感情音声では、「悲しみ」と「恐れ」、「喜び」と「驚き」、「怒り」と「嫌悪」

の3組の混同が顕著であり、CJによる感情音声では「悲しみ」と「恐れ」のみが混同されやすいという結果を得た。この結果は、先行研究の結果とは異なり、母語話者の発話よりも学習者の発話の意図する感情の同定率が高く、混同も少ないことが示したものである。本研究では、なぜこのような知覚傾向が見られたのかについてさらに検討するために、JNおよびCJが7種類の感情をこめて発話した音声を対象に音響分析を行ない、知覚との関係について考察を行う。

2. 音響分析

音響分析に用いた音声資料は、李他（2018）で収録したものをを用いた。音声に込める感情はEkman（1992）の6基本感情（「怒り」「悲しみ」「喜び」「中立」「驚き」「恐れ」「嫌悪」）に基づいて選び、「中立」は他の感情と比較するためのベースラインとして採用した。JN2名、CJ4名によるそれぞれの感情を込めた日本語音声計462トークン（6名×7感情×11タスク語）を分析の対象とした。全てのトークンについて、基本周波数F0の平均値（F0mean）、最小基本周波数（F0level（= min））、最大基本周波数（F0max）、基本周波数の範囲（F0span）、持続時間（Duration）、インテンシティ（Intensity）の6つのパラメータについて測定を行った。F0mean、F0level、F0max および F0span については、Hz から Z 得点へ変換した。JNによる音声の音響分析の結果を図1に示し、CJによる音声の音響分析の結果を図2に示す。図中のマーカ―はそれぞれの被験者の平均値を示す。

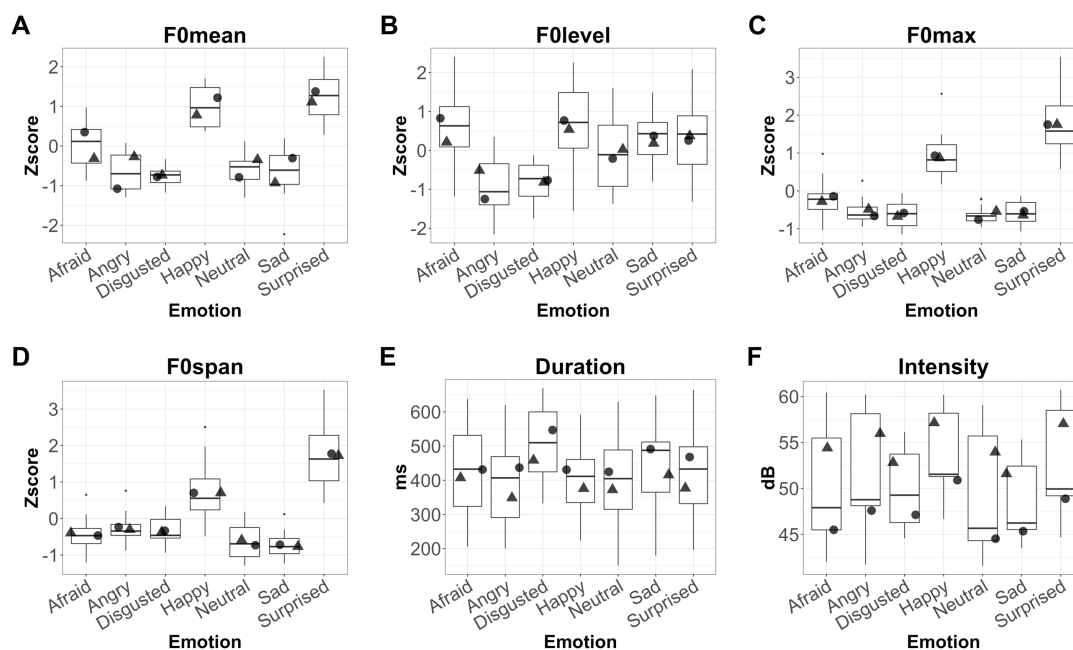


図1 JNによる感情音声の音響的特徴

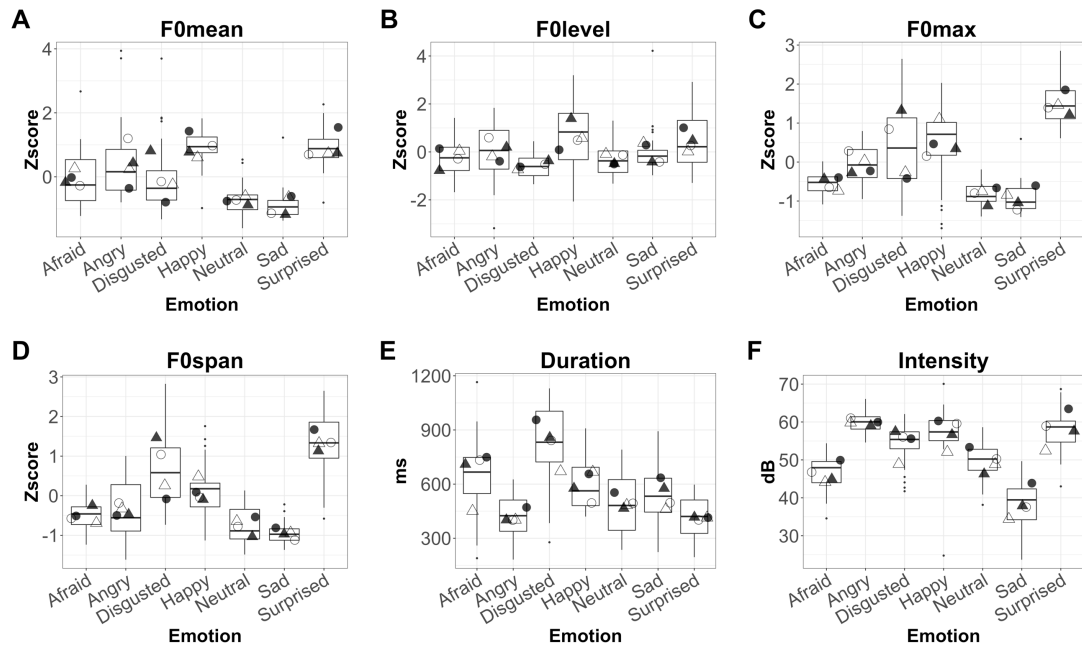


図 2 CJ による感情音声の音響的特徴

感情を要因とする 1 要因分散分析を行ったところ、すべての音響的特徴において感情の主効果が有意であった ($\text{all } ps < .001$)。そのため、それぞれの音響的特徴について多重比較を行い、有意差が見られた感情間の結果のみを表 1 に示す。

まず、李他 (2018) において、CJ による発話において混同されやすかった「悲しみ」と「恐れ」について検討する。CJ による発話の音響分析の結果では、「悲しみ」の F0mean と Intensity は「恐れ」より有意に低いことが分かった。この結果から、CJ が「悲しみ」と「恐れ」を区別するときは F0mean と Intensity を音響学の手がかりとして感情音声を生成している可能性を示唆していると考えられる。

さらに李他 (2018) においては、JN では「悲しみ」と「恐れ」の他に、「喜び」と「驚き」、「怒り」と「嫌悪」の 2 組で混同が顕著に認められたが、CJ ではそれはあまり見られなかった。表 1 の通り、CJ による発話では、「喜び」の F0max と F0span は「驚き」より有意に低く、「喜び」の Duration は「驚き」より有意に長かった。さらに、「怒り」の F0span は「嫌悪」より有意に狭かった。すなわち、CJ が感情音声を生成する際、「喜び」と「驚き」を区別する音響学の手がかりは F0max と F0span であり、「怒り」と「嫌悪」を区別する手がかりは F0span である可能性がある。一方で、JN が生成した感情音声では、これらの 3 組について音響的特徴の分析したところ、違いは見られなかった。これらの音響的特徴が、聴者が JN による発話に対して「悲しみ」と「恐れ」、「喜び」と「驚き」、「怒り」と「嫌悪」を混同する要因になっている可能性があると考えられる。以上の結果から、CJ の音声の同定率が高く、JN の音声混同されやすかったのは、JN の感情音声間の音響

的特徴の違いが CJ に比べて顕著ではなかったためだと考えられる。

表 1 各音響的特徴における多重比較の結果

CJ	F0mean	F0level	F0max	F0span	Duration	Intensity
中立 vs. 怒り	<i>n.s.</i>	<i>n.s.</i>	<i>n.s.</i>	<i>n.s.</i>	<i>n.s.</i>	<*
中立 vs. 悲しみ	<i>n.s.</i>	<i>n.s.</i>	<i>n.s.</i>	<i>n.s.</i>	<i>n.s.</i>	>*
中立 vs. 驚き	<i>n.s.</i>	<i>n.s.</i>	<i>n.s.</i>	<*	<i>n.s.</i>	<i>n.s.</i>
中立 vs. 嫌悪	<i>n.s.</i>	<i>n.s.</i>	<i>n.s.</i>	<i>n.s.</i>	<*	<i>n.s.</i>
喜び vs. 悲しみ	<i>n.s.</i>	<i>n.s.</i>	<i>n.s.</i>	<i>n.s.</i>	<i>n.s.</i>	>*
喜び vs. 驚き	<i>n.s.</i>	<i>n.s.</i>	<*	<*	>*	<i>n.s.</i>
怒り vs. 悲しみ	<i>n.s.</i>	<i>n.s.</i>	<i>n.s.</i>	<i>n.s.</i>	<i>n.s.</i>	>*
怒り vs. 驚き	<i>n.s.</i>	<i>n.s.</i>	<*	<i>n.s.</i>	<i>n.s.</i>	<*
怒り vs. 嫌悪	<i>n.s.</i>	<i>n.s.</i>	<i>n.s.</i>	<*	<i>n.s.</i>	<i>n.s.</i>
悲しみ vs. 驚き	<*	<i>n.s.</i>	<*	<*	<i>n.s.</i>	<*
悲しみ vs. 恐れ	<*	<i>n.s.</i>	<i>n.s.</i>	<i>n.s.</i>	<i>n.s.</i>	<*
悲しみ vs. 嫌悪	<i>n.s.</i>	<i>n.s.</i>	<i>n.s.</i>	<i>n.s.</i>	<i>n.s.</i>	<*

JN	F0mean	F0level	F0max	F0span	Duration	Intensity
嫌悪 vs. 驚き	<i>n.s.</i>	<i>n.s.</i>	<i>n.s.</i>	<*	<i>n.s.</i>	<i>n.s.</i>
悲しみ vs. 驚き	<i>n.s.</i>	<i>n.s.</i>	<i>n.s.</i>	<*	<i>n.s.</i>	<i>n.s.</i>

* $p < .05$

3. 生成と知覚の関係

感情音声の生成と知覚の関係を調べるために、収録した発話に現れた音響的特徴と李他(2018)で得られた感情同定率の相関分析を行った結果を表 2 に示す。前節で述べた感情の区別に手がかりとされた可能性のある音響的特徴に関して網かけを用いて示す。

CJ による発話では、「悲しみ」と「恐れ」に関して、「悲しみ」の同定率は F0mean が低いと高くなり ($r(132) = -.47, p < .001$), Intensity が弱いと高くなる ($r(132) = -.38, p < .01$)。

「恐れ」の同定率は F0mean が高いと高くなり ($r(132) = .57, p < .01$), Intensity が強いと高くなる ($r(132) = .23, p < .01$)。また、「喜び」と「驚き」に関して、「喜び」は Duration が長いと同定率が高くなる ($r(132) = .62, p < .001$)。さらに、「怒り」と「嫌悪」に関して、「怒り」は F0span が狭いほど同定率が高い ($r(132) = -.32, p < .05$) ことが観察された。

JN による発話では、感情の区別に手がかりとされた可能性のある音響特徴と同定率の相関は見られなかった。

表 2 音響的特徴と同定率の相関 (r)

		各感情の同定率						
CJ	恐れ	怒り	嫌悪	喜び	中立	悲しみ	驚き	
F0mean	.57***	.60***	-.16	.16	-.03	-.47***	-.27*	
F0level	.31***	-.09	-.18*	.14	.18*	-.53***	.22*	
F0max	.24**	.26**	-.10	.19*	-.11	.50***	-.02	
F0span	-.09	-.23**	-.07	.13	-.21*	.04	-.14	
Duration	-.08	.20	.27***	.62***	-.19*	.07	-.13	
Intensity	.23**	.27*	.09	.06	.13	-.38**	-.03	
JN	恐れ	怒り	嫌悪	喜び	中立	悲しみ	驚き	
F0mean	.13	.01	.12	-.06	.07	-.02	.27	
F0level	.48	.26*	.03	-.16	-.01	-.43***	.13	
F0max	.16	.16	.08	-.15	.22	.04	.12	
F0span	-.26*	-.05	.06	-.03	.10	.05	.05	
Duration	-.45***	.09	-.17	.04	-.13	.17	-.07	
Intensity	.05	.15	.22	.35	.41**	-.01	-.13	

* $p < .05$, ** $p < .01$, *** $p < .001$

以上の感情音声の生成と知覚の関連についてまとめる。CJ による発話では、「悲しみ」と「恐れ」に関して、「恐れ」の F0mean と Intensity は「悲しみ」より有意に大きく、「悲しみ」の同定率は F0mean が低いと高く、Intensity が弱いと高く、「恐れ」の同定率は F0mean が高いと高く、Intensity が強いと高いことが見られた。また、「喜び」と「驚き」に関して、「驚き」の F0span は「喜び」より有意に大きく、「喜び」の Duration は「驚き」より有意に長く、「喜び」は Duration が長いと同定率が高くなることが分かった。さらに、「怒り」と「嫌悪」に関して、「嫌悪」の F0span は「怒り」より広く、「怒り」は F0span が狭いほど同定率が高いことが観察された。以上の分析から、聴者は CJ の「悲しみ」と「恐れ」の区別に F0mean と Intensity、「喜び」と「驚き」に Duration、「怒り」と「嫌悪」に F0span を手掛かりとしている可能性が示された。

4. おわりに

筆者らの前研究(李他, 2018)において, CJによる発話の同定率が高く, JNによる発話が混同されやすいことが見られたが, 音響分析および同定率との相関分析の結果, JNの感情音声間の音響的特徴の違いがCJに比べて顕著ではなかったためだということが明らかになった。CJの発話では「恐れ」と「悲しみ」の区別には声の高さ(F0mean)と強さ(Intensity)を制御して表しており, 「驚き」と「喜び」ではピッチレンジ(F0span)や持続時間(Duration)を, 「嫌悪」と「怒り」でもピッチレンジ(F0span)を制御していることが示された。このような音響特徴の顕著な違いはJNには見られず, CJは日本語の感情音声の伝達にあたって, 日本語母語話者よりも多くの音響情報を利用していることが示された。このようなCJの感情音声の音響特徴が背景言語の影響によるものなのか, または学習言語であるためのオーバーリアクションと考えられるのかについては, 今後さらに, 中国語の感情音声について同様の分析を行うなどして検討する必要がある。また, 日本語における感情表出についても, 声質など本研究で行なった分析以外の面からの検討も行っていく予定である。

謝辞

本研究は, JSPS 科研費 17H02352 の助成を受けたものです。論文執筆にあたり貴重なご意見をいただいた川島朋也氏にこの場を借りて感謝いたします。

参考文献

- Ekman, P. (1992). An Argument for Basic Emotions. *Cognition And Emotion*, 6, 169–200.
- Erickson, D. (2005). Expressive Speech: Production, perception and application to speech synthesis. *Acoust. Sci. & Tech*, 26, 317–325.
- 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 (pp. 31–46) くろしお出版.
- 甲斐 朋子・田淵 咲子 (2003). 日本語の感情を含む発話に対する韓国人日本語学習者の聞き取りと発話をめぐって ポリグロシア, 7, 1–11.
- 李 歆玥・Aaron Albin・林 良子 (2018). 中国語を母語とする学習者による日本語感情音声の知覚—母語話者および学習者による発話を対象に— 音声学会第 337 回例会要旨.
- 櫻庭 京子・今泉 敏・笥 一彦 (2001). 感情判断と言語判断—母語は外国語より感情知がしやすいか— 信学技報, 20, 33–38.
- 櫻庭 京子・今泉 敏・笥 一彦 (2004). 「ピカチュウ」にこめられた感性情報 音声研究, 8, 77–84.
- Scherer, K. R., Banse, R., & Wallbott, H. (2001). Emotion Inferences From Vocal Expression Correlate Across Languages and Cultures. *Journal of Cross-Cultural Psychology*, 32, 76–92.

ベトナム人日本語学習者の日本語発話リズム —日本語母語話者との比較—

松田真希子（金沢大学）・吉田夏也（国立国語研究所）・金村久美（名古屋経済大学）
 mts@staff.kanazawa-u.ac.jp, natsuya@ninjal.ac.jp,
 kanamura@nagoya-ku.ac.jp

1. はじめに

第二言語習得において、発話のリズムは韻律特性の中でも母語話者の発話印象に影響を与える重要な要素である (Tajima et al., 1997, Gut, 2007)。モーラをリズムの基本的な単位とする日本語は、母音の長短の区別、促音、撥音、拗音、母音の無声化など、さまざまな言語を母語とする日本語学習者にとって、習得が困難な韻律特徴を多く有するといわれている (近藤, 2012)。

ベトナム語を母語とする日本語学習者 (以下、ベトナム人学習者とする) の発音では、中国など他の声調言語母語話者と比べても、リズムの不自然さがより強いことが指摘されている (松田, 2016)。ベトナム人学習者は、音節リズムであるベトナム語のリズムにあてはめて日本語を発音しようとする傾向があり (金村, 1999)、その結果、日本語のモーラリズムが乱れ、短音の長音化や長音の短音化、促音の挿入などが起こるとみられる。杉本 (2005) が行なった、ベトナム人学習者の日本語の語の読み上げにおける調査においても、長音化と短音化等に関する誤用が多く出現していると報告されている。しかし、ベトナム人学習者のリズム習得に関する研究は、語レベルでの研究 (杉本, 前掲)、知覚に関する研究 (Đỗ Hoàng Ngân, 2015) を除き、非常に限られている。

近年、言語のリズムの特徴を計測する物理的指標として、Pairwise Variability Index (PVI) が提唱されている (Grabe & Low, 2002)。PVI 値は、発話における母音および子音の時間長のばらつきを数値化したものであり、PVI 値が小さいほど、各セグメントの値のばらつきが少ないことを意味するため、これによって言語のリズムを客観的に評価することが可能となる。PVI によって学習言語のリズムを評価する可能性が示唆されており (近藤, 前掲)、日本人英語学習者の PVI を調べた研究は、新谷 (2017) などすでに数多くある。一方日本語学習者のリズムを PVI によって調べた研究は、韓国語母語話者のもの (木下, 2011)、台湾人対象のもの (呂, 2017) 等まだ限られている。同様にベトナム語話者についても、ベトナム人英語話者のリズムを調べた研究はあるが (Słowik & Volin, 2018)、日本語学習者の日本語発話を調べた研究は管見の限りない。そこで、本研究では、日本語母語話者とベトナム人学習者に同一のテキストを朗読させた音声を収集し、両音声の PVI を比較分析し、ベトナム人学習者のリズム面での発音の不自然さについて考察を試みる。

2. 資料

2.1. 分析対象

駅から大学へのアクセス方法の説明文 (全 14 文) を用意し、音声提供者に読み上げるよ

う依頼し、録音した。録音した音声の中から、本報告では、表1に示した5つの文の音声を分析対象とした。この5文を選んだのは、特殊拍を含む文を分析対象とするためである。明らかな言い間違いは音声から削除し、正しく言い直した部分は分析対象に含めた。

表 1: 読み上げ文 ([数字]は文番号)

[3]	駅に着いたら、駅の東口に行ってください
[4]	東口の前に、広島大学行のバス乗り場があります
[5]	90番, 92番, 95番です
[6]	ふつう、バス乗り場の前に、学生の行列ができていますので、すぐ、わかるはずです
[11]	整理券の番号の料金を払ってください

2.2. 話者及び録音方法

分析に使用したのは、日本語母語話者3名（関東地方出身者・女性・30代）、ベトナム人日本語学習者3名（北部出身者・女性・20代、2名、中部出身者・男性・20代、1名）の音声である。ベトナム人学習者の日本語能力については、全員上級レベルである。3名ともベトナムの大学で日本語を学び、現在は日本の大学に留学中である。ベトナム人学習者の音声提供者3名の属性を表2に示した。

表 2: 被験者の属性

名前 (仮名)	性別	JLPT	滞日期間	専攻	身分
マイ	女性	N2	9か月	日本語・日本語教育	学部交換留学生
トゥイ	女性	N1	4か月	人文学	大学院研究生
ズン	男性	N1	10か月	人文学	学部交換留学生

録音については、静音環境で読み上げを行い、MP3形式（ビットレート192kbps）、またはwav形式（サンプリング周波数：44.1kHz）で録音した。録音は次の手順で行なった。まず、録音前に読み上げ文を紙媒体で配布し、事前に練習してもらい、その際には録音担当者は誤り等の指摘は行わなかった。次に、文中に未知語がないことを確認した後で、音声提供者に紙媒体を見ながら読み上げるよう依頼した。言い間違い、ポーズを除いた各話者の音声の持続時間は、5つの文で13.8秒から20.4秒であった。分析対象とした音声の数は、両言語とも3名が各5文を読み上げたため、15文ずつである。また、計測した母音部（Vocalic）の数は、ベトナム人学習者298個、日本語母語話者278個である。

3. 分析

3.1. 方法

分析にはPraat (ver.6.0.40)を使用し、録音した音声を、波形とスペクトログラムの視認

によって、母音部と母音間部に分けた。分類基準は基本的に Grabe & Low(前掲)に従っており、母音部には、母音が含まれ、母音間部には子音と無声化母音が含まれる。なお、今回分析対象とした音声の中の拗音は、母音との分離が困難であったため、母音部に含めている。また、発話頭の閉鎖音の閉鎖部分は、閉鎖の開始部が不明のために、計測に含めていない。なお、Grabe & Low(前掲)では、いわゆる final lengthening の影響を避けるために、音韻句末のセグメントを分析対象から除外しているが、本報告では日本語学習者の final lengthening の有無が不明であるために、音韻句末のセグメントも分析対象にしている。

このセグメンテーションにしたがって、母音部と母音間部の時間長及び強度（該当区間の二乗平均平方根値）を計測した。下記では、まず、母音部の持続時間の PVI、及び強度の PVI の分布を、母語別に比較する。次に、母音部と母音間部の全セグメントの持続時間の平均値を、母語別に比較する。なお、今回、強度を分析対象とするのは、ベトナム語は音声において句末で常に伸長が起こる (Final lengthening) 言語であり (Thompson, 1965)、持続時間だけでなく強度の現れ方も日本語と異なることが推測される。このことがベトナム人学習者の日本語音声のリズムに影響する可能性を探るためである。

3.2. 結果（持続時間 PVI の分布）

図 1 に、母音部の持続時間の PVI の分布を示す。この分布から、日本語母語話者と比較して、ベトナム人学習者では、PVI 値の分布が 0-50 の間に偏っており、日本語母語話者と比べて PVI が小さいセグメントが多いことがわかる。両者の PVI の平均値は、ベトナム人学習者が 44.8、日本語母語話者が 52.0 であった (t 検定: $P=0.0096<.01$)。

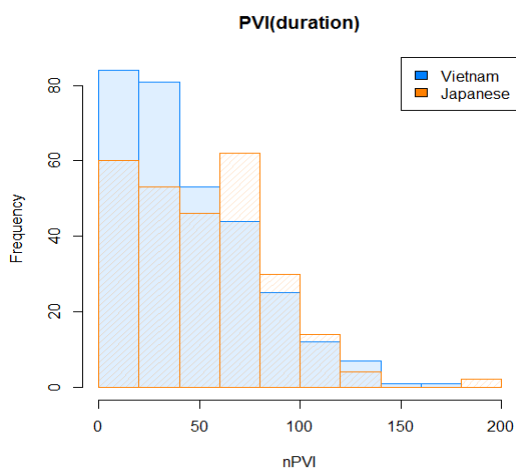


図 1: 持続時間 PVI の分布

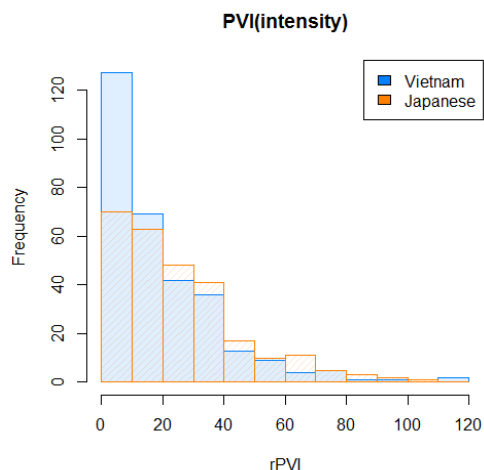


図 2: 強度 PVI の分布

3.3. 結果（強度 PVI の分布）

図 2 に母音部の強度の PVI の分布を示す。強度の分析では、発話速度を考慮しないため、rPVI を使用している。また、数値を見やすくするために、Grabe & Low (前掲) にならって、PVI 値を 1000 倍にして表示している。

ベトナム人学習者のグループは、日本語母語話者に比べ、母音部の持続時間の PVI と同様に、母音部の強度の PVI も 0-20 の間に偏って分布している。このことから、ベトナム人学習者では、母音部の強度の PVI が小さいセグメントが多いといえる。両者の PVI 平均値は、ベトナム人学習者は 19.6、日本語母語話者は 25.8 であった (t 検定: $p=0.0002<.001$)。すなわち、ベトナム人学習者の音声では、隣り合うセグメント間の強度の変化が、日本語母語話者と比べて小さい音声であるといえる。

3.4. 結果（母音部と母音間部の持続時間長の平均値の比較）

続いて、母音部と母音間部の持続時間の平均値を比較する。表 3 に、母音部と母音間部の持続時間の平均値を、日本語母語話者とベトナム人学習者を比較して示した。母音部の持続時間の平均値は、ベトナム人学習者は日本語母語話者より長く、有意差が検出された ($P=3.53 \times 10^{-7}<.0001$)。一方で、母音間部の持続時間の平均値には、母語による有意差が見られなかった。ベトナム人学習者の発話時間は、平均 3.79 秒であったのに対して、日本語母語話者は、平均 3.09 秒と短かったことと合わせて考えると、ベトナム人学習者が遅い発話速度で発話したときに母音部をより長く発音していることの結果と考えられる。

表 3：母音部と母音間部の持続時間長の平均値の比較

	持続時間(単位：msec)		
	ベトナム人学習者	日本語母語話者	
母音部	100.1(n=298)	81.2(n=278)	<.0001
母音間部	81.1(n=307)	75.5(n=277)	n.s.

4. 考察

本研究のデータにおいては、ベトナム人学習者の母音部の持続時間は日本語母語話者より有意に長いことが確認されたが、母音の持続時間の PVI 値 (44.8) は日本語母語話者 (52.0) よりも有意に小さいことが確認された。このことは、ベトナム人学習者の母音部の持続時間のばらつきが小さいことを意味する。

また、今回録音した音声を聴覚的に観察してみると、ベトナム人学習者には「となりー」「きゅじゅにーばん」「きゅじゅごーばん」のような長音化が多くみられた。こうした短母音が連続する部分において長音化する例がある一方、「乗車 (じょうしゃ)」が「じょしゃ」となるように長音が短音化したり、「整理券 (せいりけん)」が「せりけん」となるように、母音連続がある部分での母音が脱落したりする現象も顕著であった。

これはベトナム語の母語干渉によるものと考えられる。まず、Slówik & Volín(前掲)の調査によるベトナム語の母音部の持続時間と比較したい。この値は約 20-40 と、今回の日本語母語話者(52.0)、ベトナム人学習者(44.6)よりも低い。

次に対照言語学的に見ると、日本語には母音の長短の区別があり、日本語母語話者の音声ではこれを区別して発音できるのに対し、ベトナム語では、長音節・短音節の区別はなく、母音の持続時間を伸張して音節全体の持続時間を長くしてリズムに変化をつけることは原則としてない。音節の形態が CVC, CVVC, CV, CVV と多様であっても、音節中の母音の長さが伸び縮みすることによって、音節頭の開始時間の間隔、すなわち音節のリズムが一定に保たれる。そのため、学習言語である日本語においても、母音の長さの長短の差をつけ、CV または V から成る拍の長さを一定に保って発音することができず、代わりに長い母音を短く、短い母音を長く、と母音の長さを伸び縮みさせて、拍よりも音節のリズムを一定に保つことを優先してしまうのである。

今回の結果は、ベトナム人学習者が母音の長短の区別をつけられず、どちらも同程度の長さで発音したために、母音の長さの区別が日本語母語話者に比べて乏しいものになったことによると考えられる。

また、母音部の強度の PVI においても、日本語母語話者とベトナム語話者には差があり、ベトナム語話者のほうが PVI 値が低いことが確認された。このことは、日本語母語話者に比べて、ベトナム人学習者は、母音強度が一定になる傾向が強いことを示している。現状ではこの結果の原因を十分に考察するための情報がないが、日本語話者は句中のピッチのピークにおいて強度も強まり、句末に近づくにつれて弱くなるのに対し、ベトナム語は句末を除きいずれかの音節を強く発音することはなく、音節の強度が一定に保たれて発音される傾向にある。こうしたことから、日本語のように句のピッチに合わせて強度を推移させることが難しく、常に一定の強さで話す発音となっている可能性がある。

5. まとめと今後の課題

本研究ではベトナム人学習者 3 名と日本人 3 名の読み上げ音声のリズムについて分析を行った。その結果、母音の長さや強度の PVI に有意な差があり、いずれも日本語母語話者の方がベトナム人学習者より値が大きいことが確認された。日本語教育の現場では、ベトナム人学習者は拍の長音化や短音化の誤りが多いとよく指摘されるが、それは長短を区別して認識・発音することが困難なため、結果として短音を長めに、長音を短めに発音してしまい、全体として長短の区別が小さい発音にしまっていることを示唆している。

また、強度に関しては、今後より詳しい調査が必要である。ベトナム人学習者の日本語の発音は、聴覚的印象では、音節の開始部が強く、末尾で弱くなるように聞こえる。これに対し日本語母語話者の発音は、拍の持続時間中では強さにあまり違いがなく、句の高さのピークが強く、句末は弱くなる傾向にある。今後、リズム単位中の強度の変動を調べることで、ベトナム人学習者の日本語発話の特徴がより詳しく捉えられるとみられる。

なお、今回はサンプル数の小さいデータに基づく検証であり、標本数が十分であったと

はいえない。日本語の習得レベルの高いベトナム人学習者を対象にリズムの特徴を調べたが、初級や中級レベルの学習者の音声は分析対象としていない。また、他の母語話者との比較、被験者のベトナム語の朗読音声との比較なども重要であろう。こうした分析については今後の課題とする。

付記 本研究は JSPS 科研費 17H02352 の助成を受けたものです。

参考文献

- 新谷敬人 (2017) 「留学は発音を良くするか? : リズム特徴に基づく予備的研究」『The Bulletin of the Otsuma English Association』 50, 113-125.
- 金村久美 (1999) 「ベトナム語母語話者による日本語の音調上の特徴」『ことばの科学』 第 12 号, 73-91.
- 木下直子 (2011) 『日本語のリズム習得と教育』 早稲田大学出版部
- 近藤眞里子 (2012) 「日本語学習者の音声習得における第一言語特有の干渉と普遍言語的干渉—日本語教師へのアンケート調査から—」『早稲田大学文学研究科紀要第 3 分冊』 57, 21-34.
- 杉本妙子 (2005) 「ベトナム語圏日本語学習者の発音に関わる誤用について II: 音声聞き取り調査と発音調査における長音化・短音化の誤用の比較と考察」『茨城大学人文学部紀要. コミュニケーション学科論集』 (17), 73-93.
- 松田真希子 (2016) 『ベトナム語母語話者のための日本語教育』 春風社
- 呂思盈 (2017) 「台湾人日本語学習者の日本語リズム特性—自立拍・促音・拗音について—」 広島大学大学院総合科学研究科博士論文
- Boersma, Paul & David Weenink. (2018) "Praat: doing phonetics by computer."
- Đỗ Hoàng Ngân. (2015) 「ベトナム人学習者の日本語における長音・促音の知覚に関する問題」 *Tạp chí Khoa học ĐHQGHN: Nghiên cứu Nước ngoài*, Tập 31: Số 2, 31-38. (『ハノイ国家大学科学雑誌-外国研究-』 第 31 卷 2 号, 31-38.)
- Grabe, Esther & Ee Ling Low. (2002) "Durational variability in speech and the rhythm class hypothesis." In Carlos Gussenhoven and Natasha Warner (eds.) *Laboratory Phonology 7*. De Gruyter Mouton, 515-546.
- Gut, U. (2007) Foreign accent. In: Müller, C. (ed.), *Speaker classification*. Berlin: Springer, 75-87.
- Slówik, O., Volín, J. (2018) Acoustic correlates of temporal structure in North-Vietnamese English. In: Volín, J., Skarnitzl, R. (ed.), *The Pronunciation of English by Speakers of Other Languages*. Cambridge Scholars Publishing, 196-209.
- Tajima, K., Port, R. and Dalby, J. (1997) "Effects of temporal correction on intelligibility of foreign-accented English." *Journal of Phonetics* 25, 1-24.
- Thompson, L.C. (1965) *A Vietnamese reference grammar*. Seattle: University of Washington Press.

長音と促音の知覚における隣接要素間の同化効果と対比効果*

石橋 頌仁・神谷 祥之介・竹安 大（福岡大学人文学部）

1. はじめに

日本語には母音と子音に音韻的長短の対立が存在し、当該母音または子音の持続時間を主要な手がかりとして音韻的長短の知覚がなされる（藤崎・杉藤 1977）。また、母音や子音の音韻的長短の知覚には、隣接する音節の構造やその構成要素の持続時間、母音の F0、発話速度など、様々な二次的要因が影響することも知られている（関連する先行研究については、Kawahara 2015, Takeyasu & Giriko 2017 等を参照）。本稿では、日本語における母音と子音の音韻長の知覚に関わる二次的要因のうち、隣接する要素間に生じる同化効果と対比効果について論ずる。具体的には、 $C_1V_1C_2V_2$ 構造の 2 音節語において、 V_1 の持続時間と音韻長が C_2 の音韻長の知覚に与える影響と、 C_2 の持続時間と音韻長が V_1 の音韻長の知覚に与える影響について議論する。

1.1. 隣接要素間の同化効果と対比効果：先行研究

本稿では、ある要素が物理的または音韻的に長くなることにより、隣接する要素が音韻的に長いと判断されやすくなることを同化効果と呼び、ある要素が物理的または音韻的に長くなることにより、隣接する要素が音韻的に長いと判断されにくくなることを対比効果と呼ぶこととする。

促音の知覚においては、子音に先行する母音の持続時間が長くなると、その子音は促音だと判断されやすくなる（＝同化効果）ことが知られている（大深他 2005）¹。先行母音が長ければ長いほど同化効果が生じると仮定すれば、子音に先行する母音が長母音のように音韻的にも物理的にも長い場合、後続する子音には促音の知覚が生じやすくなるはずである。しかし、実際には、子音に先行する母音が長母音である（音韻的に長い）と、短母音である（音韻的に短い）場合と比べてその子音は促音だと判断されにくくなる（＝対比効果）ことが指摘されており（詳細は Takeyasu & Giriko 2017 を参照）、話は単純ではない。

Takeyasu & Giriko (2017) では、 $C_1V_1C_2V_2$ 構造の 2 音節語の V_1 と C_2 の持続時間が様々な長さを持つように操作した刺激を用い、被験者に V_1 と C_2 の音韻長に関する 4 択の判断 (CVCV, CVVCV, CVCCV, CVVCCV) を求め、先行母音 (V_1) の持続時間および被験者によって知覚された先行母音の音韻長が後続子音 (C_2) の促音知覚にどのように影響するかを調べる実験を行った。その結果、先行母音の持続時間と、被験者が知覚した先行母音の音韻長の影響は互いに独立して存在しており、前者は同化効果、後者は対比効果をもたらすことが明らかとなった。さらに、Takeyasu & Giriko (2017) の知覚実験では、母音 (V_1) の音韻長の知覚に

* 本研究は福岡大学音声学実験室の研究プロジェクト成果の一部である。

¹ 先行母音持続時間の延長により、促音だと判断されにくくなるという報告もある。関連する先行研究やこうした齟齬が生じる理由については、Takeyasu & Giriko (2017) の議論を参照。

対して、後続子音 (C₂) の持続時間や知覚された音韻長が影響を及ぼし、前者が同化効果、後者が対比効果をもたらすことも示された。しかし、興味深いことに、後続子音の持続時間による同化効果は、後続子音が促音だと判断された場合には生じなかった²。

1.2. Takeyasu & Giriko (2017)の問題点と本研究の目的

Takeyasu & Giriko (2017)の知覚実験は、従来はあまり明確に区別されていなかった持続時間の影響と音韻長の影響を分離することに成功した一例である。しかし、この実験結果が一般性の高いものであると言えるかどうかは注意深く考える必要がある。まず、Takeyasu & Giriko (2017)の知覚実験では、東京方言話者のみが被験者とされており、また、ターゲットとなる語の F₀ を平坦に設定した刺激が用いられている。方言によって、特殊拍の持続時間のコントロールが異なる場合があることが指摘されているほか (前川 1997)、東京方言をはじめとする日本語諸方言では、語アクセントによって語内の F₀ が変動するのが一般的であることを考えると、東京方言以外の話者を対象とした場合にも、また、語の F₀ に変動がある刺激を用いた場合にも、同様の実験結果が得られるとは限らない。

そこで、本研究では、福岡方言話者を対象とし、語の F₀ に変動がある刺激を用いた場合にも、Takeyasu & Giriko (2017)が指摘している隣接要素間の同化効果と対比効果が観察されるかどうかを知覚実験を通して明らかにし、Takeyasu & Giriko (2017)の指摘の一般性を検証する。

2. 知覚実験

2.1. 刺激

日本語を母語とする福岡出身の女性に、2 音節の無意味語「パーポ」 (/paRpo/、アクセント型は平板)を、「彼は__と言った。」というキャリア文に入れた状態で、話者にとって普通の発話速度で 10 回ずつ発音してもらった。この中から、言いよどみや雑音の混入がなく、無意味語の各セグメント持続時間がそれぞれの平均値にできるだけ近いトークンを 1 つ選出した。選出したトークンにおける無意味語の各セグメント持続時間は、表 1 の通りであった。

次に、無意味語の/aR/の持続時間およびそれに後続する/p/の閉鎖持続時間と、単語全体の F₀ を以下のように操作することにより、「パポ」、「パーポ」、「パッポ」、「パーッポ」のいずれかに聞こえるような刺激音声を作成した。/aR/については、隣接するセグメントとの遷移区間を除いた定常部 (母音開始点から 40 ms ほど経過した時点から 140 ms ほど) を選択し、praat (Boersma and Weenink 2017)の Manipulation 機能を用いてその区間の持続時間を短くすることにより、/a/全体の持続時間を 21 ms 刻みで 186 ms から 60 ms まで 7 段階に設定した。また、/p/の無音区間については、21 ms の無音区間を必要なだけ埋め込むまたは削除することにより、55 ms から 181 ms まで 7 段階に設定した。

² Takeyasu & Giriko (2017)によれば、これは音声産出における母音・子音持続時間の分布と音声知覚が対応しているためだと解釈できる。

さらに、praat の Manipulation 機能を用いて無意味語の F0 を操作し、F0 に関して平坦、上昇、下降の 3 つの系列を作成した。平坦系列では、V₁ の開始時点から V₂ の終了時点まで、有声区間の F0 は 240 Hz に保たれた。上昇系列では、V₁ 開始時点から V₁ 終了時点にかけて F0 が 190 Hz から 240 Hz に上昇し、その後 V₂ 終了時点まで 240 Hz に保たれた。下降系列では、V₁ 開始時点から 40 ms が経過する時点までは F0 が 240 Hz に保たれ、そこから V₂ の終了時点にかけて 240 Hz から 170 Hz まで下降した。なお、録音時点では無意味語が平板型で発音されているため、無意味語に後続する「と言った」は元の音声では F0 が高いまま続いているが、これを下降系列にそのまま接続すると不自然な音声になってしまう。また、キャリア文の F0 が刺激系列によって異なると、平坦、上昇、下降系列の間で異なる実験結果が得られた場合、それがキャリア文の F0 の違いによって生じたという可能性を排除できない。そこで、無意味語に後続するキャリア文の「と言った」の有声区間の開始点から文の終了まで、F0 が 160 Hz から 140 Hz に下降するように設定し、これを平坦、上昇、下降すべての系列で用いることとした。これによって、平坦、上昇系列は尾高型に聞こえることとなった。

以上の操作により、V₁ 持続時間 (7 段階) × C₂ 閉鎖持続時間 (7 段階) × F0 パターン (3 系列) の掛け合わせにより、147 種類の刺激音声を作成した。

表 1: 無意味語の持続時間と刺激の持続時間の設定

	C ₁ (/p/)	V ₁ (/aR/)	C ₂ (/p/)	V ₂ (/o/)
選出したトークン	閉鎖区間: 99 ms VOT: 51 ms	198 ms	閉鎖区間: 76 ms VOT: 16 ms	66 ms
刺激の設定		60 ms ~ 186 ms (21 ms 刻み、7 段階)	55 ms ~ 181 ms (21 ms 刻み、7 段階)	

2.2. 被験者

15 名の福岡在住の日本語母語話者 (20~28 歳) が実験に参加した。被験者は、キャリア文に埋め込まれた状態の無意味語を聞き、それが「パポ」、「パーポ」、「パツポ」、「パーッポ」のいずれに聞こえるかを回答した。刺激は被験者ごとにランダムな順序で計 5 回ずつ提示された。

2.3. 結果と考察

15 名の被験者のうち、実験後の質問紙調査により、高校卒業までに福岡県外で生活した経験があることが判明した 2 名を除く 13 名を分析の対象とした。被験者の回答は、V₁ が音韻的に長い (=長母音である) と判断されたかどうか、また、C₂ が音韻的に長い (=促音である) と判断されたかどうかという点から集計・分析された。以下では、C₂ の音韻長の知覚に対する V₁ の影響、V₁ の音韻長の知覚に対する C₂ の影響の順に結果を提示する。

2.3.1. C₂ の音韻長の知覚に対する V₁ の影響

各 V₁ 持続時間ごとの C₂ 促音判断境界値を probit 分析により求めた結果を図 1(a)に示す。さらに、V₁ が長母音だと判断されたかどうかによってデータをさらに分割した結果が図 1(b)

である。図 1(b)より、V₁ が短母音だと判断された場合も長母音だと判断された場合も、V₁ 持続時間が長くなるほど促音判断境界値が下がっていく様子が見て取れる。また、V₁ 持続時間が同じであれば、V₁ が長母音だと判断された場合の方が促音判断境界値が高くなることがわかる。これらは、Takeyasu & Giriko (2017)で報告されているのと同様の傾向であると言える。一方、F0 系列間の差については、全体として平坦系列の促音判断境界値が高めで、上昇系列の促音判断境界値が低めであるように見えるが、V₁ の音韻長によっても F0 系列の影響の現れ方が異なっているようであり、一般化することが難しい。

被験者の C₂ 音韻長の判断を 2 値の従属変数（非促音(0); 促音(1)）、C₂ 持続時間（連続変数）、V₁ 持続時間（連続変数）、V₁ 音韻長（名義変数: 短母音(0); 長母音(1)）、F0 の系列（名義変数: 平坦(0)、上昇(1)、下降(2)）を独立変数とするロジスティック回帰分析の結果、V₁ の音韻長×F0 の系列の交互作用が強いことが明らかとなったため、F0 系列ごとにデータを分割して再度ロジスティック回帰分析を実施した。その結果、いずれの F0 系列においても V₁ 持続時間が長いほど促音判断率が上がる（促音判断境界値が下がる）こと、また、V₁ が長母音だと判断されると促音判断率が下がる（促音判断境界値が上がる）ことが明らかとなった（いずれの要因も 0.1% 水準で有意。紙面の都合上、偏回帰係数や Wald χ^2 の値は省略する。）。つまり、V₁ 持続時間の影響は同化効果的、V₁ の音韻長は対比効果的であり、本研究の実験においても、Takeyasu & Giriko (2017)で報告されているのと同様の結果が得られたことになる。

また、F0 系列間の結果を比較するため、V₁ が短母音だと判断されたものと、長母音だと判断されたものにデータを分割し、F0 系列を独立変数に含めたロジスティック回帰分析を行い、オッズ比に基づいて 3 系列間の多重比較を行ったところ、V₁ が短母音だと判断された場合には平坦系列と下降系列の間に有意な差が見られ ($p < 0.001$)、下降系列の方が促音だと判断されやすかった。それに対して、V₁ が長母音だと判断された場合には、平坦系列と上昇系列の間、および下降系列と上昇系列の間に有意な差が見られ(いずれも $p < 0.001$)、上昇系列は他の系列よりも促音だと判断されやすいことが明らかとなった。F0 に下降があるとそうでない場合と比べて促音の判断が促進される（V₁ が短母音だと判断された場合）という点では、Kubozono et al. (2013)で報告されているのと同様の傾向が観察されたと言えるが、V₁ が長母音だと判断されると F0 変動の影響の現れ方が変わってしまう理由については、現時点では明確な回答を出すことは難しいため、今後の検討課題としたい。

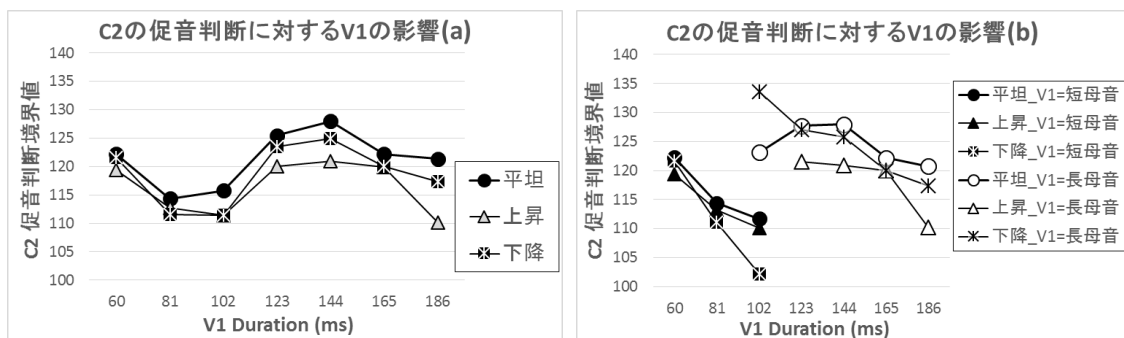


図 1: C₂ の促音判断に対する V₁ の影響---V₁ 音韻長によりデータを分けない場合(a)と分けた場合(b)

2.3.2. V₁の音韻長の知覚に対するC₂の影響

各 C₂ 持続時間ごとの V₁ 長音判断境界値を probit 分析により求めた結果を図 2(a)に示す。さらに、C₂ が促音だと判断されたかどうかによってデータをさらに分割した結果が図 2(b)である。図 2(b)より、C₂ が非促音だと判断された場合には、C₂ 持続時間が長くなるほど長音判断境界値が下がっていく様子が確認できるが、C₂ が促音だと判断された場合にははっきりとした傾向が確認できない。C₂ が非促音と判断されるか促音と判断されるかによって、V₁ の音韻長に対する C₂ 持続時間の影響の現れ方が異なることは、Takeyasu & Giriko (2017) で報告されているのと同じ傾向である。また、C₂ 持続時間が同じであれば、C₂ が促音だと判断された場合の方が長音判断境界値が高くなっており、この点においても、本研究の実験結果は Takeyasu & Giriko (2017)の実験結果とよく似た傾向を示している。一方で、F0 系列間の差については、全体に C₂ の音韻長に対する V₁ の影響で見られたのとは異なる結果となっているだけでなく、C₂ の音韻長によっても F0 系列の影響の現れ方が異なっており、やはり単純な一般化をすることが難しい。

被験者の V₁ の音韻長の判断を 2 値の従属変数（短母音(0); 長母音(1)）、V₁ 持続時間（連続変数）、C₂ 持続時間（連続変数）、C₂ 音韻長（名義変数: 非促音(0); 促音(1)）、F0 の系列（名義変数: 平坦(0)、上昇(1)、下降(2)）を独立変数とするロジスティック回帰分析の結果、複数の交互作用が強く働いていることが明らかとなったため、F0 系列ごと、また、C₂ 音韻長ごとにデータを分割して再度ロジスティック回帰分析を実施した。その結果、いずれの F0 系列においても C₂ が促音だと判断されると長音判断率が下がる（対比効果が生じる）傾向が見られた（上昇系列の C₂ 音韻長は 10%水準、それ以外はすべて 0.1%水準で有意）。一方、C₂ 持続時間については、C₂ が非促音だと判断された場合にはどの系列においても同化効果（C₂ 持続時間が長いほど長音判断率が上がる）を示したのに対し（いずれの系列でも $p < 0.001$ ）、C₂ が促音だと判断されると、系列によって C₂ 持続時間の効果が有意でなかったり、同化効果ではなく対比効果を示したりするなど、一貫した傾向が見られなかった。C₂ の音韻長が対比効果を示すという点と、C₂ が促音だと判断されると V₁ の音韻長判断に対して C₂ 持続時間の同化効果ははっきりとは生じなくなるという点で、本研究の実験結果においても Takeyasu & Giriko (2017)の報告と類似した結果が得られたと言える。

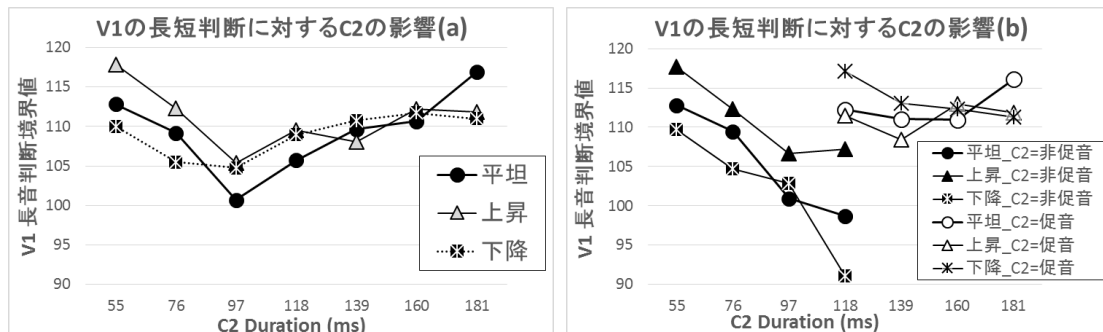


図 2: V₁ の長音判断に対する C₂ の影響---C₂ 音韻長によりデータを分けない場合(a)と分けた場合(b)

また、F0 系列間の結果を比較するため、C₂ が非促音だと判断されたものと、促音だと判断されたものにデータを分割し、F0 系列を独立変数に含めたロジスティック回帰分析によりオッズ比に基づいて 3 系列間の多重比較を行ったところ、C₂ が非促音だと判断された場合には平坦系列と上昇系列の間と、下降系列と上昇系列の間に有意な差が見られ（それぞれ $p < 0.001$ ）、平坦系列と下降系列の間には 10%水準で有意な傾向が見られた。以上の結果を図の判断境界値と合わせて判断すると、上昇系列は最も V₁ が長母音だと判断されにくく、下降系列が最も長母音だと判断されやすいと解釈できる。これは、Takiguchi et al. (2010)において F0 変動と長音の知覚の関係として指摘されているのと同様の傾向を示すものである。その一方で、C₂ が促音だと判断された場合にはいずれの F0 系列の間にも有意な差が見られなかった。なぜこのように F0 変動の影響の現れ方が異なって来るのかについては、やはり現時点では明確な回答を出すことは難しく、稿を改めて検討したい。

3. 結論

本稿では、Takeyasu & Giriko (2017)が報告した隣接要素間の同化効果と対比効果が、彼らの知覚実験とは別の方言話者に、F0 に変動のある語を刺激として用いた場合にも観察されるかどうかを明らかにすることを目的として、福岡方言話者に対して、平坦、上昇、下降の 3 つの F0 パターンを持つ刺激を用いた知覚実験を実施した。知覚実験の結果、東京方言以外の話者においても、また、F0 に変動のある語を刺激として用いても、ある要素の持続時間や音韻長が、隣接する要素に対してそれぞれ同化効果と対比効果を及ぼすことが明らかとなった。

参考文献

- 大深悦子・森庸子・桐谷滋 (2005) 「促音の知覚に対する先行・後続母音長の影響」『音声研究』第 9 巻第 2 号, 59-65.
- 藤崎博也・杉藤美代子 (1977) 「音声の物理的性質」『音韻 (岩波講座 日本語 5)』pp.63-106. 岩波書店.
- 前川喜久雄 (1997) 「日韓対照音声学管見」『日本語と外国語との対照研究 IV 日本語と朝鮮語 (下)』(国立国語研究所編), pp.173-190. くろしお出版.
- Boersma, Paul and David Weenink (2017). Praat: doing phonetics by computer (Version 6.0.28) [Computer program]. <http://www.praat.org/>
- Kawahara, Shigeto (2015) The phonetics of sokuon, or geminate obstruents. Haruo Kubozono (Ed.) *Handbook of Japanese Phonetics and Phonology*. pp.43-77. Berlin: Walter de Gruyter.
- Kubozono, Haruo, Hajime Takeyasu, & Mikio Giriko (2013). On the positional asymmetry of consonant gemination in Japanese loanwords. *Journal of East Asian Linguistics*. 22(4). pp.339-371.
- Takeyasu, Hajime & Mikio Giriko (2017) Effects of duration and phonological length of the preceding/following segments on perception of the length contrasts in Japanese. Haruo Kubozono(Ed.) *The Phonetics and Phonology of Geminate Consonants*. pp.85-117. Oxford: Oxford University Press.
- Takiguchi, Izumi, Hajime Takeyasu, & Mikio Giriko (2010) Effects of a dynamic F0 on the perceived vowel duration in Japanese. *Proceedings of Speech Prosody 2010* [CD-ROM], 100944: 1-4.

英語母語話者と機械による自動評定に基づく 日本語話者の L2 英語の流暢さの評価

近藤真理子(早稲田大学)・Lionel FONTAN(Archean Technology)・Maxime Le Coz (Archean Technology)・Sylvain DETEY(早稲田大学)・小西隆之(早稲田大学大学院)

{mkondo,detey}@waseda.jp, tkonishi@aoni.waseda.jp,
{lfontan, mlecoz}@ archean.tech

1. はじめに

第二言語(L2)でのコミュニケーションや L2 の口頭運用能力を評価するとき、“流暢さ”は評価の重要な要素の一つである(Anderson-Hsieh, Johnson, & Koehler, 1992)。一口に流暢さといっても、単に発音の良し悪しだけでなく、語彙アクセスや統語構造の複雑さ、談話プランニング等、様々な要因が関わっているが、こと L2 の発話において、流暢さは母語(L1)の訛の強さや発音の明瞭さ、分かりやすさ、また発話速度やポーズなどの発話リズム等の影響を大きく受ける(Thomson, 2015; Ghanem & Kang, 2018)。L2 音声教育において、“流暢さ”は「不必要なポーズやその他非流暢と受け取られる要因を排除した滑らかで聞きやすい発話の度合い」(Derwing & Munro, 2015)とされる。母語話者に近い流暢さの習得は L2 学習者にとっては必須であり、そのために流暢さをいかに判定できるかが重要となる。しかし、L2 発話の流暢さの評価方法・基準は一樣ではなく、人による聴覚的印象による評価や、話速や音韻単位の長さや比率、調音の速度、ポーズ長などを音響的に測定し評価する方法など、様々な方法がある(Isaacs & Trofimovitch, 2017; Kim & Ginther, 2018)。また、一般的に第二言語の評価は、その言語を母語とし、その言語を外国語として教える教師や言語学研究者などの専門家が評価することが一般的であるが、実際の言語コミュニケーションは、非専門家である一般の言語使用者同士のほうが圧倒的に多いと考える方が普通であろう。一般の言語使用者にとって流暢に聞こえる、つまり聞きやすく分かりやすい L2 発話はどのようなものか。また専門家の音声分析をもとに構築した機械による自動評定とどう異なるのか。これからの L2 音声研究では大規模コーパスを使った研究が盛んにおこなわれるようになると思われるが、その際に、人による発話レベルの評価はあまり現実的ではなく、精度の高い機械による自動判定システムを整えることが課題となる。

近藤・小西(2017)と小西・近藤(2018)が行った日本語母語話者の英語発話評価の研究では、評価者の英語音声学の知識の有無が、英語レベル判定評価に影響を及ぼしていた可能性が示唆された。アメリカ英語母語話者 5 名、日本語話者 4 名、以下の言語の母語話者各 1 名ずつ計 9 名(韓国語、広東語、スペイン語、中国語、ドイツ語、パンジャブ語、フランス語、ベトナム語、ポーランド語)に、J-AESOP コーパス(第 2 節参照)から、日本語話者がイソップ寓話の『北風と太陽』の英語版“The North Wind and the Sun”を音読したものを、“流暢さ”、“分節音の正確さ”、“韻律”、“母語訛の少なさ”、の四項目について評価してもらい、その各評定値の相関を求めた。評価者は、英会話学校や大学レベルで英語または音

声学を教えているか教えた経験のある教師か研究者である。各項目ごとの評定者間の相関係数は比較的高かったが(0.7以上)、アメリカ英語話者のうちの1名とベトナム語話者の両者の評定値は、他の評価者の評定値との相関が低かった(0.4前後が多い)。特に、このアメリカ英語話者は、アメリカ英語以外の英語方言話者の発話の“母語訛の少なさ(nativeness)”の評価値が低め(6~7/10)で、他の話者の評定値と著しく異なっていた。また日・英語以外がL1の評価者が各一名しかいないため、L1の影響なのか、個人差なのかは定かでないが、ベトナム語話者の評定値も他の評価者の評定とは大きく異なっていた。この2名は、現在は英語教育を専門とし、大学生を対象に英語音声学や発音を教えているが、元々の専門は語用論、言語教育、日本語教育であったので、他の評価者とは異なる英語音声学や音声学の教育を受けている可能性があり、専門知識の有無が評価の違いに繋がっている可能性がある¹。

そこで本研究では、英語音声学の知識が殆どない一般の英語母語話者は何を手掛かりにL2英語の流暢さを評価するか、機械による自動評定と評価が一致するかを検証した。日本語母語話者のL2英語発話の人間による評価を、機械による流暢さの判定と比較し、一般英語母語話者が流暢さをどう判断しているかを検証した。

2. 実験

今回の一般アメリカ英語母語話者による日本語母語話者のL2英語発話の流暢さの評価者として、英語音声学を専門としないアメリカ英語話者の大学院生2名に、日本語話者のL2ソップの寓話である *The North Wind and the Sun* を音読したもの72人分のデータを使用し、流暢さに特化して評価してもらった。J-AESOP-パスは、現在の時点で183人分の日本語話者と25名の英語母語話者のデータが収録されているが、今回は並行して行っているL2フランス語評価のデータベースの大きさとの釣り合い、また評定者の時間的負担等を考慮し、J-AESOPから比較的英語評定値分布のバランスの取れた(図1)、一部72名分の発話データを使用した。

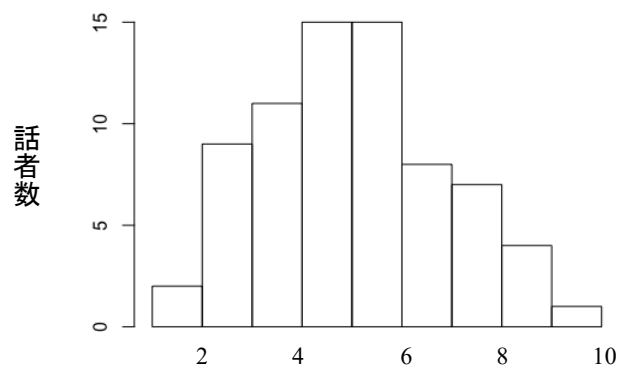


図1. 日本語話者72名の英語評定値(平均=5.06/10)

¹ 従って、J-AESOPの評定からはこの二人の評価を除いている。

J-AESOP 音声コーパスは、話者が音読中につかえたり、読み間違えたり、読み直したり、沈黙してしまったりしたものもあるが、それらに削除等の編集を施さず、そのままコーパスとしている。*The North Wind and the Sun* は 5 文からなっているので、これらのつかえや読み間違えが流暢さの判断に与える影響をできるだけ少なくするため、また特定の分節音の発音の間違ひの影響を最小限にとどめるため、各話者の音読のファイルを一文ずつ 5 つのファイルに分割した(5 文 x 72 名=360 ファイル)。日本語話者の発話データに加え、英語母語話者(方言はアメリカ英語、イングランド英語、オーストラリア英語、カナダ英語、スコットランド英語、ニュージーランド英語等と異なる)25 名が *The North Wind and the Sun* を音読したデータを、同様に 1 文ずつ五分割したもの(5 文 x 25 名=125 ファイル)を加えた 485 全ファイル(360+125 ファイル)をランダム化したものを、今回の評価者であるアメリカ英語話者(以下、英語母語話者)2 名に 0-4 の整数の 5 段階で評価してもらった。今回、発話の評定値を先に評定した整数 1-10 の 10 段階(図 1)で評価せず、0-4 の 5 段階での評価としたのは、現在他言語でも進めている機械による自動評定研究の評定法と合わせたためである。

評定はオンラインで行われ、2 名の評価者は指定されたサイトにログインし、使い方の説明を読んだのち、図 2 の画面上で自分のペースで音声を再生し、0 (Not fluent at all) から 4 (As fluent as a native speaker) から選択するようになっている。評価者は同じ音声ファイルを何度でも聞き直してよいことになっており、既に評価したファイルに戻って評価し直すこともできる。一度に全部のファイルの評価せず、何度かに分けて評価していいと伝えた。またつかえ、言い淀み、繰り返しなどは考慮しないで、流暢さのみを評価するように伝えた。2 名の評価者が要した時間は平均 5 時間程度であった。

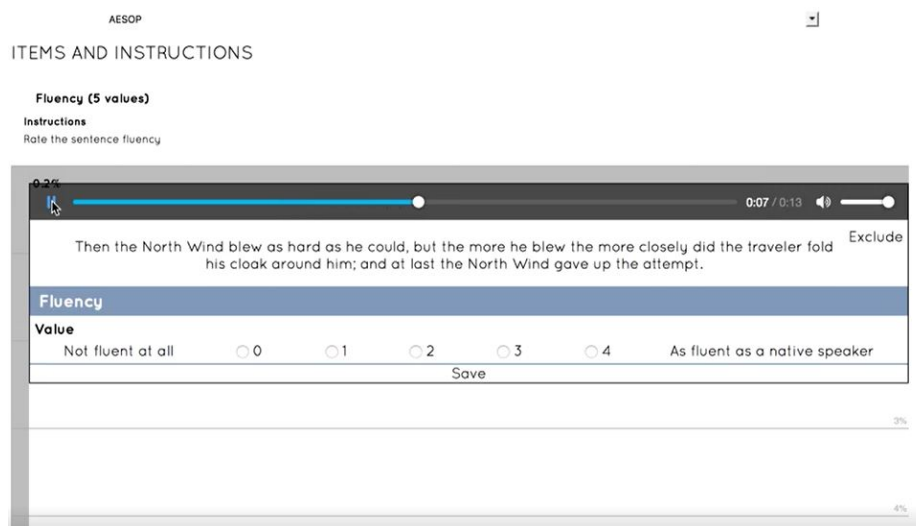


図 2 オンライン評定システム

3. 結果

3.1. アメリカ英語話者による評価

二人の英語母語話者による全発話データ 485 ファイルの流暢さの評定の一致度を、スピ

アマンの順位相関係数で検証したところとても高かった ($\rho=.88, p<.001$)。しかし、英語母語話者 25 名の 125 音声ファイルを除いた日本語話者 72 人分のみ 360 ファイルで両評価者間の評定値の一致度を求めたところ、一致度はやはり $\rho=.77, p<.001$ と高くはあるが、英語話者のデータを加えたときの一致度に比べて低かった。これは図 3 に示されているように、評価者二人が英語母語話者 25 名の 125 音声ファイルの流暢さすべてに評定値 4 をつけたためである。日本語話者の英語の流暢さの評定値の平均は 2.12 (標準偏差値 1.04) であり、一致度は若干低い。

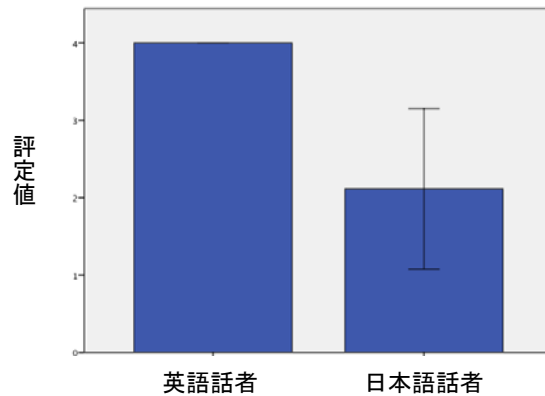


図 3 話者の母語別の評定値の平均。エラーバーは 1 SD

72 名の日本語話者の 360 ファイルそれぞれの流暢さの評定値は広く分散しており、各話者ごとの五つの文それぞれにつけられた評定値の平均値は 2.11 (標準偏差が .99)、最低が 0.1 で、最高が 4 であった。日本語話者の各文ごとの評定値の平均と標準偏差値は図 4 に表されているが、各文の平均評定値には有意な差が見られなかった。つまり、英語母語話者 2 名の評価は信頼できるもので、また両者の評定値間の一致度も高いことが示唆される。

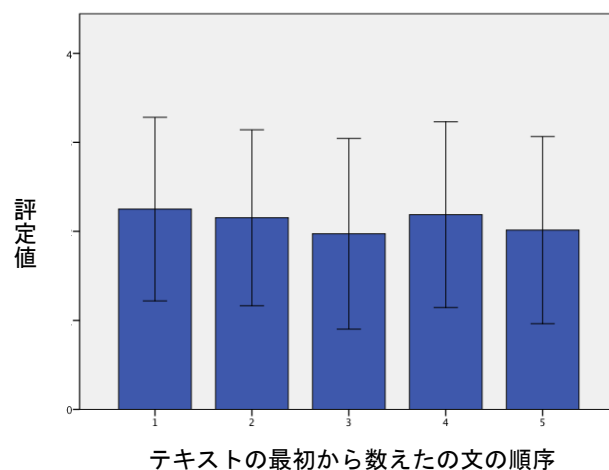


図 4 *The Northwind and the SUN* の五つの文それぞれの平均評定値。
エラーバーは 1 SD

3.2. 自動評定値との相関

上記の2名の非専門英語話者の評定値を、機械による流暢さの自動評定の結果と比較した。この判定で利用した自動評定システムは、Forward-Backward Divergence Segmentation アルゴリズムを使い、通常多くの流暢さの自動評定システムで用いられる発話長とポーズ長の比率や、標準偏差等と共に、話速や話速の規則性を推計し、またフォルマントトラッキング・アルゴリズムを使い、隣接する音の同時調音や同化、融合等の推計を行い、発話がどのくらいよどみがなく、流暢に聞こえるか評価を行っている。このシステムを用いて、日本語母語話者のL2 フランス語音声発話の流暢さの評定判定を行い、既に信頼性が検証されている(Fontan, Le Coz & Detey, 2018)。このシステムを、L2 英語の流暢さの判定にも使えるよう、応用したものである。この自動評定システムを使って同じ72名の日本語話者のL2 英語の流暢さを、同じ0~4の5段階で評価し、英語話者2名の評定値との相関を求めたところ、人間の評価と機械の評価の相関は低かった($r=0.43$)。この機械による自動評定は、これまでの英語音声学やL2 英語音声学などの先行研究の結果を元に、流暢さの判断として発話速度や一定の発話リズム、また前後の分節音との調音の連結、同化、脱落等など“発話の滑らかさ・淀みのなさ”に対する評価が重視されるよう設定されている。しかし、今回の二人の評価者は英語母語話者の発話に対する流暢さの評定がすべて最高点の4であったことを考慮すると、音声学的に異なる基準、おそらく外国語訛の少なさ(nativeness)を加味して、流暢さの評価をしていた可能性がある。両評価者の評定値が低かった音声ファイルを検証してみると、分節音の発音の正確さに問題があるものが少なくなく、両評価者はいわゆる流暢さ以外の要因にも敏感に反応していた。

4. 結論と今後の研究

結果から、一般の英語話者にとっての流暢さの判断は、単に韻律的規準だけでなく、分節音の正確さ等、他の要因の影響も大きく、L1 訛度が強く関与しているらしいことが分かった。近藤・小西(2017)では、英語音声学の知識の有無により、L2 英語発話の評価に違いがある可能性が示唆されたが、自動評定システムは一般的な英語音声学・音韻論の研究の結果を基に構築されているため、専門家の評定と比較的近いと推測される。

本稿で紹介した結果はL2 音声発話評価研究の一部であり、今後はL1の異なる非専門家によるL1日本語話者のL2英語音声の評価や、英語音声学を専門とするの英語母語話者による流暢さ及び他の項目に関する評価研究を、同様の評価システムを使い調査する予定である。それらの結果を自動判定システムの結果と照らし合わせ、専門家と非専門家が何を基にL2 英語音声発話の流暢さの判断を行うのか、また自動判定システムの精度を検証し、いわゆるL2 英語の流暢さ、聞きやすさとは何か、検証を行う予定である。

参考文献

- Anderson-Hsieh, J., Johnson, R. and Koehler, K. (1992) “The relationship between native speaker judgments of non-native pronunciation and deviance in segmentals, prosody, and syllable structure”. *Language Learning*. 42:4. 529-555.
- Bolton, K. (2004) World Englishes. In A. Davies & C. Elder (eds), *Handbook of Applied Linguistics*. (pp.367-396). Oxford: Blackwell.
- Crystal, D. (2003) *English as a Global Language*. Cambridge: CUP.
- Derwing, T. M. and Munro, M. J. (2015) *Pronunciation Fundamentals. Evidence-based Perspective for L2 Teaching and Research*. Amsterdam, Netherlands: John Benjamins.
- Fontan, L., Le Coz, M. and Detey, S. (2018) “Automatically measuring L2 speech fluency without the need of ASR: a proof-of-concept study with Japanese learners of French”, *Proceedings of INTERSPEECH 2018*. Hyderabad, India.
- Isaacs, T. and Trofimovitch, P. (2017) *Second Language Pronunciation Assessment. Interdisciplinary Perspectives*. Bristol, U.K.: Multilingual Matters.
- Kim, O. and Ginther, A. (2018) *Assessment in Second Language Pronunciation*. London, U.K.: Routledge.
- Kondo, M., Tsubaki, H. & Sagisaka, Y. (2015) “Segmental Variation of Japanese Speakers’ English: Analysis of “the North Wind and the Sun” in AESOP Corpus”, 『音声研究』, 19巻1号, 3-17.
- 近藤真理子・小西隆之 (2017) 『通じる英語のための発音教育』, *Conference Handbook 35*, 232-237.
- 小西隆之・近藤真理子 (2018) 『日本語母語話者のL2英語発音評価を構成する音声特性』, 第32回日本音声学会全国大会予稿集.
- Meng, H., Tseng, C., Kondo, M., Harrison, A. and Visceglia, T. (2009) “Studying L2 Suprasegmental Features in Asian Englishes: A Position Paper”, *Proceedings of 2009 INTERSPEECH*, 1715-1718.

パラ言語的情報を伝達するドイツ語心態詞の韻律的特徴 —ドイツ語母語話者とドイツ語学習者の発話と知覚¹

生駒 美喜 (早稲田大学)・小西 隆之 (早稲田大学大学院)
ikoma@waseda.jp, tkonishi@aoni.waseda.jp

1. はじめに

ドイツ語の話しことばには心態詞と呼ばれる語が頻繁に用いられる。ドイツ語の心態詞は、日本語の終助詞と似た機能を持ち(Kawamori 1997), パラ言語的情報(森他 2014)を伝達する。例えばドイツ語心態詞 *schon* は文脈によって「確信」「留保付肯定」「反論」の心的態度を表す(Ikoma 2007)。

Ikoma (2018)は心態詞 *schon* を含む「反論」のドイツ語母語話者による発話を「確信」と比較した結果、発話全体の F0 最大値が低く持続時間が長いという特徴が見られた。生駒 (2017) は心態詞 *schon* と同一の状況で発話した日本語母語話者による日本語発話を分析し、「反論」の発話を他の発話意図と比較した結果、発話全体の F0 最大値が高く、発話全体の持続時間が長く、「留保付肯定」における発話全体の F0 レンジが小さいことが明らかになった。

パラ言語的情報の第 2 言語習得に関しては、「母語からの転移や文化の違いによる影響を受けやすい」(福岡 2017: 2)とされるが、パラ言語的情報の知覚においては普遍性が見られることも明らかになってきている。

それでは日本語を母語とするドイツ語学習者の心態詞の発話と知覚はどうなっているだろうか。Ikoma (2016)はドイツ語学習者を対象に、留学前後におけるドイツ語心態詞 *schon* の発話と知覚を調査した。知覚実験の結果、留学後の正答率は母語話者の正答率と差がなく、留学前よりも正答率が上がっていた。また 1 名の学習者の留学後の「反論」の発話において、アクセントを持つ心態詞 *schon* の音節部分の F0 レンジが「留保付肯定」における F0 レンジよりも増大していた。但し Ikoma (2016)では発話全体の音響特徴は分析していない。

以上の先行研究の結果をふまえ、本研究はドイツ語学習者の心態詞 *schon* の発話と知覚を解明するために発話全体の韻律的特徴にも着目して分析を行い、以下の仮説を検証することを目的とする：

- 1) ドイツ語学習者の発話において、F0 の特徴はドイツ語母語話者とは異なるが、持続時間の特徴はドイツ語母語話者(以下, *de*)と共通している。
- 2) ドイツ語圏長期滞在経験のあるドイツ語学習者(以下, *ja1*)の発話は、長期滞在経験のないドイツ語学習者(以下, *ja2*)とは異なり、*de* の発話と類似する韻律的特徴が見られる。
- 3) ドイツ語学習者 *ja1*, *ja2* のドイツ語心態詞の知覚においては、*de* よりも正答率は低くなるが、誤答率はドイツ語母語話者に類似している。

¹ 本研究は、早稲田大学特定課題(基礎助成)2018K-002、特定課題(B)2018B-003、および早稲田大学現代政治経済研究所「社会行動と言語選択」特別研究部会(研究代表者:生駒美喜)の助成を受けています。

- 4) 知覚実験における ja1 の正答率は、ja2 と比較して高い。

2. 発話実験

2.1. 実験の資料・被験者・手順

本研究では、Ikoma (2018)で用いた短文 Peter kommt schon (ペーターは来る)に加え、心態詞 schon が文中に位置する Peter kommt schon zum Unterricht. (ペーターは授業に来る) の2つの短文を発話文として用いることとし、それぞれの短文が上記1.で述べた「確信」「留保付肯定」「反論」の意図となるような状況文および対話文を作成し、実験資料として用いた。

実験の被験者としての日本語を母語とするドイツ語学習者10名(男性4名、女性6名)²に協力してもらった。このうち4名はドイツ語圏に留学等で1年以上滞在した経験がある。残る6名はドイツ語圏の長期滞在経験はなく、全員が1年以上ドイツ語を日本の大学等で学習している。

以上の学習者を対象に、Ikoma (2018)における2人1組の発話実験と同様の手順にて録音を行った³。さらに母語話者の発話と比較するため、同様の発話実験をドイツ語母語話者計20名⁴の協力により実施した。

2.2. 音響分析

以下の点について発話データの音響分析を行った⁵：

- 1) 持続時間：発話全体、各音節における母音部分の持続時間、ピッチアクセントから発話末までの持続時間
- 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、ピッチアクセントのある音節におけるF0最大値、F0最小値、F0平均値、音節開始からF0ピークまでのF0レンジ、F0ピークのタイミング、ピッチアクセントから発話末までのF0レンジ
- 3) 振幅：発話全体の最大値、最小値、平均値

分析においては、ja1, ja2, de のグループ間に差が見られるかを一元配置分散分析(対応なし)を用いて調べた。また上記の音響特徴について、ja1, ja2, de

² 6名が関東地方出身者。1名が関西地方、1名が九州地方、2名が中部地方出身者である。長期滞在経験のある学習者4名のうち2名が関東地方出身者、2名が中部地方出身者。

³ 発話時に本来とは異なる語にアクセントを置くことを避けるため、発話資料のうち、多くのケースに schon が置かれるとされる「留保付肯定」の文にのみ、schon を大文字、太字で示し、アクセントが置かれることをあらかじめ伝えた。動詞部分の kommt および心態詞 schon 以外の場所にアクセントを置いた発話は分析の対象から外した。

⁴ 10名(女性8名、男性2名)は、2017年2月に Peter kommt schon のみの発話実験を実施した。分析結果は Ikoma (2018)を参照。別の10名(女性6名、男性4名)に対し、2018年2月に Peter kommt schon zum Unterricht の発話実験を実施した。

⁵ 音響分析に際し、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。

毎に、「確信」「留保付肯定」「反論」の3つの発話意図の間に有意差が見られるかを一元配置分散分析（対応あり）で調べた。

2.3. 発話実験結果と考察

2.3.1. 文1, 文2の発話におけるピッチアクセントの位置と出現頻度

ドイツ語母語話者（de）の確信の発話は文1では全て（100%）kommtにアクセントが置かれているが、文2では30%の発話にschonにアクセントが置かれていた。また、deの反論の発話において文1ではkommtにアクセントが置かれるケースが多い（73%）のに対し、文2はschonにアクセントが置かれるケースが多かった（70%）⁶。一方、留保付肯定では、deの発話は文1, 文2共にschonにアクセントが置かれるケースが大半を占めた（文1:80%, 文2:73%）。以上のように発話文によってアクセントの位置の違いが見られるため、以下の分析ではschonにアクセントが置かれる発話と、kommtにアクセントが置かれる発話に分けて分析を行った。

2.3.2. 文1, 文2の各々の発話意図におけるドイツ語学習者と母語話者の比較

「確信」「留保付肯定」「反論」の状況下における発話の韻律的特徴をde, ja1, ja2の間で比較したところ、主として以下の結果が得られた：

- 1) 文1「確信」でkommtにアクセントの置かれるdeの発話におけるkommtの母音部の持続時間（発話全体の持続時間に対する比率）が、ja1, ja2よりも長い($ps<.001$)。また、文1「留保付肯定」, 「反論」の状況下でschonにアクセントの置かれる音節において、de, ja1におけるschonの母音部の持続時間（%）が、ja2と比較して長い($ps<.001$)。
- 2) 文1「留保付肯定」の状況下でschonにアクセントが置かれる場合、ピッチアクセント～発話末までのja2のF0レンジがde, ja1と比較して小さい($ps<.001$)。同様に、文2「留保付肯定」でschonにアクセントが置かれる発話においてもピッチアクセント以降のja2のF0レンジがdeと比較して小さい($p<.05$)。文2「反論」においてもdeと比較してja2のF0レンジが小さい($p<.05$)。このことからja2の発話では発話末までのイントネーションが平坦になっていることが示唆される。
- 3) 文1「反論」でkommtにアクセントが置かれる発話において、ja1の発話全体のF0レンジがdeに比べて小さく($p<.01$)、kommtの母音開始部分からF0ピークまでのレンジがja1においてdeと比較して小さく、F0ピークまでの持続時間も短い($ps<.01$)。

2.3.3. ドイツ語学習者, 母語話者毎の発話意図における韻律的特徴

一元配置分散分析（対応あり）の検定を行った結果、母語話者の文1の発話において、発話全体の持続時間が留保付肯定>反論>確信の順に長く($ps<.001$)（図1左）、F0最小値が留保付肯定>確信>反論の順になっていた($ps<.001$)（図

⁶ 同一の発話意図で、文構造および心態詞の文中での位置が異なることによりピッチアクセントの位置が異なるという結果は、心態詞の発話のピッチアクセントの位置に音韻的さらには音声的な要因が関わる可能性を示唆しており、同一のドイツ語母語話者による発話を用いた更なる分析が必要である。

1 右). また, 発話全体の F0 レンジが反論 > 留保付肯定 ($p < .01$)であった.

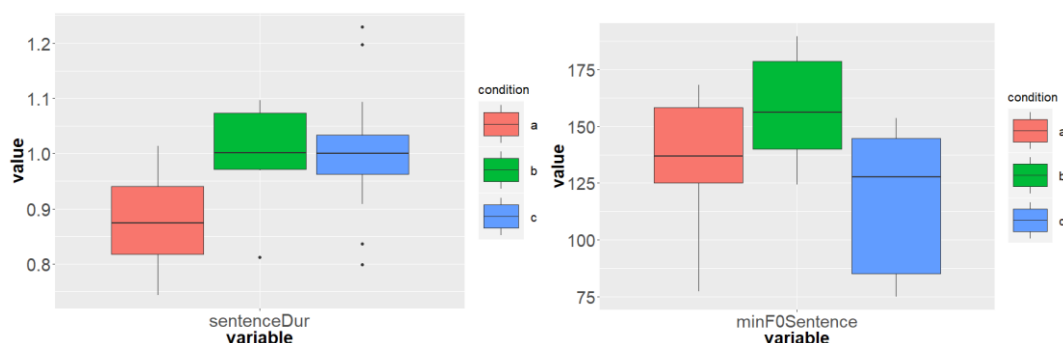


図 1: ドイツ語母語話者の文 1 (Peter kommt schon) の発話全体の持続時間 (左) (sec) および F0 最小値 (Hz) (右) (a=確信, b=留保付肯定, c=反論)

これに対し学習者の発話においては, 文 1 の発話全体の持続時間は ja1 において反論 > 確信 ($p < .05$), ja2 において留保付肯定 > 反論 ($p < .05$) という結果が見られたが, その他は 3 つの発話意図間での有意差は見られなかった.

文 2 の発話では, schon にアクセントが置かれる発話に関して, ピッチアクセント ~ 発話末の F0 レンジが de, ja1 において反論 > 確信 (de: $p < .05$; ja1: $p < .01$) となっていたが, ja2 の発話には有意差が見られなかった.

さらにアクセントの位置に関わらず全ての発話について分析を行ったところ, de の発話全体の振幅の最小値が反論 > 留保付肯定 > 確信 ($p < .001$) となっていた (図 2) のに対し, ja1, ja2 のいずれの発話にも有意差は見られなかった.

一方, ja2 の発話全体の F0 平均値は反論 > 確信 ($p < .01$) となっていたが, de の発話には有意差が見られなかった.

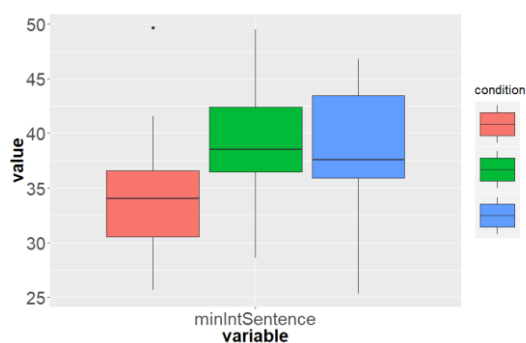


図 2: ドイツ語母語話者の文 2 (Peter kommt schon zum Unterricht) の発話全体の振幅最小値 (dB) (a=確信, b=留保付肯定, c=反論)

3. schon を含む短文を用いた知覚実験

3.1. 実験の資料, 被験者および手順

上記 2. の発話実験に参加した被験者 10 名のうち 7 名が知覚実験に参加した. 比較のため, ドイツ語母語話者 4 名に知覚実験に参加してもらった. 実験は Ikoma (2018) と同様の手順にて, Praat の知覚実験プログラムを用いて行った. 知覚実験に用いた発話は 2. の発話実験にて母語話者が発話した文 1 "Peter

kommt schon”および文 2”Peter kommt schon zum Unterricht”の発話データ⁷である。知覚実験の被験者はランダムに提示された音声聞き、確信、留保付肯定、反論のいずれか 1 つの発話意図を選択した。

3.2. 知覚実験結果

de, ja1, ja2 毎の知覚実験の回答率 (%) を以下の図 3 に示す。de, ja1, ja2 で回答率に有意差が見られるかカイ二乗検定を行った結果、文 2 の反論の発話を除き、全ての発話意図において有意差が見られた ($p < .01$)。しかしながら ja1 の正答率は ja2 の正答率と比較して de すなわちドイツ語母語話者に近似しており、ja2 においても、母語話者にも知覚が困難であった反論を除く発話において正答率は 40% となっていた。

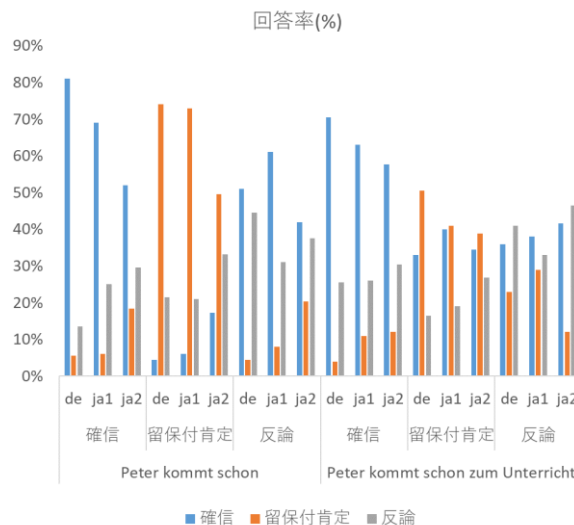


図 4: 文 1(左: Peter kommt schon)および文 2(右: Peter kommt schon zum Unterricht)の知覚実験の回答率 (%)

4. まとめと今後の展望

以上行ったドイツ語心態詞 schon を含む短文のドイツ語学習者およびドイツ語母語話者の発話の音響分析、およびドイツ語学習者、ドイツ語母語話者による知覚実験の結果から、以下の点が明らかになった:

- 1) 仮説 1 については発話全体の持続時間が留保付肯定 > 反論 > 確信の順に長いというドイツ語母語話者と共通の特徴がドイツ語学習者にも見られた一方、アクセントの置かれる音節の持続時間や、発話全体の F0 の特徴がドイツ語母語話者とは異なっていた。さらに発話全体の振幅最小値がドイツ語母語話者において反論 > 留保付肯定 > 確信の順に大きいのに対し、学習者にはそのような違いは見られなかった。
- 2) 仮説 2 に関して、一部においては長期滞在経験のない学習者と異なる特

⁷ 2. の発話実験にて被験者は同じ状況の文を 3 回繰り返して発話している。知覚実験では、それぞれ 2 回目の発話のみを資料として用いた。

徴が見られたものの、長期滞在経験がある学習者においても、アクセントの置かれる音節の母音の持続時間が短いなど、母語話者とは異なる韻律的特徴が明らかになった。

- 3) 仮説3については、ほぼ仮説通りの結果が得られた。ドイツ語学習者への正答率はドイツ語母語話者よりも全体的に低いが、誤答率はドイツ語母語話者と共通しており、ドイツ語母語話者と類似した傾向がある。
- 4) 仮説4についても、正しいことが検証された。ドイツ語圏長期滞在経験のあるドイツ語学習者は長期滞在経験のない学習者と比較して総じて正答率は高くなっていた。

今後の研究では、ドイツ語学習者の数を増やし、学習歴を考慮したより細かなグループ分けを行い、学習者の発話の音声特徴についてより詳細な分析を行いたい。知覚実験においてもより多くのデータで分析を進め、ドイツ語母語話者による評定実験を実施し、その結果に基づきより詳細に音響分析を行い、ドイツ語学習者の発話の音声特徴を明らかにしたい。

参考文献

- Ikoma, Miki (2007) *Prosodische Eigenschaften der deutschen Modalpartikeln*. (Schriftenreihe PHONOLOGIA, Band 103). Hamburg, Dr. Kovač.
- 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 回日本音声学会全国大会予稿集』日本音声学会, 123-128.
- Ikoma, Miki (2016) „Produktion und Wahrnehmung der deutschen Modalpartikel *schon* durch japanische Deutschlernende“. In: Zhu, Jianhua, Zhao, Jin and Szurawitzki, Michael (eds.) *Akten des XIII. Internationalen Germanistenkongresses Shanghai 2015*. Peter Lang, 253-259.
- 生駒美喜(2017) 「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」. 『第 31 回日本音声学会全国大会予稿集』日本音声学会, 216-221.
- Ikoma, Miki (2018) „Prosodie und Bedeutung der unbetonten und betonten Modalpartikel *schon*“, *Akten des 44. Linguisten-Seminars, Tokyo 2016*. München, Iudicium, 53-68.
- 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.
- 福岡昌子(2017) 「パラ言語的情報の強調順位に関する日本語学習者の知覚」『音声研究』第 21 巻第 3 号. 日本音声学会, 1-14.
- 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.

現代モンゴル語のアクセントについて

包桂蘭 (内モンゴル大学モンゴル学院)
axita@163.com

1. はじめに

モンゴル語のアクセントについて、少数の学者達が一般言語学で言う意味のアクセントが存在しないと考えているが、多くの学者達は、やはりモンゴル語にはアクセントがあるという見方を持っている。伝統的な学説の主な観点は「第一音節に当たる強弱アクセント」である。例えば、「モンゴル語のアクセントは、主に第1音節に当たっている音勢的なアクセント」(Б·Я·Владимирцов、1988)、「モンゴル語の音勢的なアクセントには第1アクセント、第2アクセントと無アクセントの3種類があり、第1アクセントは第1の音節、第2アクセントは第2の音節にあり、その後の音節にはアクセントがない」(Ц·луvsанвандан、1982)、「モンゴル語のアクセントは、第1音節に固定されている、その主な特徴は音勢が強いのである」(清格爾泰、1991)等の論点を述べている。最近25年間、学者達がモンゴル語のアクセントについていろいろ音響学分析をした結果、「第一音節に当たるアクセント」という見方を根本から否定したものであるが、今までモンゴル語のアクセントの本質と単語における位置などの基本的な問題について意見が一致していない所が多いのが現状である。彼らの討論の重点は主に以下の2点の問題をめぐっている：(1)アクセントの位置については、モンゴル語のアクセントは固定しているか？固定であれば、一体どの音節に固定されているか？もし非固定であるならば、分布の規則がどうなるのか？(2)アクセントの性質については、モンゴル語のアクセントは強弱に基づいたストレスアクセントか？高低に基づいたピッチアクセントか？長短に基づいた音長アクセントなのか？あるいはいくつかの要素が共に作用するものであるか？

1.1. モンゴル語のアクセントの位置について

確精扎布(1993)はモンゴル語に音響学分析を行った後、「単語において母音の構造が短一長である場合、アクセントが確かに非語頭の長音音節に当たっているし、母音が短一短と長一長である場合も、非語頭の音節に当たっている」ということを発見した。そして、「チャハル方言を代表した標準モンゴル語のアクセントは第一音節に固定されていない」という結論が出た。白音門徳(1997)は、バーリンとホルチン方言のアクセントを分析して、「モンゴル語にアクセントがあるとすれば、ストレスアクセントとピッチアクセントのいずれも、アクセントが第一の音節に固定されていない場合である。分析したデータから見ると、アクセントが第2音節にあると考えざるを得ない」と述べている。呼和(2009)は「モンゴル語のアクセントは独特、習慣的な規則を持つアクセントで不規則的なものではない」と「類型学の角度から見ると、モンゴル語のアクセントは自由なアクセントであり、固定

的なアクセントではない」と論じている。敖登其木格（2013）もハルハ方言に音響学分析と聴解実験を行って、以上述べられた論点と大体一致した結論を出した。すなわち、「単語には長音がある場合、アクセントが長音の音節に当たり、2 つ以上の長音音節がある場合、アクセントが前の長音の音節に当たることである。いくつかの短音音節で構成された多音節語のアクセントが非語頭の音節に当たりますが、具体的にはどちらかの音節に固定されていない」ということである。

これら以外、多くのモンゴル語の方言における音響学分析を行ったデータから見ると、モンゴル語のアクセントはやはり第1の音節に固定されていないと考えられる。

学者たちがまとめた高さや強さのデータによると、高さや強さのピークが第2の音節にある割合が多い。それで、モンゴル語のアクセントが第2の音節にはあると判断している方もいる（宝玉柱、2007）。これによってモンゴル語のアクセントの標準的な位置がその第2の音節に位置することを断定するならば、われわれはアクセントが他の音節にある変異の規則を探求する必要がある。もし規則があるならば、この結論を論証することができる。しかし、現在の多くの研究結果を分析して、アクセントは一体どの音節に当たっているのか、まだ固定モードが見つかっていないし、モンゴル語の諸方言にも分布が一致していないのである。白音門徳（2014）は、モンゴル語の25の方言に分析した音声データに基づいて、「モンゴル語には一般言語学で言っている意味でのアクセントは存在しない。モンゴル語の多音節語のある音節を強くあるいは高く発音する現象はあるが、それはアクセントではなくて、発音の必要に応じて発生した変化である」と述べた。

1.2. モンゴル語のアクセントの本質について

郑玉玲と鲍怀翹（2001）はモンゴル語のアクセントには「最も重要な音響学関連物は音長」と思う。また、「アクセントに関する諸要因の中には、相対的に変化しているその要素がアクセントとして感知られていることで、モンゴル語のアクセントは躍動的なアクセントになっている」と分析した。宝玉柱（2007）は「ピッチピークが相対的に安定していて、他の要素の影響を受けるのが長さより小さいので、モンゴル語はピッチアクセント」と断定している。呼和（2007）はモンゴル語のアクセントについて「高さ、長さや強さなど様々な要素による「卓立」(Prominence) や「高揚」(Culminative) 現象である」と言っている。敖登其木格（2013）は「高さの影響は他の要素より著しい、高さでアクセントを分けることができない場合、長さや強さが役に立つ」という。

上述の実験研究によると、モンゴル語のアクセントには音の高さ、強さと長さが多かれ少なかれ作用することで、モンゴル語の諸方言の中でそれらの影響の程度も異なっている。ある要素の影響は他より明らかであること、いくつかの要素が共に作用することや互いに交替して作用することもある。そして、2014年（呼和）に発表された『再論モンゴル語のアクセント』では、モンゴル語のアクセントは「ある要素に基づく単一の性質的なアクセントではない」という観点を堅持し、モンゴル語のアクセントを「絶対的なアクセント」と「相対的なアクセント」という2種類に分けた。また、「絶対的なアクセントモードでは、

音色の作用が相対的に大きい」とし、「相対的なアクセントモードは、音声の四つの要素の変化による総合的な効果で、その中には長さや強さの役割が他より大きいかもしれない」としている。

2. 音色とアクセント

学者たちがモンゴル語のアクセントについて、よく「アクセントが落ちている音節が、はっきりしている」と言っているが、これはその音節の音素の構造がより完備しているし、発音にも着目していることを表明しているため、感知の上で他の音節よりはっきりしているというわけである。モンゴル語の母音と子音の先行研究により、語頭音節の母音の発音はより完全で、着位して、非語頭音節の短音はある程度に「央化」する、すなわち、「弱化」しているということである。また、語頭の子音の発音もより着位し、前後の音素の影響を受けにくいということがわかる。それで、ある学者（李兵、2010）は、モンゴル語のアクセントは第1音節に落ちている音色的なアクセントであるとも考えている。林茂灿（1990）中国語の普通話の軽声について「軽音音節の音色は明らかに減縮している。主に母音の音響学的な空間が縮小するとか、子音の発音が着目できないとか現象がある。そのため、高さ、強さと長さなどからアクセントを分析するだけでは不十分で、音色からアクセントを分析することも重要である」と述べている。2006年（Yuling Zheng & Jianfen Cao）に発表した『Coarticulation and prosodic hierarchy』一文にはEPGを用いて、舌と口蓋の接触面積から発音の強化や減縮などの表現により中国語普通話のリズムの階層の特徴を探究した。

以上に述べていることに対して、本文ではアメリカのKay社製の6300型エレクトロパトグラフィ（EPG）、3700型Multi-Speechと南開大学のMinistechLbなどの分析機械やソフトウェアを用いて、モンゴル語の単語を分析し、舌と口蓋の接触面積から音色がアクセントとなにか関係があるのかを調べてみた。

図1は実験に利用したEPGの人工口蓋床に配置された96個の電極の分布図である。上から下までR1、R2……R12の12行に分かれ、真ん中の1列（C6）に対称して、両端から中へ、C1、C2……C5の11列に分かれている。

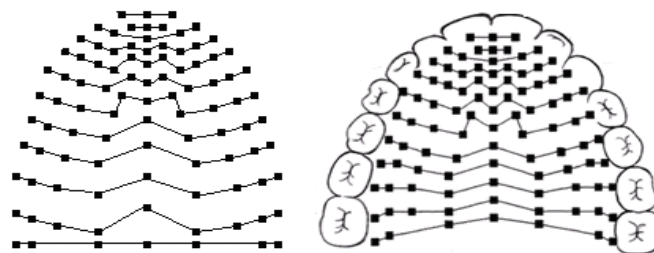


図 1: 人工口蓋床の電極の分布図

舌と口蓋の接触面積が音節における表現を比べるために、第1と第2の音節に同じ音段がある二音節語を選んで分析する対象とした。例：[lep/lɜx]（問い合わせ）、[tʰe/tʰɜx]（引く）、[tʃi/tʃik]（小さい）、[sɛn/sɜr]（宇宙）、[ʃiφ/ʃik]（恥辱）等。計18の実験語には、音

節における高さや強さの分布は、3つの単語がH-LとS-W、その他の単語がL-HとW-Sである。母音の音長分布（物理的長さ）は基本的にS-LとL-S構造が半分ずつとなっている。では、舌と口蓋の接触面積のピークは高さや強さのピークがある音節に落ちるか？やはりより長い母音の音節に落ちるか？または他の分布法則があるのか？

次は、[ʃiϕik]（恥辱）という一語の語図と同期の舌と口蓋の接触面積（RCA、ratio of contact area）の動的な図を例にあげてみよう。RCA値は、毎行の電極の接触率を合わせて、各フレームにおける舌が口蓋に接触する面積率を得ること。RCAがゼロになるとは、この時点で舌が口蓋に接触していないことを示す。RCA値の増加は、舌と口蓋の接触が増えていることを示し、減少した場合には、舌と口蓋の接触が減少していることを示している。RCA動的な図は音声を発声する過程には舌が口蓋に接触する変化状況をよく反映できるものである。

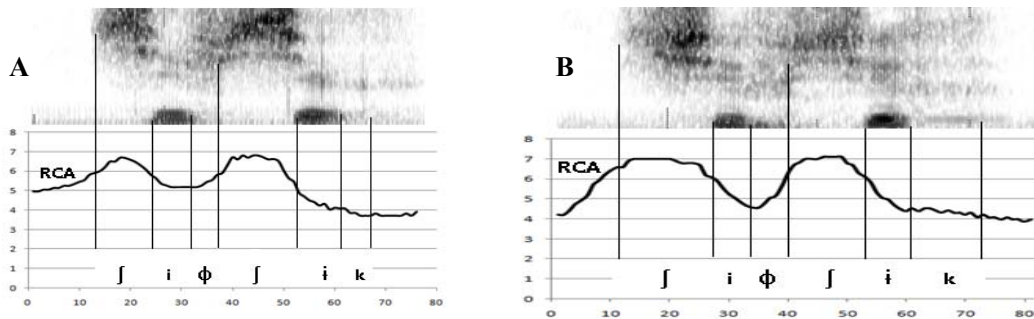


図 2: [ʃiϕik]（恥辱）一語の語図と同期の RCA 図

図 2 は 2 人の被験者が発声した音声サンプル（図 A はチャハル方言地区の被験者（W）、図 B はホルチン方言地区の被験者（H））を分析した図である。2 人の被験者が発声した [ʃiϕik] 一語の高さ、強さと長さの分布は同じ L-H、W-S、S-L 型である。図 2 によると、A と B のいずれにしても第 1 音節と第 2 音節の最高 RCA 値が大体同じで、はっきりした差が見えない。また、2 人の発声には、音節における舌と口蓋の接触最高点の持続時間の分布も一致していない。続いて、もう一つの例を挙げてみよう。

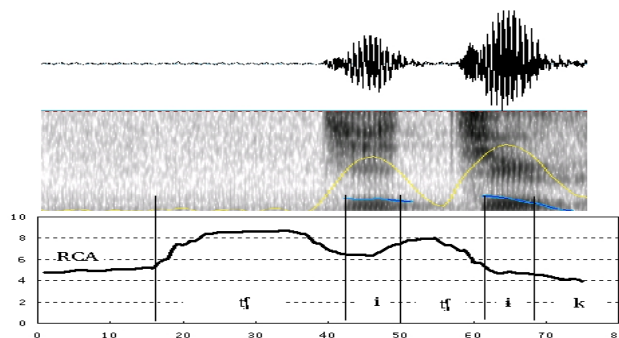


図 3: [tʃi/tʃik]（小さい）一語の語図と同期の RCA 図

図 3 の [tʃi/tʃik] という語の高さ、強さと長さの分布は L-H、W-S、S-L 型であるが、第

1 音節の R C A 値が第 2 音節より少し大きい。つまり、その音色の表現は強さ、高さ、長さのピーク表現と簡単に呼応していないということである。

同じように、他の実験語にも、舌と口蓋の接触面積のピークは第 1 の音節にも、第 2 の音節にも来るのがある。したがって、今回の実験の限りでは、舌と口蓋の接触面積が音節に於いて何が固定的な分布規律なのかわからないのである。その舌と口蓋の接触面積の最高点が高さと強さと同じ音節にくる場合と、そうでない場合がある。それで、音色がモンゴル語のアクセントに何が影響を与えているのか、あるいは何が関係しているのかについて、もっと研究を進める必要がある。

3. アクセントにおける聴解実験

郑玉玲と鲍怀翘（2001）は 30 人（10 人はモンゴル言語文学の専門、10 人は数学の専門であるモンゴル人、他の 10 人はモンゴル語がわからない中国人）の学生を対象して、モンゴル 3 音節語に於いて聴解実験を行った結果、みんなはそのアクセントの位置について聴解した判断結果がほぼ同じで、その音響パラメーターでも密接な関係があることがわかった。敖登其木格（2013）もハルハ方言に音響学分析と聴解実験を行って、「母語者が音節における高低と強弱をはっきり聞き分けている」と述べている。しかしながら、呼和（2007）は郑玉玲と鲍怀翘（2001）が行った聴解実験のデータを違う角度から分析して、「聴覚のみによってアクセントの位置を判定するのは、主観的な要因の影響を受ける可能性がある」と述べている。

3.1. 対象及び方法

今回の試験で、アナウンサー専門の学生が発した録音をモンゴル語が全然わからない中国人（日本に留学している修士 6 人）と日本人（ボランティアしている日本語の教師 6 人）に 3—5 回聞かせて、二音節語と三音節語におけるモンゴル語のアクセントの位置とアクセントにかかわる主な要素について調べた。

3.2. 結果

表 1 は実験語の音節構造の分布、個数とそれらに落ちるアクセントの位置について被験者が感じた状況を統計したデータである。音節構造の S(short)は短母音、L(long)は長母音（二重母音を含む）を示す。このデータによると、アクセントの位置について、長音のある場合、日本人も中国人もその音節にはアクセントがあると思うが、長音がない場合、ほぼ第二音節に、他より卓立していると感じている。アクセントにかかわる要素について、日本人も中国人も長音の音節は他より長くて、はっきりしていると感じている。長音がない場合、中国人の留学生達は、そのアクセントが強さによって違っているという割合がより多いが、やはり強さや高さや長さのどちらかを理解することがちょっと難しいということである。日本人の教師の場合では、主に強弱と長短によってアクセントを感じ、高低にはあまりかかわらないと感じる。それで、みんなアクセントを聴解弁別する主な根拠は音長であり、次は、強さ、高さの順番になる。

この 100 単語における音響分析によると、二音節語の中で、第二音節には強さのピーク

が 85%落ち、高さのピークが 93%落ちている。三音節語の中で、分布状況が少しい複雑でもっと多い実験語に調べる必要がある。

英語、日本語とモンゴル語のアクセントを比較すると、モンゴル語は他の二言語より高低の差も強弱の差も少ない。例えば、城生佰太郎（2014）の実験データにより、第二音節にストレスがある英語の「subject」一語のストレスのある位置では、高さが 219.1Hz で強さが 52.7dB であり、ストレスのない位置では、高さが 179.7Hz で強さが 44.4dB である。つまり、物理量では、高低の差は 39.4 Hz で強弱の差は 8.3 dB である。また、第一音節にアクセントがある日本語の「朝」一語の「ア」は、高さが 157.1Hz で強さが 59.8 dB であり、「サ」は高さが 107.1Hz で強さが 51.9dB である。つまり、高低の差は 50 Hz で強弱の差は 7.9 dB である。今回の実験によると、モンゴル語の音節における、物理量的な高低の差は 3-22 Hz で強弱の差は 0.1-5dB である。

表 1: 聴解実験の統計データ

アクセントの位置 音節の構造		留学生（中国人、6人）					日本語教師（日本人、6人）				
		第1音節	第2音節	第3音節	ない	わからない	第1音節	第2音節	第3音節	ない	わからない
2 音節 語 (70)	S-S (28)	29%	62%	---	8%	1%	7%	75%	---	10%	8%
	S-L (22)	20%	79%	---	---	1%	4%	91%	---	---	4%
	L-S (14)	70%	30%	---	---		35%	14%	---	21%	30%
	L-L (6)	45%	50%	---	---	5%	25%	50%	---	---	25%
3 音節 語 (30)	L-L-L (3)	31%	37%	25%	2%	4%	15%	23	52%	---	10%
	L-L-S (3)	42%	38%	17%	---	3%	26%	48%	12%	---	14%
	L-S-S (3)	56%	26%	15%	---	3%	41%	26%	21%	---	12%
	S-S-S (8)	14%	58%	28%	---	---	7%	47%	40%	---	6%
	S-S-L (7)	27%	27%	34%	---	12%	2%	3%	95%	---	---
	S-L-S (3)	32%	38%	22%	---	8%	9%	61%	30%	---	---
	S-L-L (3)	22%	24%	39%	---	15%	5%	26%	69%	---	---

4. 結論と討論

本研究では、代表的な研究を取り上げながら、音響分析と生理分析及び聴解実験を通じて、単語における現代モンゴル語のアクセントについて検討した。今回の実験を先行研究に結びつけて、以下の四点に概括する。

(1) 高さも強さも第二音節にくる単語が圧倒的に多く、高さについては、第二音節にくることで特に安定している。

(2) EPG を用いた実験の結果、舌と口蓋の接触面積が音節に於いて、固定的な分布規律が見つからないのである。それで、音色がモンゴル語のアクセントに何が影響を与えているのかについてもっと研究が進める必要がある。

(3) モンゴル語がわからない日本人と中国人を対象にした聴解実験を行った結果、アクセントを聴解弁別する主な根拠は音長であり、次は、強さ、高さの順番になる。

(4) 英語、日本語とモンゴル語のアクセントを比較すると、モンゴル語は他の二言語より高低の差も強弱の差も少ない。それはモンゴル語のアクセントが語の意味を区別する機能がないからと考えられ、よって、単語におけるアクセントの位置が他より移り易い現象の一つの原因でもある。

参考文献

- 符拉基米尔佐夫 (Б·Я·Владимирцов 1988) 『蒙古语书面语与喀尔喀方言比较语法』(陈伟、陈鹏翻译) 西宁：青海人民出版社。
- 舍·罗布苍旺丹 (Ш·луvsанvандан 1982) 『现代蒙古语』呼和浩特：内モンゴル人民出版社。
- 清格爾泰 (1991) 『蒙古语语法』フフホト：内モンゴル人民出版社。
- 确精扎布 (1993) 「关于蒙古语重音-语音实验中间报告」『内蒙古大学学报』第 1 期、pp.1-14.
- 白音門德 (1997) 『巴林土语研究』フフホト：内モンゴル人民出版社。
- 呼和 (2009) 『蒙古语语音实验研究』沈阳：辽宁民族出版社。
- 宝玉柱 (2007) 「现代蒙古语正蓝旗土语重音研究」『中央民族大学学报』第 6 期。
- 敖登其木格 (2013) 『蒙古语喀尔喀方言重音研究』内モンゴル大学博士論文。
- 白音門德 (2014) 「モンゴル語のアクセントの音響音声学的分析」『千葉大学ユーラシア言語文化論集 16』p.101-122.
- 郑玉玲・鲍怀翹 (2001) 「蒙古语三音节词韵律模式」『新世纪的现代语音学-第五届全国现代语音学学术会议论文集』北京：清华大学出版社。
- 呼和 (2007) 「再论蒙古语词重音问题」『民族语文』第 5 期、pp.1-14.
- 李兵・贺俊杰 (2010) 「蒙古语卫拉特方言双音节词重音的实验语音学分析」『民族语文』第 5 期。
- 林茂灿 (1990) 「普通话轻声与轻重音」『语言教学与研究』第 3 期。
- Yuling Zheng · Jianfen Cao (2006) "Coarticulation and prosodic hierarchy" . *Second International Conference on Tonal Aspects of Languages*(pp.145-150) .La Rochelle.
- 城生佰太郎 (2014) 「モンゴル語学における若干先行研究：音声学・音韻論関係の紹介とコメント」『文学部紀要』27(2)、51-70.

母語話者日本語音声を対象にした各種文脈における 語頭 2 モーラの F_0 上昇に関する分析

吉澤 風希・峯松 信明・齋藤 大輔（東京大学大学院）
{yoshizawa, mine, dsk_saito}@gavo.t.u-tokyo.ac.jp

1 はじめに

日本語東京方言における 1 型以外の単語の孤立単語発声において、「第 1 モーラから第 2 モーラにかけてピッチが上昇する」と説明されることが多い。音声において心理量であるピッチに対応する物理量は基本周波数 (F_0) であるが、第 2 モーラが長母音の二拍目または撥音である場合には F_0 の上昇が抑えられる。これは異音的変動と呼ばれる (Hattori, S., 1954; Lei, 2016; Pierrehumbert, J. B., Beckman, M. E., 1988)。

近年、日本語学習者用に任意の入力文を東京方言で読上げる際に必要なピッチ制御を表示する、Online Japanese Accent Dictionary (OJAD) と呼ばれる web 教材が運用されている (峯松ほか, 2013)。その機能の一つに韻律読み上げチュータ「スズキクン」がある。これは任意の文章に対し東京方言での発声に必要な韻律制御を可視化し、その通りに読み上げた手本となる合成音声を提供する機能である。スズキクン利用者の利用による発音の向上は、実験的に確認されている (峯松・平野・中村, 2016)。また弁論大会参加者の練習にも利用されるツールになっている。OJAD では、ピッチ制御をシンボルを用いて離散的に表示するのではなく、入力テキストに対する自動解析結果のみを利用し、仮想的なピッチパターンに対応する曲線を用いて示している。OJAD におけるピッチパターン表示は異音的変動を考慮せず、語頭の 2 モーラで常にピッチが明確に上昇するカーブを描く。本研究ではこのピッチパターン表示をより実際の現象に近づけることを念頭に置き、任意の文入力における異音的変動に対応できるよう、各種文脈における語頭 2 モーラにおける F_0 上昇について音響分析を行なった。

2 分析実験

第 2 モーラが長母音の二拍目または撥音である単語の、語頭での F_0 上昇の変化を分析するために実験を行った。実験は、韻律的なコンテキストを考慮するために、該当単語の直前が無音、L トーン、H トーンの 3 つの場合で行った。

注目する単語が含まれる文と、比較対象となる単語が含まれる文をそれぞれ読み上げた音声を取録し、音声の F_0 のモーラごとの代表値を求め、語頭でのアクセント上昇を F_0 の変化として定量的に分析した。更に注目単語の直前が無音の場合だけでなく、L である場合と H である場合についても音声を録音し、同様の評価を行った。これは (Lei, 2016) では、単独での発声のみに言及していたが、それに加えて、文中に置かれた時の発声での変化を観察するためである。このように注目単語の直前として、無音・L・H の 3 パターンを想定し、これらを以下ではコンテキストと呼ぶ。

表 1: 実験文リスト。1 モーラ目の母音が異なる 2 文のペアを用意した。

	a	i	u	e	o
2モーラ目が通常拍	ラムネを見つける	千鳥ヶ淵を見学する	熊本に向かう	セミプロになりました	駒沢大学を見学する
2モーラ目が長音	ラー油を見つける	地域おこしを見学する	空港に向かう	成人になりました	工業大学を見学する
	a	i	u	e	o
2モーラ目が通常拍	窓枠を見つける	締め出しをする	ブルジョアを訪ねる	ゲレンデが見える	小麦粉を破棄する
2モーラ目が撥音	万力を見つける	申告をする	文豪を訪ねる	限界が見える	婚約を破棄する

2.1 実験文

第2モーラが長母音の二拍目である単語と撥音である単語それぞれについて、実験文を以下のように構成した。まず、上記の条件を満たす単語から第1モーラの母音がそれぞれア・イ・ウ・エ・オになるように計5単語を選出した。さらに比較対象の単語として、これらの単語と同じモーラから始まり、単語長とアクセント型が同じである単語の内、第2モーラが有声音となる単語を比較対象として選出した。このように5ペアの単語を選びだし、各ペアの単語を文頭に持ち後続の部分は同一である文のペアを5ペア構成した。これらの5ペアの文に「無事 (LH)」と「今朝 (HL)」を文頭に置いた場合を加え、3種のコンテキストを反映した文を実験文とした。実験文は総計 $2 \times 5 \times 3 = 30$ ペアとなる。表1に文頭に何も置かない場合の実験文を示す。

2.2 実験データ

東京方言話者である20代男性5名が実験文を読み上げた音声を実験データとした。読み上げの際には、なるべく自身にとって読みやすいスピードで、かつ文ごとにスピードを変えないように指示した。被験者とは事前に個々の単語について正しいアクセントを確認したが、例となる音声は聴かせなかった。また、「無事」や「今朝」が文頭に付く場合に、ポーズを置かないように指示をした。

録音はサンプリング周波数48kHz、量子化bit数が16bit、モノラルで行い、これを16kHzにダウンサンプリングしたものを利用した。データ数は、60文を5人が読み上げたため300発声である。

2.3 分析手法

ピッチの上昇量を定量的に比較する為に、実音声から $\log F_0$ の上昇量を求めた。 F_0 抽出にはSPTK (SPTK, 2017) を利用した。SPTKでの F_0 抽出は、SWIPEアルゴリズムを用い、フレームシフト80、SWIPE閾値0.3、最小 F_0 を50Hz、最大 F_0 を400Hzとして行った。

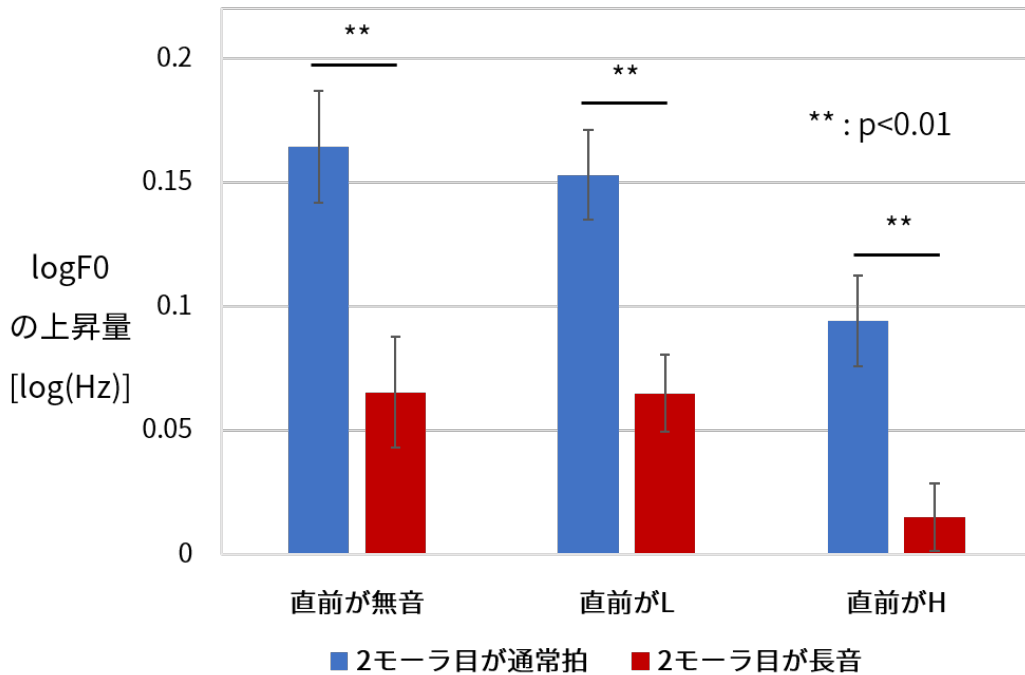


図 1: 長母音の有無と F_0 上昇量の関係 (エラーバーは標準誤差を表す)。

また、Julius (Lee & Kawahara, 2009) の音素セグメンテーションキットによって音声と音素の強制アラインメントを取り、実験文を構成する音素に対応する時間情報を得た。時間情報を元に、文頭の単語、あるいは「無事」、「今朝」の後にある注目している単語について、第1モーラの母音部の開始時刻と第2モーラの母音部の終了時刻を求め、これら2時刻での $\log F_0$ の差分を計算した。 F_0 の検出に失敗した部分がある場合は、 F_0 が存在している時刻をモーラ内で探索し、それぞれ開始時と終了時に最も近い時点での値を採用した。

分析を行う際には、注目単語の第1モーラの母音で区別せずにデータを利用した。

3 実験結果

3.1 第2モーラが長母音の二拍目である単語

第2モーラが長母音の二拍目である単語についての分析結果を図1に示す。

平均値を見ると、直前が無音・L・Hのいずれのコンテキストであっても、第2モーラが長母音の単語はそうでない単語より $\log F_0$ の上昇量が有意に小さかった ($p < 0.01$)。

コンテキストが無音やLである場合、第2モーラが長母音の二拍目である単語の語頭での F_0 上昇は、抑えられるものの完全に消失してはいなかった。一方コンテキストがHであると、 F_0 の上昇はほぼ消失していた。コンテキストごとに見た場合、傾向としての差は存在したが5%の有意差は示されなかった。

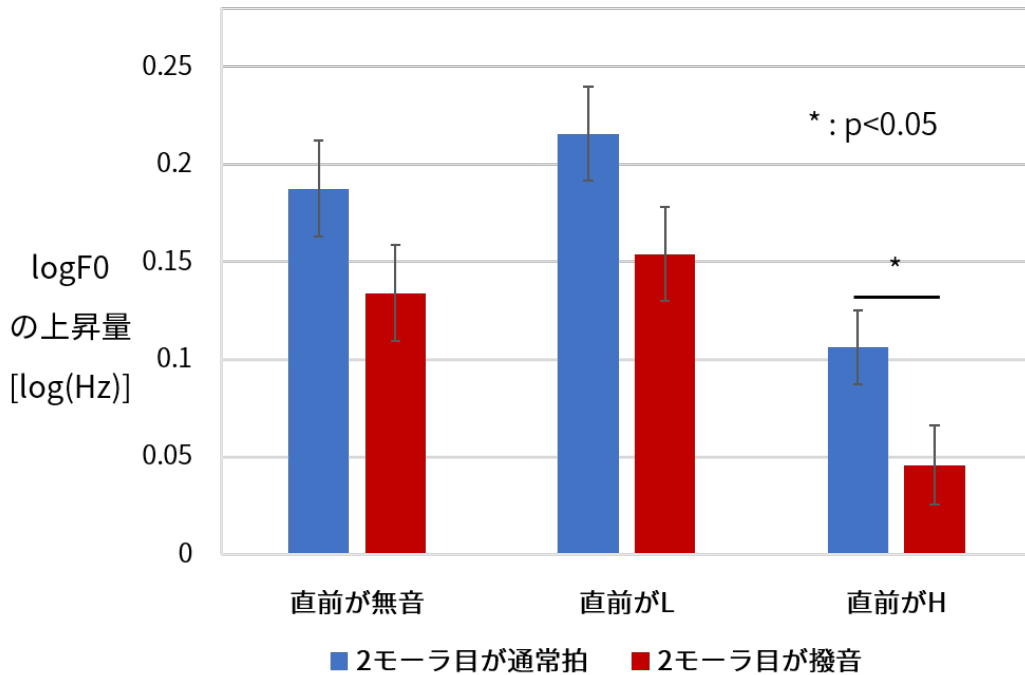


図 2: 撥音の有無と F_0 上昇量の関係 (エラーバーは標準誤差を表す)。

3.2 第 2 モーラが撥音である単語

第 2 モーラが撥音である単語についての分析結果を図 2 に示す。

平均値を見ると、 F_0 の上昇は抑えられている傾向にある。しかし分散が大きいため、直前が無音・L のコンテキストである場合には、5 % の有意差が示されなかった。一方、直前が H コンテキストである場合には、5 % の有意差が示された。

3.3 母音ごとの比較

第 2 モーラ目が撥音となっている単語について、 F_0 パターンを観察した。その結果、第 1 モーラの母音によって語頭のアクセント上昇の抑えられ方が変化する傾向にあることを見いだせた。第 1 モーラの母音ごとについての分析結果を図 3 に示す。

第 1 モーラの母音が「イ」である場合は、コンテキストに依らず $\log F_0$ の上昇が抑えられる傾向が見られる。また、第 1 モーラの母音が「オ」である場合は、単語の直前が無音か H コンテキストであると、上がり方が抑えられる傾向にある。

なお、いずれの場合も 5 % の有意差は示されていない。

4 考察

第 2 モーラが長母音の 2 拍目または撥音である単語について、全体として F_0 の上昇は抑えられる傾向にあった。

またコンテキストに注目すると、単語直前が H であると、語頭での F_0 上昇は他のコンテ

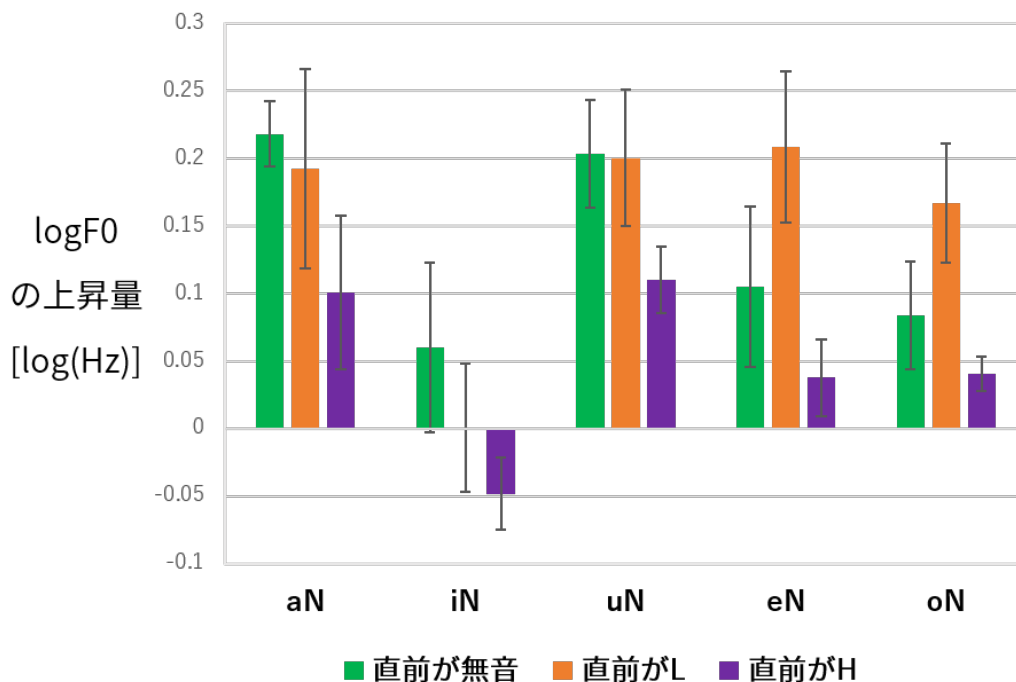


図 3: 撥音の直前にある母音と F_0 上昇量の関係 (エラーバーは標準誤差を表す)。

クスト下と比較して、大きく抑えられている。第 2 モーラが長母音の 2 拍目または撥音である単語の直前に、「無事 (LH)」のような H で終わる語が存在していると、当該単語は 1 モーラ目から 2 モーラ目へかけてのピッチの上がりがないように読まれる。つまり単純にアクセントを並べた場合に LH **L** HHH となる文が、LH **H** HHH と読まれていることを意味している。この傾向は第 2 モーラが長母音の 2 拍目である単語では特に顕著で、語頭での F_0 上昇はほぼ消失していた。

また、単語の 2 モーラ目が撥音であることの語頭でのアクセント上昇への影響は、1 モーラ目の母音によって差が生じる可能性が示唆された。第 1 モーラの母音で区別せずに分析を行った場合には有意差は示されなかったが、データを十分な数用意して第 1 モーラの母音ごとに分析を行えば、第 1 モーラの母音によっては明確な有意差が示される可能性がある。

本実験では、 F_0 上昇量の分散が大きくなってしまっていた。その原因の 1 つにアクセントの混在が考えられる。被験者には「無事ラムネを見つける。」といった「無事 (LH)」から始まる文を読み上げてもらった。しかし、被験者に何も指示をせずにこれらの文を発声させると、LH **H** HHH…と読む被験者と LH **L** HHH…と読む被験者が混在していた。「ラムネを」という部分のを単独で読み上げると、そのアクセントは LHHH である。しかしこれを「無事」という語の後に続けて読んだ場合に、「ラムネを」の語頭が L になるかどうかは、話速などの各個人の発話スタイルに依存する。そのため、アクセントに差異が生じる。

これらの差異を区別していないことが、ノイズになっている可能性がある。ただ、これらの差異をコントロールするためにどちらかのアクセントで読むように指示をすると、発声の自然性が失われる可能性が生じる。そのため、特に被験者には指示を出さずに読んでもらい、後でどちらのアクセントで読まれているかをラベリングし、データを分けて分析をすべきだと考えられる。

5 おわりに

本研究では、各種文脈下における語頭2モーラでの F_0 上昇に関する音響分析を行った。分析実験の結果、第2モーラが長母音の二拍目または撥音である場合には語頭での F_0 の上昇が抑えられる、異音的変動と呼ばれる現象が実際に起きていることが確認できた。また、この異音的変動は、単語直前のアクセントの影響を受けていることを実験的に検証した。この帰結より、OJAD のスズキクンにおけるピッチパターンをより現実のものに近づけるためには、異音的変動を考慮して、特定の文脈・単語において語頭2モーラでのピッチの上昇を緩やかにして提示するべきであると考えられる。

参考文献

- Hattori, S (1954) "The Japanese accent in a view of phonology." *Kokugo kenkyu* 2(2), 37-65.
- Lee, A., & Kawahara, T. (2009) "Recent development of open-source speech recognition engine julius." In *Proceedings: APSIPA ASC 2009: Asia-Pacific Signal and Information Processing Association, 2009 Annual Summit and Conference* (pp. 131-137).
- Lei, H.-Y. (2016) "An acoustic analysis on the allotonic variation of the initial rise in Tokyo Japanese in native speakers and learners." In *Proc. International Symposium on Applied Phonetics*.
- Pierrehumbert, J. B., Beckman, M. E. (1988) *Japanese Tone Structure*. The MIT Press.
- Speech Signal Processing Toolkit (SPTK)*. (2017). <http://sp-tk.sourceforge.net/>. (最終アクセス日 2018 年 1 月 27 日).
- 峯松 信明・中村新芽・鈴木雅之・平野宏子・中川千恵子・中村則子・田川恭識・広瀬啓吉・橋本浩弥 (2013) 「日本語アクセント・イントネーションの教育・学習を支援するオンラインインフラストラクチャの構築とその評価」 『電子情報通信学会論文誌 D』 96: 10, 2496-2508.
- 峯松 信明・平野 宏子・中村 則子 (2016) 「OJAD を用いた日本語韻律トレーニングによる自然性向上に関する実験的検証」 『電子情報通信学会技術研究報告音声 (SP)』 115: 392, 13-18.