



日本語破裂音の有声性弁別について一音響的・知覚的特徴を中心に一

瞿, 琦

(Degree)

博士 (学術)

(Date of Degree)

2022-03-25

(Date of Publication)

2023-03-01

(Resource Type)

doctoral thesis

(Report Number)

甲第8234号

(URL)

<https://hdl.handle.net/20.500.14094/D1008234>

※ 当コンテンツは神戸大学の学術成果です。無断複製・不正使用等を禁じます。著作権法で認められている範囲内で、適切にご利用ください。



博 士 論 文

日本語破裂音の有声性弁別について

—音響的・知覚的特徴を中心に—

Experimental Confirmation of Japanese Voicing
Contrasts of Stop Consonants: Focusing on their
Acoustic and Perceptual Characteristics

令和 4 年 1 月

神戸大学大学院国際文化学研究科

瞿 琦

要 旨

本研究は、これまで表立って議論されてこなかった「日本語の文末の無声破裂音を含む音節を切り出すと、日本語母語話者が子音部を無声子音ではなく、有声子音と知覚する」という現象に着目し、それに関与している要因について考察したものである。本研究では、日本語母語話者が日本語の語中から切り出した「無声子音+母音」を「有声子音+母音」に聴こえるのかどうかについて検証する。「有声子音+母音」と知覚する場合、日本語母語話者の破裂音の知覚にどのような音響的パラメータが影響を与えているのかを明らかにするため、「無声音として発声した無声無気音」と「有声音として発声した無声無気音」の音声的特徴に焦点を当て、5つの知覚実験を行った。

本論文は、序論から結論まで、全8章から構成されている。第1章では、本研究の目的について述べ、第2章では日本語の破裂音に関する先行研究を概観している。

第3章の知覚実験Iでは、2モーラの無意味の録音語から第2音節の「無声子音+母音」を切り出し、日本語母語話者がその音節頭の無声子音を有声子音と知覚するかどうかについて検証した。録音語のアクセント型は頭高型と平板型であり、刺激音は録音語から切り出した高ピッチと低ピッチの第2音節の単独音である。知覚実験の結果、日本語母語話者が切り出した音節の子音部を無声子音[p]ではなく有声子音[b]と知覚することが確認できた。しかし、この現象について、低ピッチの場合には、子音部[p]が有声子音[b]と知覚されたものの、高ピッチの場合には、子音部が高い比率で無声音と知覚されることが明らかとなった。また、知覚実験Iで用いた2種類の刺激音の間で音響的特徴(VOT・後続母音長・後続母音の*f₀*)が異なっていたため、どのような音響的パラメータがこの現象に影響を与えているのかは分からない。そこで、知覚実験II, III, IVでは、知覚実験Iで用いた語中から切り出した2種類(高ピッチ・低ピッチ)の「無声子音+母音」をベースとして加工した。

第4章の知覚実験IIではVOTを、第5章の知覚実験IIIでは後続母音長を、それぞれ操作した加工音声を用いて追実験を行った。その結果、知覚実験Iと同様に、高ピッチの刺激音はVOTと後続母音の長短に関わらず無声音と知覚され、低ピッチの刺激音もVOTと後続母音の長短に関わらず依然として有声音と知覚された。日本語母語話者の破裂音の知覚に変化が見られなかったことで、有声・無声の弁別にVOTと後続母音長という要因は

関わっていないことが明らかとなった。第 6 章の知覚実験 IV では、刺激音に対し子音部の後続母音の *f* を加工し、後続母音の *f* が日本語母語話者の知覚に影響を及ぼすのかを検証した。その結果、低ピッチの場合は、後続母音の *f* が日本語母語話者の有声・無声の弁別に関与していることが確認できた。

なお、知覚実験 I から IV において、「無声音として発声した無声無気音」に焦点を当て、日本語母語話者が有声性の弁別に VOT やその後続母音長は積極的に関わっていないことが確認され、後続母音の *f* に関しては「条件付き」で有声性の弁別に寄与することが明らかとなった。しかし、知覚実験 IV では、高ピッチの刺激音において後続母音の *f* の影響が確認できなかったため、それ以外にも他の要因があると考えられる。1 つの可能性としては、無声音として発声した「無声子音+母音」を語中から語頭に移動したため、日本語母語話者がその単独の音節を有声音と知覚したことが考えられる。

そこで、第 7 章の知覚実験 V では、「有声音として発声した無声無気音」に焦点を当て、無声音化した語頭有声音を語中に移動し、日本語母語話者が無声音と知覚するかどうかについて検証した。その結果、高・低ピッチの刺激音ともに有声音に知覚された。さらに、反応時間も併せて測定することで、ピッチの違いにより処理時間に差があるかどうかについて検証した結果、低ピッチの刺激語の反応時間は高ピッチの刺激音より有意に長いことが確認された。

最後に、語中における低ピッチの「無声子音+母音」を語中から切り出すと、日本語母語話者は確かに有声子音として知覚し、高ピッチの場合が無声音として知覚した。加工タイプの間では有意な差は確認されなかったが、分散分析の結果、低ピッチと高ピッチともその音節の後続母音の *f* の変化により、有声音に知覚した音が無声音（低ピッチの場合）・無声音に知覚した音を有声音（高ピッチの場合）、として知覚する傾向が見られた。

本研究により、無声音として発声した無声無気音を日本語母語話者が知覚する際に一定の条件下で、ピッチの高低が母語話者の知覚に影響を及ぼしていることが明らかとなった。しかし、「有声音として発声した無声無気音」と「無声音として発声した無声無気音」の間には、現段階ではまだ分からない何らかの違いがあると考えている。破裂音のカテゴリ一区分にかかわる音響的パラメータについて、先行研究で既知の音響的特徴の他に何かあるのか、今後のさらなる検討が必要である。また、日本語破裂音の知覚的特徴を明らかにした上で、この特徴は日本語のみの現象なのか、普遍的な現象なのかについて、日本語と同じ無声無気音がある英語や中国語などの言語を対象に、さらなる調査を予定している。

Abstract

This study focuses on the perceptual phenomenon of how “devoicing word-medial plosives in Japanese influences Japanese native speakers’ perceptions by moving the devoicing plosives from the word-medial to the word-initial,” and explores whether and why this happens. The acoustic characteristics of devoicing word-medial plosives in Japanese are produced as unaspirated plosives; the same hold true for voicing word-initial plosives as they are almost always produced as unaspirated devoicing plosives in Japanese. Therefore, this study has prepared five perceptual experiments for perceptual decision tasks focusing on these two types of unaspirated devoicing plosives to identify the acoustic parameters used by native Japanese speakers, and the acoustic characteristic differences between these two types of plosives.

The study is divided into 8 chapters. Chapter 1 details the problem statement, purpose statement and research questions. Chapter 2 is a literature review of Japanese and Mandarin Chinese plosives’ acoustic characteristics along with perceptual characteristics of Japanese as a Foreign Language (JFL) / Japanese as a Second Language (JSL) learners’ and native Japanese speakers’ perceptions. Chapters 3 to 7 outline perceptual experiments I to V, respectively. The last chapter is the conclusion of the thesis and feedback for JFL and JSL education, and future tasks.

In Chapter 3, perceptual experiment I was set to confirm the above phenomenon. The type of words taking the devoicing plosive (i.e. [pa]) from 2-mora pseudowords (i.e. [apa]) were recorded from a native Japanese speaker. The pitch accents of the voice recordings were High Low (HL) and Low High (LH). The stimulus words are in the second mora of the recorded words and are divided into two groups by different pitch accent (High, H; Low, L). Twenty university students, all native Japanese speakers, heard each stimulus word. The result of the experiment I found that only L pitch group stimulus sounds [pa] are identified in voicing plosives [ba] after being moved from the word-medial to the word-initial. This may be due to one or more factors. The two groups of stimulus sounds had more than three distinguishing acoustic characteristics, and they also had significantly different voice onset time (VOT), different subsequent vowel strengths and different fundamental frequencies (f_0). Chapters 4 to 7 are additional experiments set to clarify which characteristics caused

the result.

In the experiments in chapters 4 to 6, five stimulus sounds are chosen randomly from the H/L groups in the first experiment, as original sounds. Then the VOT, subsequent vowel strength and *f₀*, are processed and used in experiment II Chapter 4, experiment III Chapter 5, and experiment IV Chapter 6, respectively. The results of experiments II and III were the same as in experiment I. The sound [pa] from H group could still be identified in [pa], and the sound [pa] from L group could be identified in [ba]. This argues that changing the duration of VOT and subsequent vowels' duration didn't influence the native Japanese speakers' perceptions. However, experiment IV Chapter 6, had different results. Changing the same original sounds of *f₀* from the first experiment, the stimulus sounds from H group had the same result as the experiments above, however, the stimulus sounds from L group had different results. By processing the original sounds of *f₀* to be higher or lower, the stimulus sounds from L group were able to be differentiated between [ba] and [pa]. Under these conditions, the results show *f₀* to be one of the characteristic features that influences perception of voicing contrast, but this only occurred in L pitch.

Changing the devoicing plosives from the word-medial to the word-initial influences a native Japanese speaker's perception as shown by the stimulus sounds from L group. Experiment V Chapter 7 explores whether processing the devoicing of word-initial plosives into word-medial plosives influences perception of voicing contrast by native Japanese speakers. The perception results were unanticipated, we found that moving plosives produced as unaspirated devoicing plosives from word-initial to word-medial had no effect. However, the reaction times for identifying the sounds from L group were significantly longer than that of H group.

This study highlights that in certain conditions native Japanese listeners hear stop consonants and use the perceptual characteristic of pitch to identify the voicing contrast, whether voiced or devoiced. The results of these five perceptual experiments are expected to prove that native Japanese speakers categorize these two types of devoicing plosive to recognize stop consonants by pitch. Furthermore, whether the phenomenon outlined in this study occurs in other languages is a desired topic for further research.

目次

要 旨.....	II
ABSTRACT	IV
第1章 序 論.....	1
1.1 はじめに.....	1
1.2 本研究の目的.....	4
1.3 研究課題と意義.....	5
1.3.1 知覚実験I.....	5
1.3.2 知覚実験II.....	5
1.3.3 知覚実験III.....	5
1.3.4 知覚実験IV.....	6
1.3.5 知覚実験V.....	6
1.4 研究の対象と研究方法.....	6
1.5 本論文の構成.....	7
第2章 先行研究.....	9
2.1 日本語・中国語破裂音の概要.....	9
2.2 日本語・中国語破裂音の音声的特徴.....	11
2.2.1 日本語破裂音の音声的特徴.....	11
2.2.2 中国語破裂音の音声的特徴.....	23
2.3 中国語話者による日本語破裂音の習得研究.....	25
2.4 日本語破裂音の知覚的特徴.....	27
第3章 知覚実験 I	29
3.1 実験の目的.....	29
3.2 録音の対象と方法.....	29

3.2.1	録音協力者.....	29
3.2.2	録音語.....	29
3.2.3	音声の録音.....	30
3.2.4	録音手順.....	31
3.3	刺激音の作成.....	32
3.3.1	刺激音における音響的特徴の比較.....	35
3.4	知覚実験の方法.....	36
3.5	結果と考察.....	37
3.6	まとめ.....	38
第4章	知覚実験 II : VOT の影響.....	41
4.1	実験の背景と目的.....	41
4.2	刺激音の作成.....	41
4.3	知覚実験の方法.....	45
4.4	実験の結果.....	46
4.5	考察とまとめ.....	47
第5章	知覚実験 III : 後続母音長の影響.....	49
5.1	実験の背景と目的.....	49
5.2	刺激音の作成.....	49
5.3	実験方法.....	54
5.4	実験の結果.....	55
5.5	考察とまとめ.....	56
第6章	知覚実験 IV : ピッチの影響.....	58
6.1	実験の背景と目的.....	58
6.2	刺激音の作成.....	58
6.3	実験の結果.....	64

6.4	考察とまとめ	67
第7章	知覚実験 V：語中位置の影響.....	71
7.1	実験の背景と目的	71
7.2	実験材料の作成方法	72
7.2.1	録音協力者	72
7.2.2	録音語	72
7.2.3	音声の録音	72
7.2.4	録音手順	73
7.3	刺激音の作成	77
7.4	知覚実験の方法	80
7.5	実験の結果	81
7.5.1	知覚率の結果	81
7.5.2	反応時間の結果	82
7.6	考察	83
7.6.1	語中位置（語頭・語中）の影響	84
7.6.2	閉鎖区間（内破）の持続時間	85
7.6.3	ピッチの影響	86
7.7	まとめ	86
第8章	結論	89
8.1	本研究のまとめ	89
8.2	教育現場への提言	91
8.3	今後の課題	92
	参考文献	94
	謝辞	98

図リスト

図 1	破裂音の発音の調音方法 [斎藤 (2009) p.21 より作成]	9
図 2	有声破裂音の判定例	13
図 3	SIMULTANEOUS VOICING の無声無気音の判定例 (VOT=0)	14
図 4	プラスの値を取る無声無気音の判定例 (VOT=7 MSEC)	14
図 5	無声有気音の判定例	15
図 6	語頭無声音化の有声破裂音の判定例「バナ」	17
図 7	語中無声音についての観察例「イト」	18
図 8	語中有声破裂音についての観察例：パターン①	19
図 9	語中有声破裂音についての観察例：パターン②	20
図 10	語中有声破裂音についての観察例：パターン③	21
図 11	発話語読み上げリスト	31
図 12	高ピッチ「パ」の具体例	32
図 13	低ピッチ「パ」の具体例	32
図 14	高ピッチ「パ」の具体例	33
図 15	低ピッチ「パ」の具体例	33
図 16	高ピッチ「パ」の具体例	33
図 17	低ピッチ「パ」の具体例	33
図 18	高ピッチ「パ」の具体例	34
図 19	低ピッチ「パ」の具体例	34
図 20	高ピッチ「パ」の具体例	34
図 21	低ピッチ「パ」の具体例	34
図 22	実験の流れ	37
図 23	ピッチ別の「パ」を無声音とした知覚率 (%)	37
図 24	オリジナル音声の音声波形・スペクトログラム (VOT=13.7 MSEC)	43
図 25	VOT+3 倍の音声波形・スペクトログラム (VOT=40.0 MSEC)	43
図 26	VOT+6 倍の音声波形・スペクトログラム (VOT=81.3 MSEC)	44
図 27	VOT-3 倍の音声波形・スペクトログラム (VOT=4.5 MSEC)	44

図 28	VOT-6 倍の音声波形・スペクトログラム (VOT=2.2 MSEC)	45
図 29	高ピッチ「パ」の無声音知覚率 (%)	46
図 30	低ピッチ「パ」の無声音知覚率 (%)	47
図 31	後続母音区間の測定	50
図 32	オリジナル音声の音声波形・スペクトログラム (後続母音長: 143.888 MSEC)	52
図 33	後続母音+2 倍の音声波形・スペクトログラム (後続母音長: 288.162 MSEC)	52
図 34	後続母音+3 倍の音声波形・スペクトログラム (後続母音長: 409.107 MSEC)	53
図 35	後続母音-2 倍の音声波形・スペクトログラム (後続母音長: 87.674 MSEC)	53
図 36	後続母音-3 倍の音声波形・スペクトログラム (後続母音長: 62.3MSEC)	54
図 37	高ピッチ「パ」の無声音知覚率 (%)	55
図 38	低ピッチ「パ」の無声音知覚率 (%)	55
図 39	各加工タイプのピッチ曲線の具体例	62
図 40	各加工タイプのピッチ曲線の具体例	63
図 41	無声音の知覚率	65
図 42	高ピッチ「パ」の無声音知覚率	65
図 43	低ピッチ「パ」の無声音知覚率	66
図 44	発話語読み上げリスト (高ピッチの場合)	74
図 45	発話語読み上げリスト (低ピッチの場合) ①	75
図 46	発話語読み上げリスト (低ピッチの場合) ②	76
図 47	刺激語の音声波形・スペクトログラム	78
図 48	日本語母語話者の知覚率 (%)	82
図 49	日本語母語話者の平均反応時間 (MSEC)	83

表リスト

表 1	先行研究における日本語の語中破裂音の弁別パラメータ	3
表 2	日本語・中国語の破裂音	10
表 3	日本語語頭破裂音における後続母音の <i>F₀</i> の平均値 (Hz)	12
表 4	語頭と語中破裂音の VOT の平均値 (MSEC)	22
表 5	日本語・中国語における語頭破裂音の VOT の平均値と最小値・最大値 (MSEC) ..	24
表 6	先行研究結果のまとめ	27
表 7	録音語内容	30
表 8	刺激音の音響的特徴	35
表 9	知覚実験 I の結果と刺激音の音声特徴	40
表 10	無声音と知覚した「パ」(H) の VOT 加工結果 (単位: MSEC)	42
表 11	有声音と知覚した「パ」(L) の VOT 加工結果 (単位: MSEC)	42
表 12	無声音と知覚した「パ」(H) の母音持続時間の加工結果 (単位: MSEC)	51
表 13	有声音と知覚した「パ」(L) の母音持続時間の加工結果 (単位: MSEC)	51
表 14	無声音と知覚した「パ」(H) の後続母音の <i>F₀</i> の加工結果	60
表 15	有声音と知覚した「パ」(L) の後続母音の <i>F₀</i> の加工結果	61
表 16	加工倍率が異なる刺激音の知覚結果の多重比較 (低ピッチ)	67
表 17	高・低ピッチの刺激音の加工タイプ毎の後続母音の <i>F₀</i> 平均値	69
表 18	高ピッチ刺激音の音響特徴	79
表 19	低ピッチ刺激音の音響特徴	79
表 20	刺激語の音声継続時間と平均 (単位: MSEC)	80
表 21	知覚率と反応時間における語中位置の影響	84
表 22	知覚実験 I と知覚実験 V の結果のまとめ	87

第1章 序論

1.1 はじめに

日本語では、破裂音や破擦音¹において声帯振動の有無により有声音と無声音の対立がある。特に日本語の破裂音は、第二言語習得研究の分野において、習得困難な項目の1つとして多くの研究の対象とされてきた。言語習得において音声・音韻は母語干渉が最も顕著に現れる分野であると言われており、日本語の破裂音の習得にも母語の影響が見られる。特に中国語話者による日本語破裂音の習得については数多くの研究成果が報告されている（杉藤・神田 1987, 清水 1993, 1999, 2018, 朱 1994, 2010, 福岡 1995, 杉藤 1996, 山本 2004, 2009, 劉 2005 など）。これは、母語（L1）と目標言語（L2）における音韻体系の違いが主な原因だと考えられている。

日本語の破裂音の音韻体系は、帯気性による無声有気音（ㄆの/p^h/, ㄊの/t^h/, ㄎの/k^h/）・無声無気音（ㄆの/p/, ㄊの/t/, ㄎの/k/）の対立を持つ中国語に対し、声帯振動の有無による無声音（パ行の/p/, タ行の/t/, カ行の/k/）と有声音（バ行の/b/, ダ行の/d/, ガ行の/g/）の2項対立となっている。すなわち、中国語系話者が日本語の有声破裂音と無声破裂音の知覚と生成の習得が困難である要因として、日本語では有声性対立を持つのにに対し、中国語では帯気性対立を持つことであり、2つの言語の間に差異があるからである。

これまでに行われてきた日本語破裂音の研究は日本語学習者を対象としたものが中心であり、日本語母語話者のみを対象とした研究は2000年代以前盛んであったが、それ以降わずかしか行われていない。さらに、筆者の指導教員である朱春躍教授は授業やゼミで、日本語母語話者が発話する「そうですか。²」の「か」を文脈から切り出して単独で母語話者に聴かせると、発話者本人も含め、それを「か」ではなく「が」として知覚する、という現象を確認している。

そこで本論では、この現象に注目し、日本語母語話者を対象にして日本語破裂音の知覚に関する特徴を考察する。

まず音響的には、日本語を含む多くの言語においては、有声性の対立を捉える際の特徴

¹ 日本語の破擦音は、歯茎破擦音と前部硬口蓋破擦音の2種類で、声帯振動の有無により有声音と無声音に分かれる。歯茎破擦音の場合は、有声音が「ず」([dz])、無声音が「つ」([ts])であり、前部硬口蓋破擦音の場合は、有声音が「じ」([dʒ])、無声音が「ち」([tʃ])という音として現れる。

² ここでの「そうですか。」とは、相手の話や意見に賛同する場合の下降調のことである。

として、声帯振動の開始時間（Voice Onset Time, VOT, Lisker and Abramson 1964）が挙げられる。VOT とは、声道閉鎖の開放から声帯振動が始まるまでの時間差であり、声道閉鎖の開放前に声帯振動が始まる場合はマイナス（-）の値で、声道閉鎖の開放後に声帯振動が始まる場合はプラス（+）の値で表される（Lisker and Abramson 1964）。日本語では、有声破裂音がマイナスの、無声破裂音がプラスの値をとると言われてきたが（Homma 1980, 清水 1993, 杉藤 1996 など）、日本語の語頭有声音においては、VOT がマイナスではなく、プラスの値を取ることがあるという（大友・佐藤・高倉 1957, Homma 1980, 杉藤 1996）。高田の一連の研究（2004, 2006, 2008, 2011）によると、これには地域差と世代差が関わっているようであり、地域差に関しては、東北地方において VOT がプラスの値を取ることが観察され、世代差に関しては、若い世代ほど VOT がプラスの値を取ることが観察されている。このように日本語の単独または語頭の有声音は、声帯振動が破裂の瞬間よりも前に始まるようなもの（-VOT）から、声帯振動が破裂の瞬間よりも後に始まるように（+VOT）変化しつつあることが報告されている（高田 2004）。

また、VOT の観点からみると、無声音化した語頭有声音はプラスの VOT 値を取るだけでなく、その持続時間も短いことが報告されている（朱 2010）。しかし、日本語母語話者においては、無声音化した語頭有声音を問題なく有声音として知覚できることが明らかとなっている（高田 2004, 朱 2010, 瞿 2021）。つまり、有声音が無声音化するか否かということは、ある条件の下では（単独、または語頭）日本語母語話者の有声・無声の弁別に大きく関わっていないことを示唆している。このことを踏まえ、本研究では、日本語母語話者を対象として、VOT 以外の音響的特徴も考慮に入れ、日本語破裂音の知覚を詳細に分析する。

VOT の他には、日本語の語中有声・無声破裂音の音声的特徴として、後続母音の持続時間、閉鎖区間、呼気流量などが挙げられている（Homma 1973, 1981, 杉藤・神田 1987, Shimizu 1996, 鄭・桐谷 1998, 朱 2010）。さらに、日本語の破裂音の音声的特徴については、語中位置（語頭・語中）によって大きく異なることが分かっている。本研究では、1.1 で呈示した「そうですか。」という例のように、語中の無声破裂音を対象にして、日本語母語話者の知覚に関する特徴を考察する。表 1 は日本語破裂音の語中の有声音・無声音の弁別パラメータについて、いくつかの先行研究の結果をまとめたものである。この表からわか

るように、語中の有聲破裂音と無聲破裂音では、無聲破裂音の方が VOT（外破³の持続時間）と、閉鎖区間の持続時間（内破⁴）が長いことが知られている（Homma 1981, 杉藤・神田 1987, 朱 1994, 2010）。また、語中の破裂音の違いは、前後の母音の持続時間にも両方の破裂音から影響を受けることがあり、Homma（1981）は、無声子音に隣接する方が有声子音に隣接するより短くなると述べている。

表 1 先行研究における日本語の語中破裂音の弁別パラメータ

パラメータ 文献	語中の有声音	語中の無声音
① VOT 杉藤・神田（1987） 朱（1994, 2010）など	短い 15～20 msec 以下	長い 15～20 msec 以上
② 後続母音長 Homma（1973, 1981）など	長い	短い
③ 閉鎖区間の持続時間 Homma（1981） 杉藤・神田（1987） 朱（1994, 2010）など	短い 100 msec 以下	長い 100 msec 以上

以上のように、日本語の語中の有声音と無声音は各弁別的パラメータにおいて相違があ

³ ここでの「外破」とは、「舌が硬口蓋に接触して破裂音を作ってから後続の母音が始まるまでの部分」（杉藤・神田 1987:6）のことである。すなわち、語中の場合、破裂から母音が始まるまでの時間が外破の区間である。

⁴ ここでの「内破」とは、杉藤（1996）の用語で、「子音の閉鎖区間」のことであり、「語中に来る子音の破裂前の口腔の閉鎖による無音区間」（朱 2010:7）のことをいう。

ることが指摘されている。しかし、これらのパラメータにおいて、有声音・無声音の音声的特徴が挙げられたとしても、前述の例のような語中に無声音を含む「無声子音+母音」の音節を文脈から切り出すと、その音節が「有声音+母音」に聴こえるという現象を解釈することはできない。そこで本研究では、語中の無気音のような無声音、つまり「無声音として発声された無声無気音」を、先行研究で指摘された語頭の無声音化した有声音（すなわち「有声音として発声された無声無気音」）と比較検討する。

既述したように、日本語破裂音の研究では、これまでは主に中国語話者を含む日本語学習者を対象としたものがほとんどであった。これは日本語破裂音の習得には母語の影響が顕著に現れることが分かっているからである。しかし、日本語母語話者が発話した日本語破裂音の知覚的特徴を分析することで、有声性の弁別に関わっている音響的特徴はなんであるかが明らかになるのではないかと考えられる。こうして得られた研究結果は、日本語学習者に対する音声教育にも重要な知見となるであろう。

1.2 本研究の目的

日本語の破裂音に関する研究は、日本語音声学、日本語教育学、社会言語学、音響学など、様々な分野の研究者が有声音・無声音破裂音の実態や知覚・生成の特徴を取り上げた研究が多く行われてきた。しかし、日本語破裂音の音声的特徴は変化しつつあり、「破裂前の声帯振動の有無」から、「呼気の強さや aspiration の有無」など、音韻論的に有声性対立が存在していない中国語などの言語のような有気・無気の対立に類似してきたことにより、日本語の「無声音として発声した無声無気音」と「有声音として発声した無声無気音」の相違が日本語においてますます重要性を持つようになったと言える。日本語の「有声音として発声した無声無気音」、すなわち語頭の無声音化した有声音に関する現象は、高田の一連の研究（2004, 2006, 2008, 2011）により、さまざまな面が明らかとなった。その後、この現象を中心に、日本語学習者の知覚に関する習得研究も増えてきた（金 2014, 邊 2019）。しかし、日本語母語話者における無声音化した語頭有声音と語中の無声音の有声性弁別についての研究はほとんど見られないのが現状である。

以上に述べた先行研究の結果を踏まえ、本研究は日本語母語話者が日本語の語中から切り出した無声音を含む音節が有声音に聴こえる現象を中心に、その知覚に与える要因として、「無声音として発声した無声無気音」と「有声音として発声した無声無気音」の音声的特徴に焦点を当て分析を行う。

本研究の目的は以下の3点である。日本語母語話者が日本語の無声無気音を知覚する際、

- (1) 語中の無声音を文脈から切り出すと有声音として知覚するかどうか。
- (2) (1) が確認できた場合、無声音を有声音と知覚する要因として、その音響的パラメータである①VOT や②後続母音長、③後続母音の *f₀* がどのように関与しているか。
- (3) もしも (1) が確認できて、(2) は確認できなかった場合、語中位置 (語頭・語中) がどのような影響を与えているか。

1.3 研究課題と意義

本研究は日本語母語話者を対象にした日本語無声無気破裂音の知覚と音響特徴の関連についての研究であり、どんな状況で母語話者が無声音を有声音と知覚するか、また知覚に影響を及ぼす要因は何かを探ることを目的としている。上記の研究目的を達成するために行った知覚実験の研究課題は以下の通りである。

1.3.1 知覚実験 I

知覚実験 I の研究課題は、指導教員である朱春躍教授が提案した例を参考に、「無声子音+母音」の音節を語中から切り出すと、日本語母語話者は無声子音を有声子音に聴こえるかどうかという現象を確認することである。

1.3.2 知覚実験 II

知覚実験IIの研究課題は、日本語母語話者が語中から切り出した無声子音を含む音節を知覚する際、VOT の長さの伸縮加工がどのような影響を与えるかについて検証することである。

実験では、知覚実験 I で用いた刺激音をベースとし、破裂音の VOT の長短の操作により、日本語母語話者の破裂音の知覚に変化が見られるかを調べる。

1.3.3 知覚実験 III

知覚実験IIIの研究課題は、日本語母語話者が語中から切り出した無声子音を含む音節を知覚する際、後続母音の長さの伸縮加工がどのような影響を与えるかについて検証することである。

実験では、知覚実験 II と同様に知覚実験 I で用いた刺激音をベースとし、破裂音の後続

母音の長短の操作により、日本語母語話者の破裂音の知覚に変化が見られるかを調べる。

1.3.4 知覚実験 IV

知覚実験IVの研究課題は、日本語母語話者が語中から切り出した無声子音を含む音節を知覚する際、破裂音の後続母音の基本周波数（fundamental frequency, f_0 ）の高低加工がどのような影響を与えるかについて検証することである。

実験では、前述の知覚実験 II と III と同様に知覚実験 I で用いた刺激音をベースとし、破裂音の後続母音の f_0 のシフトにより、日本語母語話者の破裂音の知覚に変化が見られるかを調べる。

1.3.5 知覚実験 V

知覚実験 I~IV の研究課題において、語中の「無声子音+母音」の語中位置を語中から語頭に移動すると、日本語母語話者がその音節の頭子音を有声音と知覚するかどうか分析し、さらに有声音と知覚する要因を VOT・後続母音長・後続母音 f_0 の計 3 種類の弁別的パラメータをそれぞれに変動して知覚実験を行った。知覚実験 V では、無声音化した有声音を語頭から語中に移動すると、日本語母語話者の知覚に知覚実験 I と同じ影響を与えるかどうかについて検証する。

実験では、語中の無声音と同じ音響的特徴を持つ「有声音として発音された語頭の無声無気音」、すなわち無声音化した語頭有声音を、語頭から語中に移動すると、無声子音に聴こえるかどうか分析する。

日本語の破裂音については、従来の多くの研究は有聲・無声に関する研究であり、無声無気音に関する研究は少なかった。以上の課題を明らかにすることで、日本語母語話者の自然発話中に現れる無声無気音の知覚に寄与する要因が明確になり、日本語破裂音の習得に困難を感じる中国人や韓国人学習者のための音声教育により効率的な音声習得の方法を提示することができると考えられる。

1.4 研究の対象と研究方法

本研究で行った調査は知覚調査であり、具体的やり方は、日本語母語話者が発声した語中の「無声音として発声した無声無気音」を、その「無声子音+母音」の音節を語中から切り出すと無声子音が有声音に聴こえるかどうかを確認した上、その無声子音が有声音に知覚する際の特徴をより明確に捉えるため、先行研究で指摘されていた複数のパラメータが弁別への影響について聴取実験を用いて確認した。さらに、音響的特徴が「無声音

として発声した語中の無声無気音」と類似している「有声音として発声した無声無気音」、つまり無声音化した語頭有声音との比較を行い、語中位置（語頭・語中）が有聲・無声の知覚に影響を与えるかどうかを検証した。研究方法としては、コンピュータのモニター画面の中央に提示された仮名表記の文字と、ヘッドフォンから提示される刺激音音声と一致であるかどうかを判断する課題を用いて行った。

なお本論文では、論述の便宜上、語中から切り出した「無声子音+母音」を含む音節のことを「語中から切り出した無声音」と称することもある。同様に、「無声音化した語頭有声音+母音」の音節のことを「無声音化した語頭有声音」と称することもある。

1.5 本論文の構成

本論文は全8章から構成されている。

まず、第1章（本章）では、研究の目的と意義について述べる。第2章では、破裂音に関する既存の研究成果をまとめ、言語学的な立場から、日本語と中国語における破裂音の音韻体系と音声的特徴について言及し、それぞれにおいて知覚的・音響的側面の究明に努めてきた先行研究を概観する。

第3章では、2拍の無意味語（例えば、「アパ」など）から無声音が含まれる語中の音節を切り出して作成した刺激音（例えば、「パ」など）を用いた知覚実験を日本語母語話者20名に対して行った結果を報告し、その結果が何を意味するかを考察する。また、第3章の結果を受けて、破裂音の有声性知覚にかかわる音響的特徴を明らかにするために、第4章、第5章、第6章、第7章の追実験を行い、日本語の有声・無声破裂音の弁別特徴に関する実態調査を行なった。

第4章、第5章、第6章においては、第3章の実験で使用した「文中から切り出した無声音の刺激音」2種類（高ピッチ・低ピッチ）の無声音に対し、さらにある種特定の音響的パラメータを加工して作成した刺激音を用いて知覚実験を行い、無声音を有声音と知覚する音響的パラメータの影響について考察する。それぞれの詳細内容について以下のようにある。第4章では「パ」のVOT、第5章では後続母音の持続時間を伸縮操作して作成した加工音声を刺激音とし、日本語母語話者10名に対して知覚実験を行った。第6章では、ピッチの影響を確認するため、2種類（高ピッチ・低ピッチ）の刺激音「パ」の後続母音の*f₀*をシフトし、日本語母語話者10名に対して知覚実験を行った。

さらに、音響的無声音が有声音に知覚される現象について、各パラメータ（VOT・後続母音長・後続母音の *f₀*）を変えて確かめた結果を踏まえ、第7章では、語中位置（語頭・語中）の知覚への影響を検証すべく知覚実験を行った。調査方法は、日本語語頭の無声音化した有声音を語中に移動し、日本語母語話者 20 名を対象に即時的正誤判断課題を実施した。最後に、第8章では、本研究の結論をまとめ、今後の課題について言及する。

第2章 先行研究

2.1 日本語・中国語破裂音の概要

子音とは、気流が口腔内でなんらかの阻害を受けて調音される音である。調音上の特徴として、声の有無、その阻害はどこで起こるか（調音位置, **place**）、どのように阻害されるか（調音方法, **manner**）という3つの基準によって定義される。例えば、[p]という子音は、無声音であり、両唇音であり、破裂音なので「無声両唇破裂音」と呼ばれる。破裂音（又は「閉鎖音」とも呼ぶ）は、閉鎖の形成、閉鎖の持続、そして閉鎖の開放（破裂）という3つの段階を経て生成される。

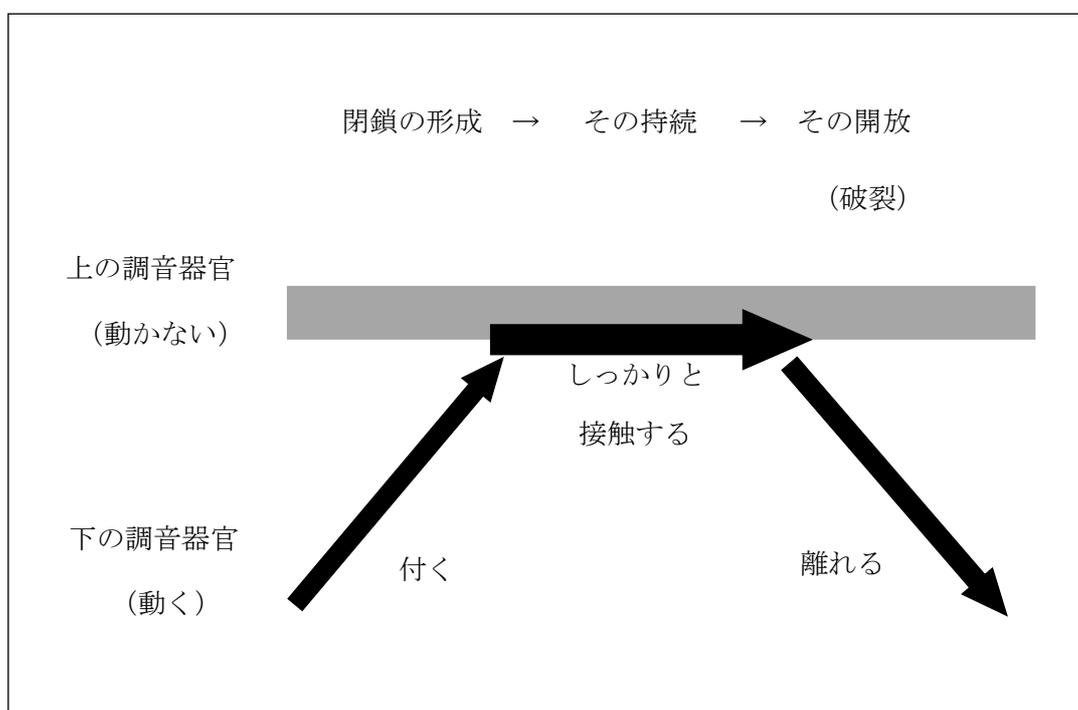


図1 破裂音の発音の調音方法[齋藤(2009)p.21より作成]

図1に示すように、破裂音の生成過程については、まず、第1段階として、唇や舌などが動く下の調音器官と、動かない上の調音器官である歯茎、上唇、口蓋（軟口蓋、硬口蓋）などに接触させることで、気流が逃げないように口腔内の声道を完全に閉鎖する。次に、第2段階として、その声道の閉鎖状態を持続させ、空気の流れを完全にストップさせることで口腔内部の気圧を高める。最後に、第3段階として、声道の閉鎖状態を一気に開放する

ことで瞬間的に空気の破裂を生じさせる。以上のような3つの段階を経て生成される音が破裂音である。

破裂音の調音位置に関しては、日本語、中国語ともに両唇音、歯茎音、軟口蓋音という3種類の子音を有している。日本語の場合、両唇音が/p/と/b/、歯茎音が/t/と/d/、軟口蓋音が/k/と/g/で対立しているのに対して、中国語の場合、両唇音が/p/と/p^h/、歯茎音が/t/と/t^h/、軟口蓋音が/k/と/k^h/の対立である(表2)。表2を見ると、それぞれにおいていかなる対立に基づく何種類の子音を持つかは、両言語間で大きく異なっていることが分かる。

表2 日本語・中国語の破裂音

		日本語			中国語		
無声音	有気	p	t	k	p ^h	t ^h	k ^h
	無気				p	t	k
有声音		b	d	g			

音韻論的レベルにおいては、日本語の破裂音には有声性 (voice) によって対立する2種類の子音、有声破裂音 (/b/, /d/, /g/) と無声破裂音 (/p/, /t/, /k/) の2項対立である。それに対して、中国語の破裂音は、有声音がなく、帯気性により無声有気音 (/p^h/, /t^h/, /k^h/) と無声無気音 (/p/, /t/, /k/) の2項対立である。すなわち、日本語の破裂音には声帯振動の有無による有声音・無声音の対立に対して、中国語には破裂時の呼気の強弱とそれによる調音器官の摩擦による有気音・無気音の対立で、それが意味の弁別に重要な働きをしている。しかし、音声レベルでは、それらの特徴だけでなく、さらに知覚的・音響的面から説明する必要がある。なぜなら、中国語を母語とする日本語学習者は、日本語の語頭音節にある無声音「カッコウ (格好)」を有声音の「ガッコウ (学校)」に、または語中音節にある有声音「コウガク (高額)」を無声音の「コウカク (広角)」に、などの例のようによく日本語の有声と無声を聴き間違えることが挙げられる。このような事実は、有声子音・

無声子音と有気音・無気音の間に違いが存在していることを示しており、中国語話者による日本語有声・無声破裂音の調音と知覚の両方に困難をもたらしている。以下では、日本語と中国語の破裂音の音声的特徴について、それぞれ 2.2.1 と 2.2.2 節で述べる。

2.2 日本語・中国語破裂音の音声的特徴

2.2.1 日本語破裂音の音声的特徴

日本語における語頭破裂音に関しては、様々な音響的特徴が指摘されている。その中でも代表的な特徴として、破裂音に後続する母音の基本周波数 (fundamental frequency, f_0) がある。 f_0 はピッチ (音の心理的高さ) に対応する音響的特性として知られており、語頭の有声音と無声音の間に f_0 の立ち上がりに差が見られる。しかし、有声・無声の区別に f_0 が関与するのは語頭の場合のみである。なぜなら、Gao and Arai (2018:59) によると、日本語はピッチアクセントを有する言語であり、語頭拍ではない後続拍の場合、ピッチアクセントの機能を担う f_0 と破裂音に由来する f_0 が共存している可能性があるためであると報告している。また、清水 (1993:169) や Shimizu (1996:37-38) では、無声子音に後続する母音の f_0 の立ち上がりは、有声破裂音に後続する場合よりも高い f_0 値になると報告している。さらに、清水 (2018:74) では、日本人女性の破裂音の f_0 の平均値について調査した結果、無声音が有声音より高いと指摘している。表 3 に清水 (1993:169) と (2018:74) で収集した日本語の語頭破裂音における後続母音の f_0 の平均値 (Hz) を示す。

表 3 日本語語頭破裂音における後続母音の f_0 の平均値 (Hz)

[清水(1993, 2018)より作成]

		清水(1993) n=72	清水(2018)
有声音	/b/	213.8 (SD=6.2)	236
	/d/		234
	/g/		227
無声音	/p/	248.5 (SD=19.5)	267
	/t/		252
	/k/		264

そこで、破裂音の有声・無声に対して発声タイプの調音と基本周波数の関係について清水(1993, 2018)では生理的・空気力学的な観点から説明された。まず、有聲破裂音に f_0 の低下が見られるのは、生理的に喉頭の位置が低く、空気力学的に気流量は少ないためである。それに対して、無聲破裂音に f_0 の上昇が見られるのは、喉頭の位置が高く、気流量が多いであることと知られている。すなわち、有聲破裂音を発声するとき、声門の閉鎖により喉頭の位置が低くなり、気流量も少なくなることで(図1の「閉鎖の形成」を参照)、内部喉頭筋の弛緩と喉頭位置が下降することによって、 f_0 の低下が見られることに対する一方、無聲破裂音を発声するとき、声門の閉鎖による気流量の増大、内部喉頭筋の緊張と喉頭位置が上昇することによって、 f_0 の上昇が見られる。このような f_0 の相違は、破裂音の有声性に固有のものであり、破裂音の生成と知覚における有聲・無声の弁別に関わっている可能性があるともされている(清水 2018:74)。また、 f_0 が始まる時から 100 ms 前後までの区間における f_0 曲線についても、有声音と無声音の間に差が見られ、無声音では開

始が高くその後低くなる下降パターンを、有声音では低い領域から開始しその後高くなる上昇パターンを示すことが報告されている (Shimizu 1996:31)。さらに、鄭・桐谷 (1998) では、日本語母語話者は有声・無声の弁別能力を持っていながら、有声と無声のピッチパターンを入れ替えた一部の音声に対して、誤聴する傾向が見られたと報告している。つまり、日本語母語話者の有声性の弁別に f_0 が少なからず影響を与えていることが窺える。

一方、日本語を含む大部分の言語においては、破裂音の有声・無声の区別に関する音響的特徴として VOT が弁別の有効な尺度であることが指摘されている (Lisker and Abramson 1964, 本間 1985, 清水 1993, Shimizu 1996 など)。VOT とは、声道閉鎖の開放から声帯振動が始まるまでの時間差であり、Lisker and Abramson (1964:403) によれば、声道閉鎖の開放前に声帯振動が始まる場合はマイナスの値で、声道閉鎖の開放後に声帯振動が始まる場合はプラスの値を示す。さらに、VOT の測定は声道閉鎖の開放を始点 (onset of release), すなわち破裂を基準時点 (0 点) とし、声帯振動が破裂に先行する場合 (voicing lead) はマイナスの値を取る有声破裂音 (図 2), 破裂と後続母音がほぼ同時に開始する場合 (simultaneous voicing) (図 3) またはプラスの値を取る場合 (図 4) 無声無気音, そして破裂の瞬間から後続母音のための声帯振動がかなり遅れ (voicing lag), 氣息 (aspiration) が伴う場合はプラスの値を取る無声有気音 (図 5) の, 3 つのタイプに分けられる。

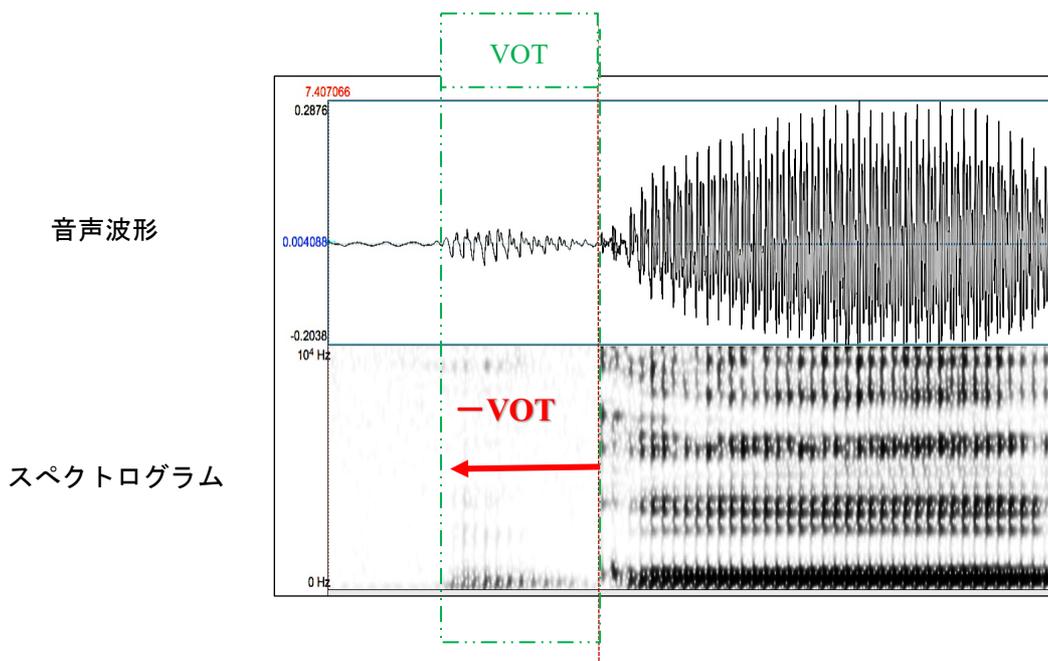


図 2 有声破裂音の判定例

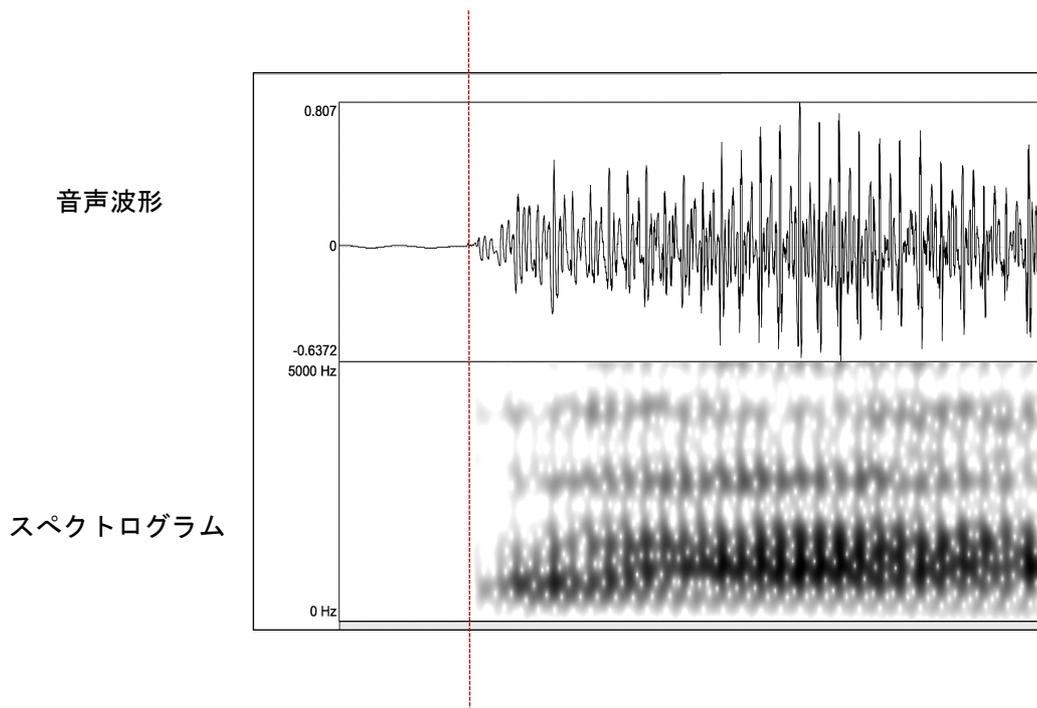


図3 simultaneous voicing の無声無気音の判定例 (VOT=0)

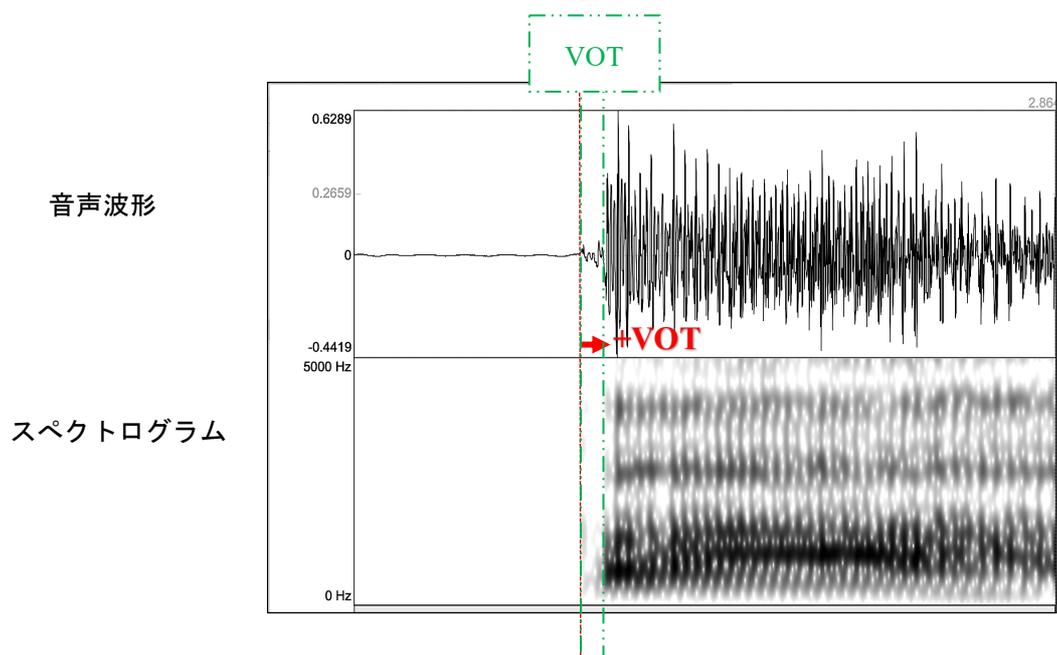


図4 プラスの値を取る無声無気音の判定例 (VOT=7 msec)

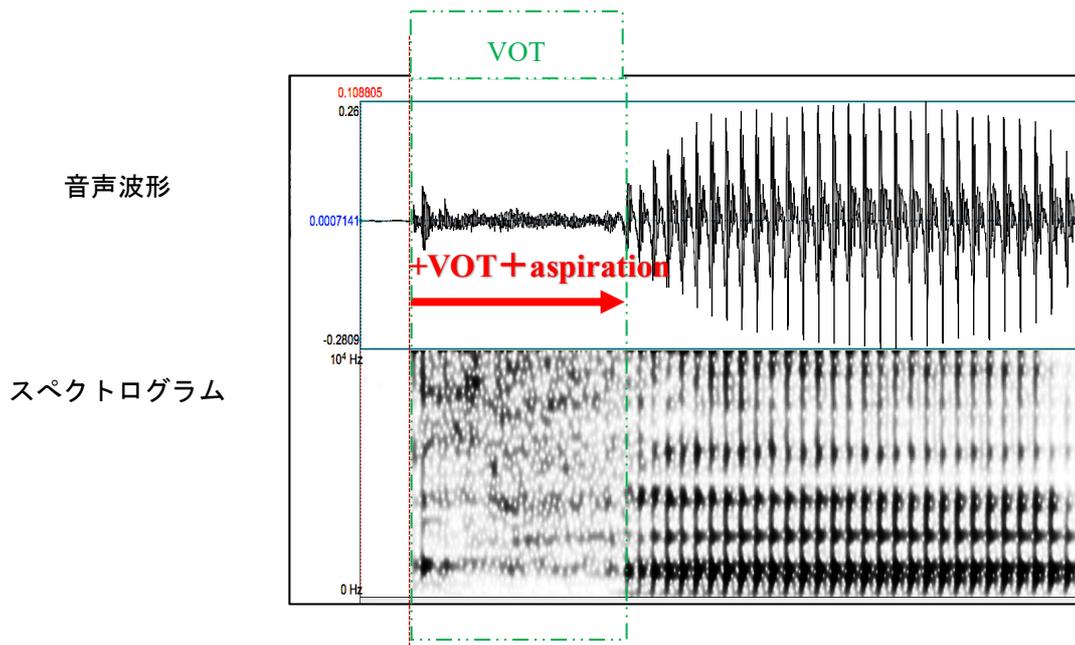


図5 無声有気音の判定例

図2から図5に示すように（図の上段は音声波形，下段はスペクトログラム），破裂の瞬間（赤の縦点線で示した時点）を基準点0とし，スペクトログラム上の voice bar を参照しながら，そこから声帯振動の開始時点までの時間の長さ実測値は VOT 値（緑の点線枠で示したところ）である。分析では，有声性を示す音声波形と，スペクトログラム上のボイス・バー（voice bar）が破裂の瞬間より前に現れる場合（すなわち破裂の前に声帯振動が始まる場合）は $-VOT$ の有声破裂音とし（図2），破裂前に有声性を示す音声波形が見られず，破裂（閉鎖開放）後における氣息音（aspiration）が見られる場合は $+VOT$ の無声有気音とする（図5）。さらに，図3及び図4のVOT値の判定に関しては，破裂の前に有声性を示す音声波形が見られない，図5と同様にプラスのものと判定するが，破裂後における aspiration が見られなく $+VOT$ の値の無声無気音とする。また，図3においては破裂の直後から声帯振動が始まるもので，VOT値が0のものと判定する。図4のVOT値の判定においては，スペクトログラム上のボイス・バーが始まるまでの区間がVOTと測定し，中国語の無気音と類似している。

Lisker and Abramson (1964:403) のデータによると，それぞれのVOTの尺度に関しては，有声破裂音は -100 ミリ秒，無声無気音は $+10$ ミリ秒，無声有気音は $+75$ ミリ秒程度のVOT値を中心に分布しているとされている。すなわち，無声有気音 $>$ 無声無気音 $>$ 有声音の順にVOT値が小さくなる。

日本語の破裂音は、有声破裂音ではVOT値がマイナスの値を取り、無声破裂音はプラスの値を取ると言われている（Homma 1980, 清水1993, Shimizu 1996, 杉藤 1996）。しかし、日本語の語頭破裂音のVOT値はばらつきが大きく、語中位置（語頭・語中）によって条件付けられない自由異音として出現することが報告されている（Homma 1980, 清水1993, 2018, 朱1994, 2010, 高田2004）。つまり、日本語の語頭有声音の中には、音韻的には有声音であるが、無声音化の現象が起こることでVOTがプラスの値を取るものがあるということである。これについては、高田の一連の研究（2004, 2006, 2008, 2011）によれば、日本語の語頭有声音のVOT値は、地域差や世代差が関わっていることが指摘されている。地域差に関しては、関東以西（近畿地方を中心として）では マイナスのVOT値が観察されるのに対して、東北地方ではプラスのVOT値が観察されるという。世代差に関しては、若い世代において大きな変化が見られ、関東以西においても東北地方と同様にプラスのVOT値が多く観察された。なお、聴取実験では、有声音において破裂前に声帯振動が伴うかどうか、すなわち日本語の語頭有声音が無声音化するか否かという現象が、日本語母語話者の有声性の弁別に積極的に関わっておらず、有声音と判断されることもあることが確認されている（高田2004:63, 朱2010:53）。さらに、図6は日本語の無声音化した語頭有声音の具体例である。典型的な有声破裂音は図2に示したように、破裂の瞬間より前に声帯振動によって、スペクトログラム上にvoice barが現れるが、図6ではそれが確認されなかった。図6のように、破裂音の閉鎖を開放した後に母音のための声帯振動が始まることもあり、この場合は有声破裂音のVOT値はマイナスではなくプラスと判定する。

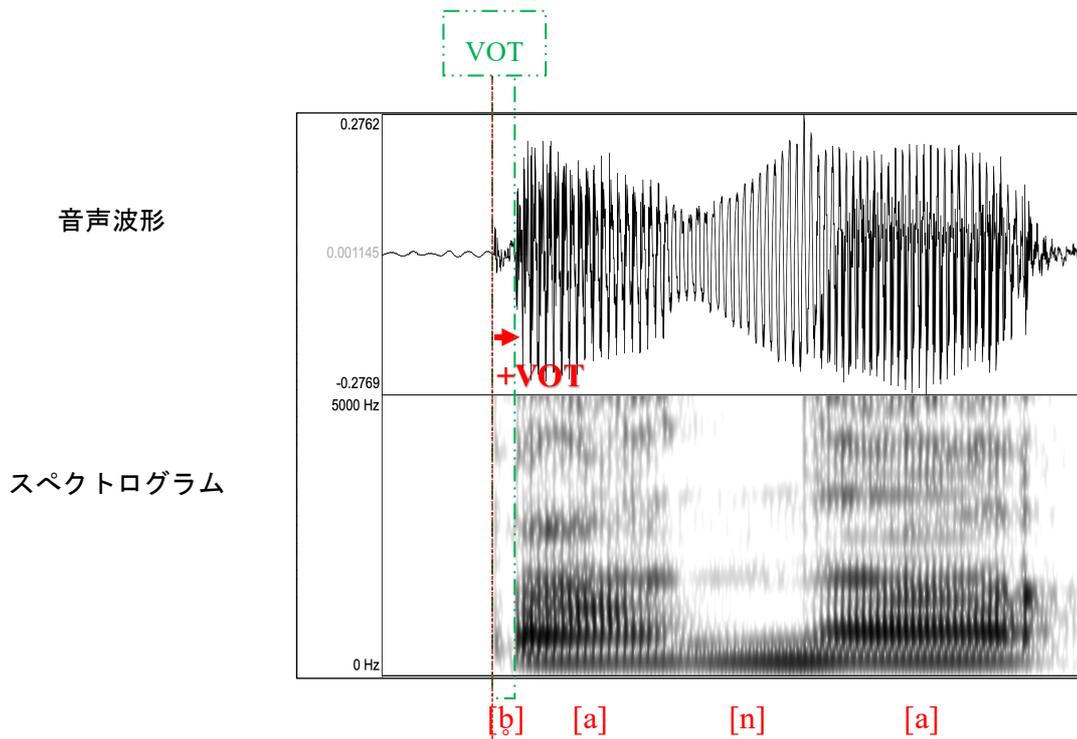


図6 語頭無声音化の有声破裂音の判定例「バナ」

一方、語中の破裂音に関しては、VOTを測定できるが、語頭の場合と全く同じようにVOTを測定することは難しい。なぜなら、有声音においては、先行する母音から閉鎖区間、後続母音に至るまで継続して声帯振動が観察されることが多く、声帯振動の開始点が破裂音周辺に存在しないためである。杉藤・神田（1987:16）でも、有聲・無聲の知覚特徴において、声帯振動の有無やVOTが重要な手掛かりとは言えないと報告している。以上の理由で、語中破裂音の有聲・無聲破裂音の音響的特徴の観察において、①「先行母音」、②語中にくる子音の破裂前の口腔の閉鎖による子音の「閉鎖区間」（杉藤・神田（1987:2）では「内破」と呼ぶ）、③閉鎖が解除された時点から後続母音が始まるまでの「VOT」（杉藤・神田（1987:6）では「外破」と呼ぶ）、④「後続母音」、この4つの区間の持続時間を測定する（図7）。図7について、上段は音声波形、下段はスペクトログラムと呼ばれるものである。下段のスペクトログラム上、縦軸が周波数、横軸が時間、色の濃さがエネルギーの強さ、そしてスペクトログラム上の横の赤色点線はその周波数帯域でエネルギーがピークになっているところを結んだものでフォルマント（formant）という。音声のスペクトルから観察されるフォルマントは、下から数えて「第1フォルマント周波数（F₁）」、「第2フォルマント周波数（F₂）」、「第3フォルマント周波数（F₃）」と呼ぶ。

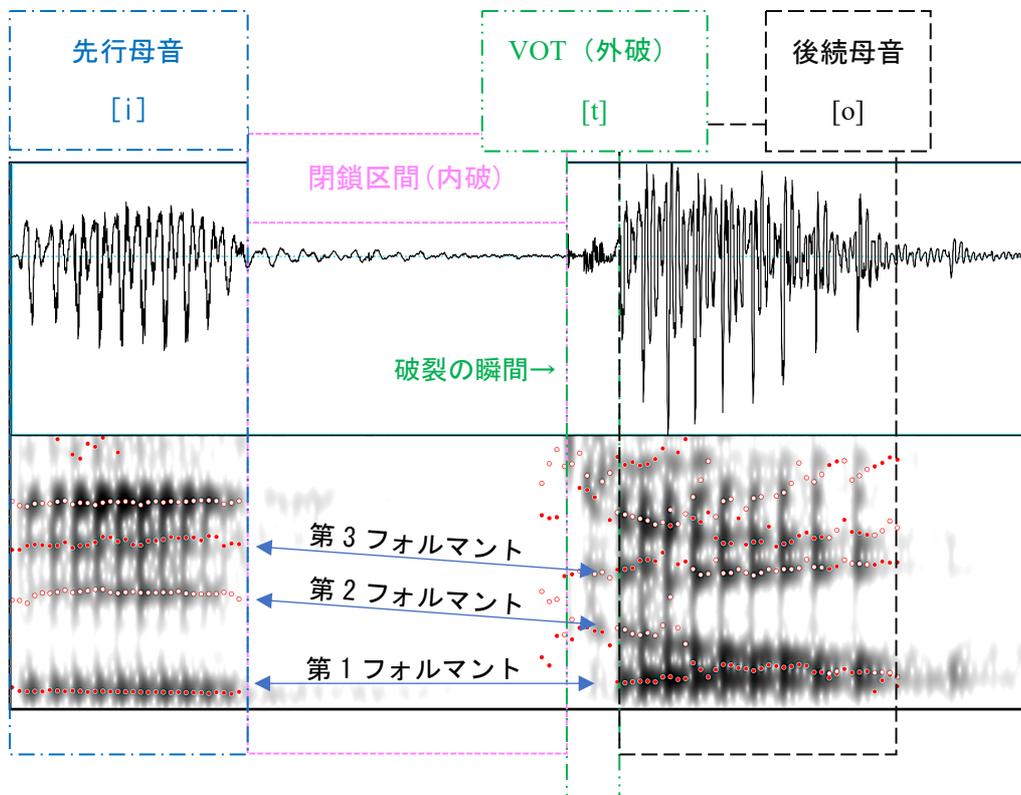


図7 語中無声音についての観察例「イト」

図7に示すように、VOTの測定方法については、語中無声音は語頭の場合と同様に、スペクトログラム上のvoice barを参照しながら、破裂の瞬間から声帯振動の開始時点までを測定することができる（プラスの場合）。本論文では先行母音の「第3フォルマント周波数（ F_3 ）」の終点を閉鎖区間の開始点とし、VOT区間の開始点（いわゆる「破裂の瞬間」）を閉鎖区間の終点として、その区間の持続時間を測定する。閉鎖区間（内破）の持続時間については、先行母音の終了時点から破裂の瞬間までの時間長である（ピンクの点線枠で示したところ）。

さらに、語中有声音のVOTの測定については、閉鎖区間内のvoice barの有無の判定、voice barが観察されるとしたらどのように現れるか、またvoice barが破裂の後に始まる場合、破裂から後続母音の声帯振動の開始まで（VOT、つまり「外破」）の時間がどれほど持続したかに注目することができる。なお、閉鎖区間内のvoice barは「先行母音より引き続く声帯振動の残留」と「破裂音のためのprevoicing」を分けて考えられる（朱1994、高田2013、2014、邊2017）。閉鎖区間に破裂音のためのprevoicingがあれば、語頭有声音と同様

に、 $-VOT$ として認定できる。すなわち、語中有声音の VOT の測定については、3つのパターンがある（図8, 9, 10）。

パターン①：閉鎖区間において、「先行母音より引き続き声帯振動の残留」と「破裂音のための prevoicing」による声帯振動が全体にわたって観察される。これは典型的な有声音の特徴のパターンである（図8）。

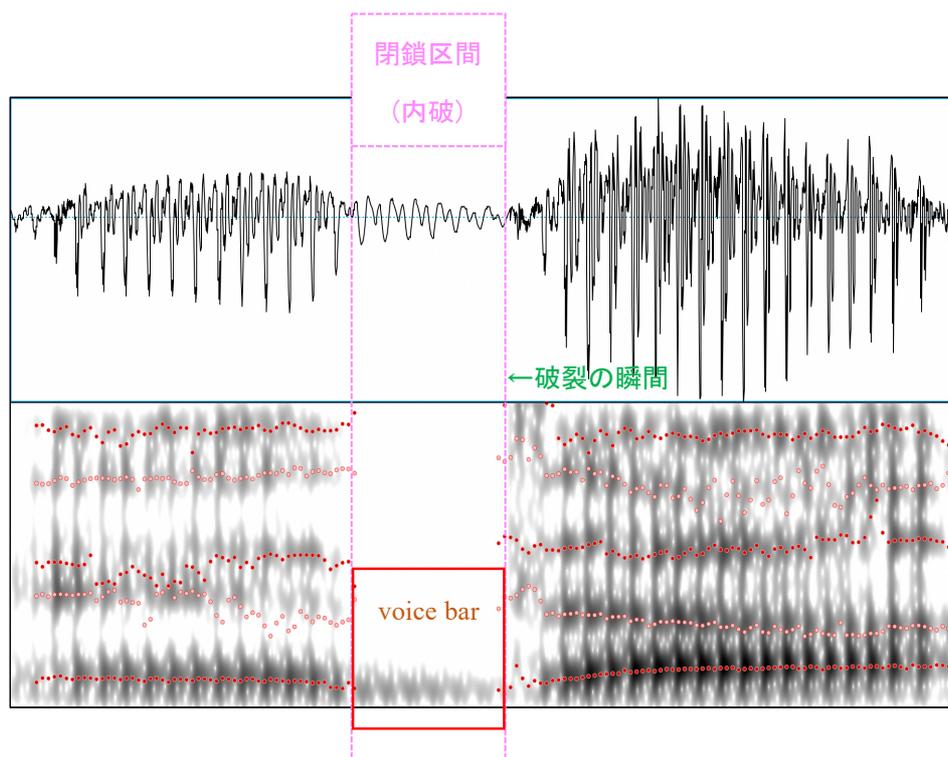


図8 語中有声破裂音についての観察例：パターン①

図8に示したのは、閉鎖区間の全体にわたって声帯振動が観察される典型的な $-VOT$ の有声音パターンである。破線の枠で括られた「閉鎖区間(内破)」の終点は破裂の瞬間であり、破裂の前に声帯振動、すなわち **voice bar**（実線の枠で括られたところ）が観察できるのは VOT 値がマイナスの特徴である。

しかし、このパターンにおいて、声帯振動が先行する母音から閉鎖区間、後続母音に至るまで観察されることがあり、声帯振動の開始点が破裂音周辺に存在しないため、語頭の場合と全く同じように VOT を測定することはできない。

パターン②：声帯振動が閉鎖区間に観察され、閉鎖の解放（破裂の瞬間）の直前に弱くなるパターン（図9）

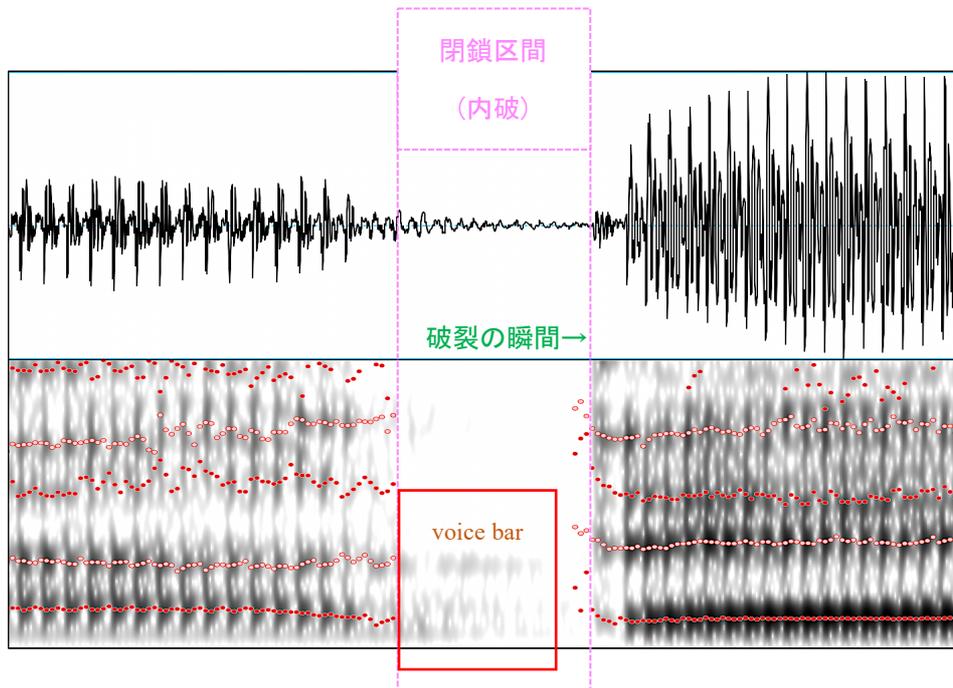


図 9 語中有声破裂音についての観察例：パターン②

図 9 では、閉鎖区間（破線の枠で括られたところ）に、「先行母音から引き続く声帯振動」が観察されるが、図 8 と違い、声帯振動が破裂（閉鎖の開放）の直前に弱くなる観察例で、VOT が測定可能な語中有声音である。

パターン③：語中の有声破裂音が「破裂音のための prevoicing」がなく、+VOT の特徴を持つパターン（図 10）

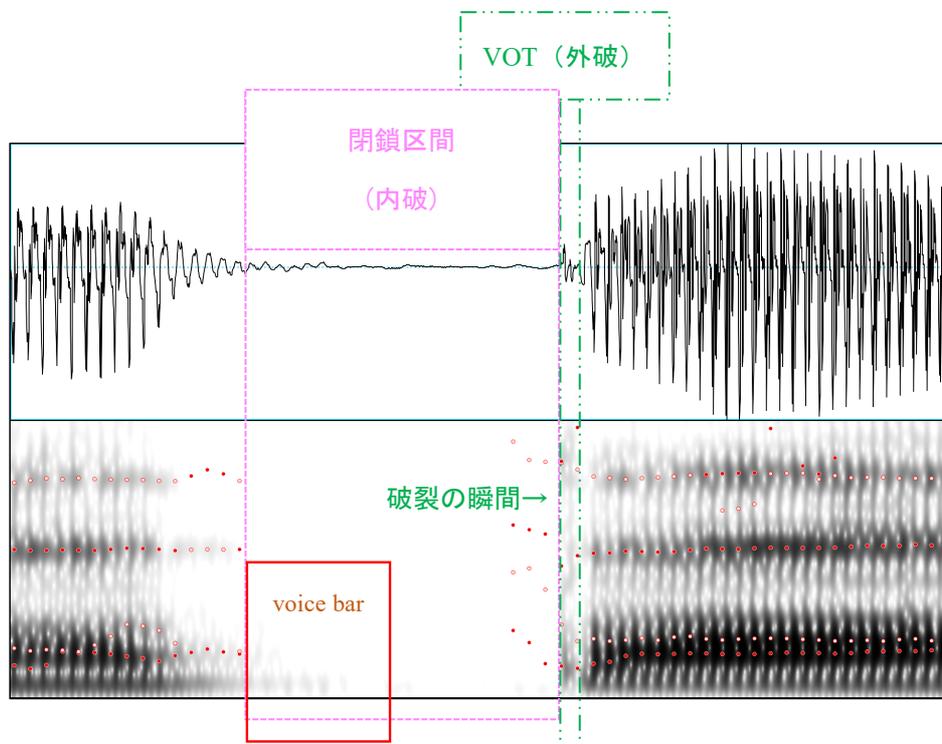


図 10 語中有声破裂音についての観察例：パターン③

図 10 においては、先行母音の F_3 が消えた時点を閉鎖区間の始点としているが、その後も弱い声帯振動が観察され（実線の枠で括られたところ）、その部分は「先行母音より引き続く声帯振動の残留」と判断したほうが合理的である。破裂前の声帯振動（voice bar）がないほか、破裂の瞬間から後続母音の始点までの外破の区間が観察されるので、この場合の VOT はマイナスの値ではなく、プラスの値と判定できる。また、このパターン③の音響的特徴は、中国語の無声無気音と類似している。

以上が日本語の有声・無声破裂音が語頭及び語中にくる場合の音響的特徴と測定方法である。続いて、これらの測定方法を用いて、さらに詳しく日本語の語頭及び語中破裂音の音響的特徴を分析した先行研究を見ていく。

杉藤・神田（1987:10-11）によれば、日本語の無声音は、語頭と比べると、語中の VOT の値が短いことが確認されている。また、語中破裂音の VOT（外破）の時間に関しても、破裂に先立ち声帯振動が先行していない有声音の値が無声音とほぼ同じか、無声音の方がやや長いという。表 4 は、杉藤・神田（1987:9-10）による語頭と語中の有声・無声破裂音

の VOT（「外破」を含め）をまとめたものである。

表 4 語頭と語中破裂音の VOT の平均値 (msec)

[杉藤・神田(1987) p.9-10 より作成]

		話者 1		話者 2	
		語頭	語中	語頭	語中
無 声 音	/t/	28 (SD=0.88)	19 (SD=0.20)	33 (SD=0.52)	18 (SD=0.27)
	/k/	68 (SD=1.77)	31 (SD=0.49)	51 (SD=1.16)	33 (SD=0.61)
有 声 音	/d/		-61 (SD=0.92)		-69 (SD=0.58)
		-13 (SD=0.42)	18 (SD=0.61)	-19 (SD=1.16)	14 (SD=0.20)
	/g/		-102 (SD=0.61)		-56 (SD=3.10)
		25 (SD=0.32)	0	19 (SD=0.38)	20 (SD=0.45)

さらに、VOT の持続時間長と知覚との関係について、朱 (2010:52) は無気音のように発話された有声子音の外破の持続時間が 20msec 以上を超えると、有声子音と知覚する正聴率が低下することがあると報告している。田村・森田・伊藤・廣瀬・森 (2016) は日本語母語話者に合成音声/ta/-/da/を用いて聴取実験を行なった結果、VOT が 15~20msec を境に有声の/da/から/ta/へと急激に切り替わると VOT の知覚境界値が報告されている。

次に、語中の破裂音の有声性に関しては、VOT と同様に閉鎖区間（内破）の持続時間も

注目すべきである。一般的に、閉鎖区間以外の条件が同じであれば、語中において有声・無声破裂音では、無声音の方が閉鎖区間が長いことが知られている (Lisker 1957, Homma 1981)。そして、日本語においても、閉鎖区間 (内破) の持続時間は、有声音より無声音の方が長いことが確認されている (Homma 1981, 杉藤・神田 1987)。破裂音の有声・無声を知覚する際に、杉藤・神田 (1987:16) では、中国語系日本語学習者の日本語有声・無声音の知覚の特徴について調査した結果、閉鎖区間 (内破) の持続時間は VOT (外破)・破裂の強さ・声帯振動の有無より重要な役割を果たしていることを報告している。閉鎖区間の持続時間が短いほど有声音に聴こえることが報告されている (杉藤 1996, 朱 2010)。

最後に、山崎 (2014) は語頭破裂音における新しい音響パラメータとして、「子音区間の高周波数帯域の噪音成分の強度」を提案した。それについて、山崎 (2014) は、朝鮮語⁵にある 3 種類の破裂音の破裂の強さを分析した結果、語中の場合において、平音 (/p/, /t/, /k/) と激音 (/p^h/, /t^h/, /k^h/) の VOT 区間では、4000Hz~6000Hz 付近の噪音成分の大きさに違いがあり、高周波数帯域の噪音成分のパワーが異なっていることを明らかにした。そこで、李 (2021) では、このパラメータが取り扱われ、中国語母語話者が発音した日本語の無気音のような有声破裂音を対象に考察したが、山崎 (2014) で指摘された朝鮮語破裂音ほど顕著な差が見られなかったと報告している。

上述した日本語の有声・無声破裂音の弁別パラメータをまとめると、(1) VOT および外破、(2) 後続母音の持続時間、(3) 後続母音の f_0 、(4) 閉鎖区間 (内破) の持続時間、(5) 高周波数帯域の噪音成分のパワー、の合計 5 種類の特徴が先行研究で言及されている。

2.2.2 中国語裂音の音声的特徴

中国語の破裂音は、無声有気音と無声無気音に分類される。呼気の強さとそれによる調音器官の摩擦が有気・無気の重要な区別素性であり、強い呼気と摩擦があれば有気音に聴こえ、摩擦がなければ無気音に聴こえる (朱 1994, 2010:27)。朱 (1994, 2010) によると、中国語の有気音は、音声環境 (語頭・語中) に関わらず、日本語の破裂音よりも呼気が強い。また、有気音の呼气流率の平均値は無気音の 2 倍以上で後続母音にまでかぶさり、語頭よりも語中の呼气流率の方が高くなる傾向が見られた (後続母音が [u] のときを除く)。なお、中国語「轻声」音節中の無声無気音や先行音節の声調が第 3 声で先行母音が /i/, /u/、

⁵ 朝鮮語の破裂音は、有声・無声の対立がなく、無声領域にある平音 /p/, /t/, /k/・激音 /p^h/, /t^h/, /k^h/・濃音 /pʷ/, /tʷ/, /kʷ/ の 3 項対立である。

/ei/以外の母音に後続する場合は有声化する率が高いと報告されている（朱・杉藤 1998）。

一方、清水（1993）では、日本語母語話者と中国語母語話者を対象に、両言語の破裂音のVOT 値について調査を行った。日本語・中国語におけるVOTの平均値、最小値・最大値の分布は表5に示す範囲であることが確認された。中国語話者の破裂音においては、無声無気音・無声有気音ともにプラスの値を取り、無声無気音は無声有気音に比べてVOT値が小さい。さらに、日本語のVOT値が広い範囲に分布しているため、中国語話者の無声無気音（VOTがプラスの値）は日本語の有声音と無声音の両方に跨っているといえる。

表 5 日本語・中国語における語頭破裂音の VOT の平均値と最小値・最大値（msec）

[清水（1993）p.167 より作成]

		日本語 n=72	中国語 n=18		
無声音	/p/	41 (SD=17.1) 15~65	無気音	/p/	7 (SD=2.3) 5~10
				/t/	12 (SD=2.1) 10~15
				/k/	19 (SD=3.8) 12~25
	/t/	30 (SD=12.7) 15~50	有気音	/p ^h /	96 (SD=13.3) 80~115
	/k/	66 (SD=12.1) 50~100		/t ^h /	98 (SD=16.1) 80~120
				/k ^h /	112 (SD=20.7) 90~130
有声音	/b/	-89 (SD=28.5) -65~-125			
	/d/	-75 (SD=32.7) -40~-135			
	/g/	-75 (SD=27.0) -35~-125			

2.3 中国語話者による日本語破裂音の習得研究

中国語話者による日本語の破裂音の習得に関しては、これまで多くの研究結果が報告されている。2.1でも述べたように日本語と中国語の破裂音の音韻体系には差異があるため、中国語には存在しない日本語の有声・無声音の習得は難しいとされてきた（蔡1976, 鈴木1985, 杉藤・神田1987, 横山2000）。これは、日本語では有声音・無声音の対立が存在していることに対して、中国語では有声性対立がなく有気音・無気音の対立が存在していることに起因していると言われている。また、中国語話者は日本語の無声音、特に語中の無声音を有声音と聴く誤聴が多いと報告されている（朱1994, 福岡1995, 山本2004, 劉2005）。その理由について福岡（1995:45）は「日本語の語中無声音は帯気が弱くVOT値も北京語の無声無気音と類似している」からだと言っている。つまり、中国語話者は日本語の語中無声音を母語である中国語の無声無気音として聴いた結果、日本語の有声音に聴き間違えるということである。

一方、語頭破裂音に関しては、有声・無声に関わらず、正聴率が高いことが報告されている（福岡1995, 山本2004, 劉2005）。その要因として朱（1994）は、日本語の無声音が語頭にある場合は、中国語の有気音ほどではないが呼気が強いため、呼気の強さが主要な弁別的特徴である中国語を母語とする日本語学習者は、有声音と混同する確率が低く、正聴率が高くなったと指摘している。しかし、実際語頭有声音の音声特徴といえば、VOTの値のばらつきが大きく（清水1993:168）、呼気流率が中国語の無気音と同等かやや弱い（朱2010:35）、すなわち音韻的には有声音であるが、音声的には破裂前に声帯振動が伴わない無声無気音となることもある。この語頭有声音のVOT値がマイナスの値を取らない現象については「無声子音化」（朱2010:50）、もしくは「+VOT化」（高田2004:57）と呼ばれている。さらに、朱（2010:50-53）では日本語母語話者と中国語母語話者が発話した「無気音のように発話された語頭の有声子音」を用い、日本語母語話者を対象に聴取実験を行った結果、発話者の母語と関係なく、無声音化しても有声音と知覚することが可能であるほか、VOTの値が6 msec以下の無気音は有声音と知覚されると報告している。さらに、この語頭有声音の無声音化現象については、高田の一連の研究（2004, 2006, 2008, 2011）によると、地域差や世代差でVOTの値がマイナスではなくプラスの値、すなわち無声音化が多く観察されるという。この無声音化した語頭有声音の音響的特徴は無気音と類似しているが、日本語母語話者は無声音化した語頭の有声音を無声音に知覚することがないと指摘されて

いる（高田2004:62-63, 朱2010:53）。そこで瞿（2017, 2021）では、この無声音化の現象に着目し、日本語母語話者と初級の台湾人日本語学習者（台湾華語⁶話者と台湾閩南語⁷話者）を対象に、語頭有声音の知覚実験を行った。その結果、台湾華語話者（中国語話者）は日本語母語話者や台湾閩南語話者よりも無声音化した語頭有声音の正聴率が低かった。つまり、母語（母方言）における有声性の対立の有無が、無声音化した語頭有声音の弁別に影響していることが示唆された。

なお、無声音化した語頭有声音のVOTが短く、音響的特徴も無声無気音と類似していることから考えると、中国語話者は日本語の語中無声音を有声音と知覚することから、無声音化した語頭有声音は有声音と知覚することができるはずである。しかしながら、日本語母語話者や台湾閩南語話者と比べて中国語話者は無声音化した語頭有声音を無声音と誤聴する傾向が確認された。先行研究の結果は表6に示す。

⁶ 台湾華語（Taiwan Mandarin）とは、中国北方方言音を標準語音とする言語であり、台湾で60年以上を経て、台湾閩南語や客家語など他の言語からの影響を受けつつ変化し、中国北方方言から独立して台湾社会に定着している言語である（何2009, 王2013）。そして、1956年に台湾政府が打ち出した国語政策（台湾では「説國語運動」という）によって、台湾華語が「国語」とされている（李2010）。

⁷ 台湾閩南語とは、中国語閩方言に属しており、台語や閩南語とも呼ばれている。台湾閩南語には、有声破裂音（/d/の子音を除き、/b/と/g/がある）が存在し、無声有気音・無声無気音・有声音の3項対立となっている。

表 6 先行研究結果のまとめ

			日本語母語話者の 日本語破裂音の知覚	中国語母語話者の 日本語破裂音の知覚
日 本 語 の 破 裂 音	無 声 音	語頭 無声有気音 (例：[p ^h])	無声音 (例：[p ^h])	有気音（無声音） (例：[p ^h])
		語中 無声無気音 (例：[p])	無声音 (例：[p])	無気音（有声音） (例：[b])
	有 声 音	語頭の 無声音化した有声音 (例：[p])	有声音 (例：[b])	無気音 (無声音・有声音) (例：[p]・[b])
		有声音 (例：[b])	有声音 (例：[b])	無気音（有声音） (例：[b])

表6に示した通り、日本語の破裂音、特に音響的特徴が類似している無声音化した語頭有声音と語中無声音との間にどのような違いがあるのか、さらに検証する必要があるだろう。すなわち、日本語母語話者の破裂音の識別に重要な判断材料は何なのか、それについて検証すべき点がある。

2.4 日本語破裂音の知覚的特徴

第1章で述べたように、日本語母語話者が、日本語の語中無声音を含む音節を文脈から切り出すと、音響的無声音を有声音として知覚する現象がある。ではなぜ語中の無声音を単独で聴くと有声音として知覚するのか。この現象に関与している要因を考察することで、日本語破裂音の新たな音声的特徴が見えてくるのではないかと考えられる。

そこで本論では、日本語の語中無声破裂音を含む音節を文脈から切り出すと有声音として知覚される現象を中心に、日本語母語話者の有聲・無声音の知覚にかかわる音響的・知覚的特徴を「無声音として発声した無声無気音」と無声音化した語頭有声音（いわゆる「有声音として発声した無声無気音」）に焦点を当て分析を行う。

第3章 知覚実験 I

3.1 実験の目的

2.2.1 で述べたように、日本語破裂音の音声的特徴においては、中国語の無気音のような無声無気音が存在している、という報告がなされている(朱 2010)。それは語中における無声音として発声した無声無気音と、語頭における有声音として発声した無声無気音である(高田 2004, 朱 2010)。無声音として発声したものに関しては、筆者の指導教官である朱春躍教授が授業やゼミで、文脈「そうですか。」の文末にある無声子音を含む音節「か」を文から切り出して、その音節だけを聴かせると、発話者本人を含め、日本語母語話者が有声音の「が」と知覚する、という現象を確認している。しかし、この現象は、本格的な知覚実験により確認されたものではない。そこで本研究では、日本語の2音節無意味語から切り出した第2音節の音を、日本語母語話者がどのように知覚するのか確認するため、知覚実験 I を行う。

この知覚実験 I を通して、主に以下の3点を観察していく。

- (1) 日本語母語話者は日本語の語中無声音を語中から切り出して単独に聴かせると、無声音を有声音として知覚するかどうか。
- (2) もし(1)が確認できた場合、その要因は例文の下降調の「そうですか。」のように、「か」のピッチが低いことから、ピッチの影響を受けてこの現象が生じた可能性があるかどうか。
- (3) もし(2)が確認できた場合、その反対に高ピッチの「か」も同様に、語中から切り出すと無声音が有声音に聴こえるかどうか。

3.2 録音の対象と方法

3.2.1 録音協力者

発話音声の録音には、日本語母語話者1名に協力してもらった。実験協力者は神戸市出身の20代男性で、高等教育機関で4年以上の教育歴を有するものである。

3.2.2 録音語

実験対象である語中無声音の子音は/p/で、後続母音は/a/に統一した。どの先行母音の環境から切り出しても上記の現象が見られるかどうか確認するため、先行母音は/a/, /i/, /u/, /e/, /o/の5種とした。さらに、ピッチの影響を検証するため、アクセント型は頭高型(HL)

と平板型 (LH) の 2 種類とした。以上の組み合わせにより、発話してもらった破裂音は 2 拍の無意味語、計 10 種である (表 7)。

表 7 録音語内容

頭高型 (H <u>L</u>)	平板型 (<u>L</u> H)
ア <u>パ</u>	<u>ア</u> パ
イ <u>パ</u>	<u>イ</u> パ
ウ <u>パ</u>	<u>ウ</u> パ
エ <u>パ</u>	<u>エ</u> パ
オ <u>パ</u>	<u>オ</u> パ

* 下線は低ピッチを示す。

3.2.3 音声の録音

音声の録音の詳細は以下のとおりである。

- (1) 実験用音声の録音：上記の 10 種の無意味語を図 11 のように配列して、録音協力者に発音してもらい、録音を行った。音声の録音は、大学の無響室で単一指向性マイクを使って、メモリレコーダーに収録した。
- (2) 実験用音声の録音機材：ポータブルリニア PCM レコーダー (SONY PCM-D50)。
- (3) 実験用音声の録音条件：録音のサンプリング周波数は 44100Hz、量子化ビット数は 16 bit。
- (4) 実験用音声の日本語母語話者による知覚確認：得られた発話音声につて、音声知覚実験に参加していない日本語母語話者 (近畿圏出身) 5 名に、録音内容が発話語読み上げリスト (図 11) のとおりに正確に発音されたかを評定してもらい、聴いた第 2

音節の破裂音の部分を紙面に仮名表記で書くように協力してもらった。その結果、5名の母語話者はすべて録音語の第2音節を無声音と判断した。

3.2.4 録音手順

録音の際、録音協力者に指示を正しく理解してもらうため、録音リストを用意した(図11)。図11に示すように、読み上げリストは、平仮名で書かれた無意味語(例:「あば」)の文字を10行並べたものである。

* () 内の「o回目」というのは発音しないでください。	
(1回目)	あば
(2回目)	あば
(3回目)	あば
(4回目)	あば
(5回目)	あば
(6回目)	あば
(7回目)	あば
(8回目)	あば
(9回目)	あば
(10回目)	あば

図11 発話語読み上げリスト

以上の録音材料を用意し、録音を開始する前に発話者に、収録の手順、発話内容、発話スピードに関する説明を行った後に読み上げリストを渡し、1語につき5回ずつ発話してもらい、5回目と6回目の間に30秒の休憩を挟む。最終的に、1語につき計10回の発話をしてもらった。

発話の際の留意点として、必ず1語ずつ区切って、「あば、あば、あば…」と話者にとって普通の発話速度で読み上げるように、と指示した。なお、刺激音に選んだのは2回目から4回目までの3発話と、7回目から9回目までの3発話で、計6発話である。

3.3 刺激音の作成

刺激音は、平板型と頭高型2種類の録音語（例：「あば」など）から第2音節の「ば」のみを切り出すものとする。第2音節の「ば」の種類は、録音語の第1音節の母音により5種類のもので、ピッチにより高（H）と低（L）2種類のもので、計10種類である。さらに、各種類の音を10回の録音語から6発話を選び、計60個を用意した。したがって、ピッチの高・低（2種類）×先行母音/a/, /i/, /u/, /e/, /o/（5種類）×6発話、とするものである。実験に無関係な60個のダミーを混ぜ込み、計120個の刺激音とした。種類毎の具体例は、図12から図21に示す。

具体例1：録音語「アバ」から切り出した刺激音「パ」

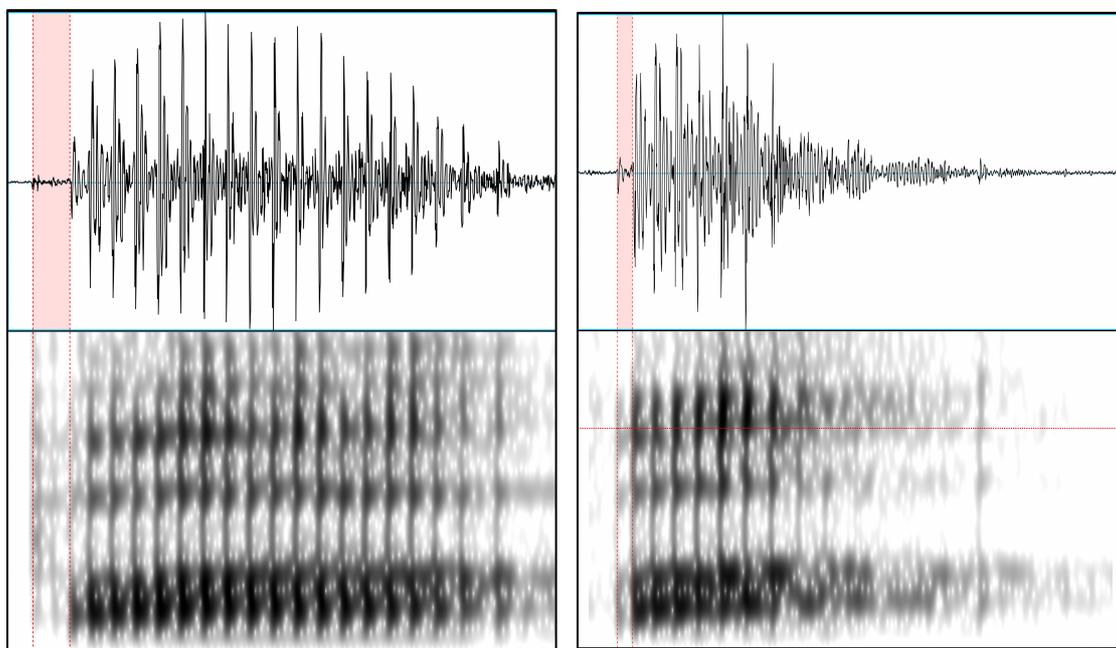


図12 高ピッチ「パ」の具体例
(VOT= 14.2 msec)

図13 低ピッチ「パ」の具体例
(VOT= 7.0 msec)

具体例 2 : 録音語「イパ」から切り出した刺激音「パ」

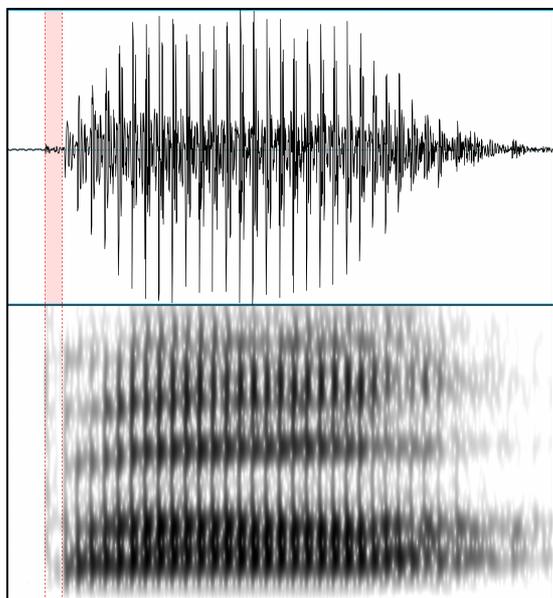


図 14 高ピッチ「パ」の具体例
(VOT= 10.5 msec)

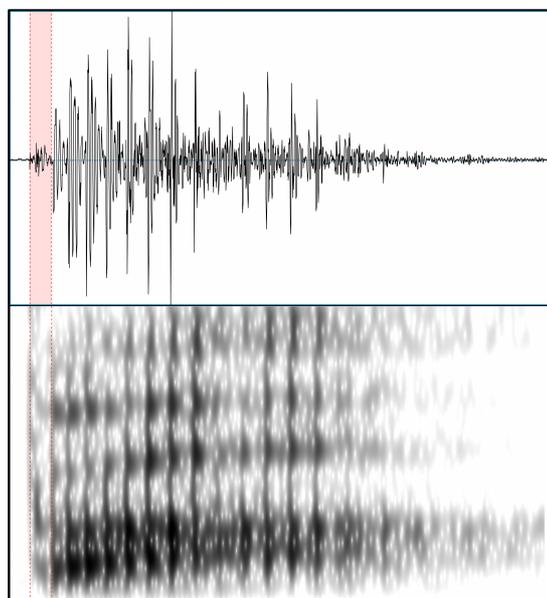


図 15 低ピッチ「パ」の具体例
(VOT= 10.4 msec)

具体例 3 : 録音語「ウパ」から切り出した刺激音「パ」

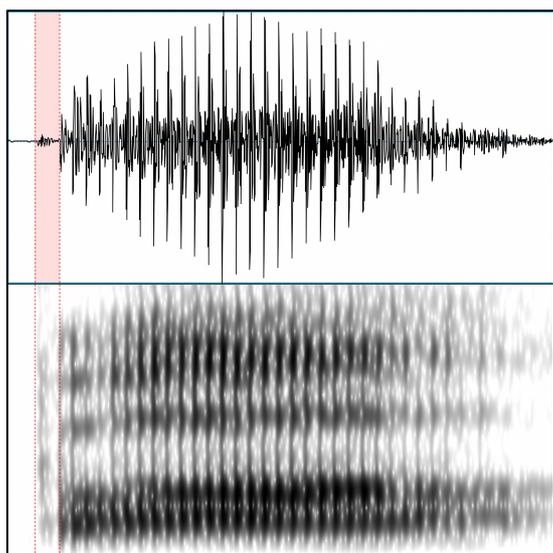


図 16 高ピッチ「パ」の具体例
(VOT= 14.2 msec)

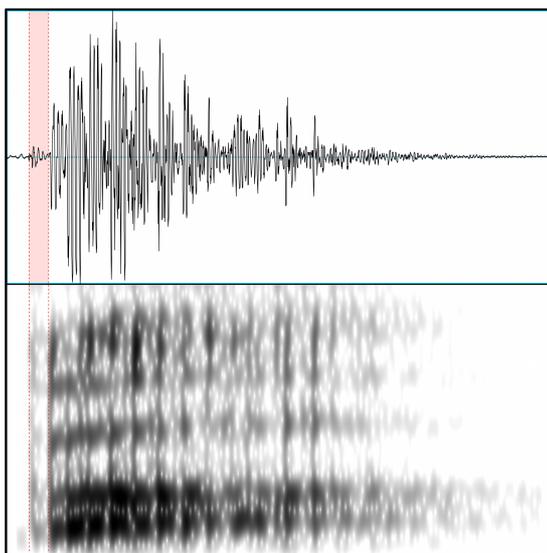


図 17 低ピッチ「パ」の具体例
(VOT= 9.3 msec)

具体例 4 : 録音語「エパ」から切り出した刺激音「パ」

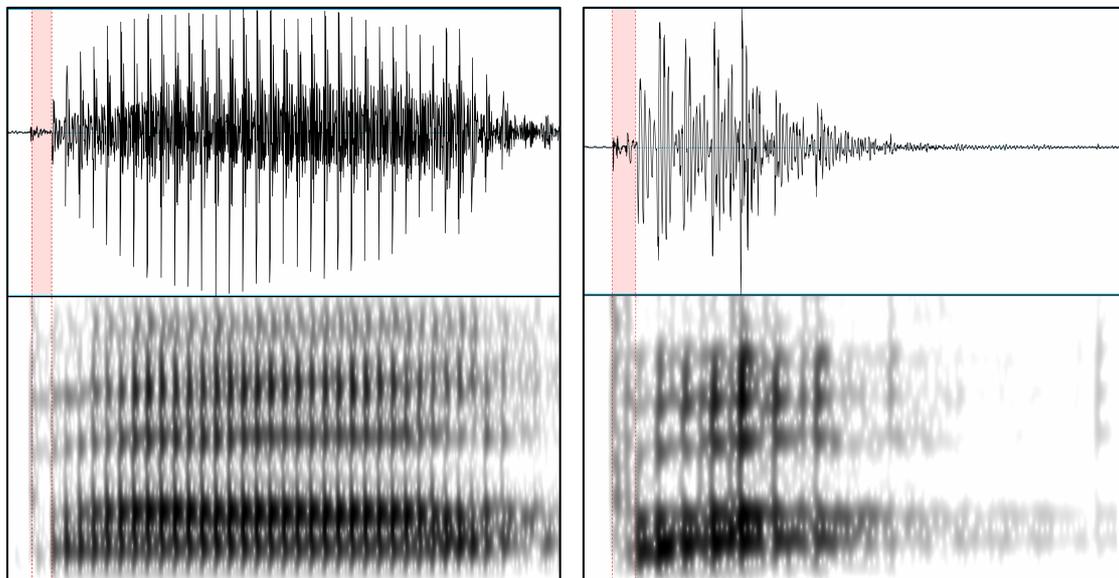


図 18 高ピッチ「パ」の具体例
(VOT= 12.4 msec)

図 19 低ピッチ「パ」の具体例
(VOT= 9.5 msec)

具体例 5 : 録音語「オパ」から切り出した刺激音「パ」

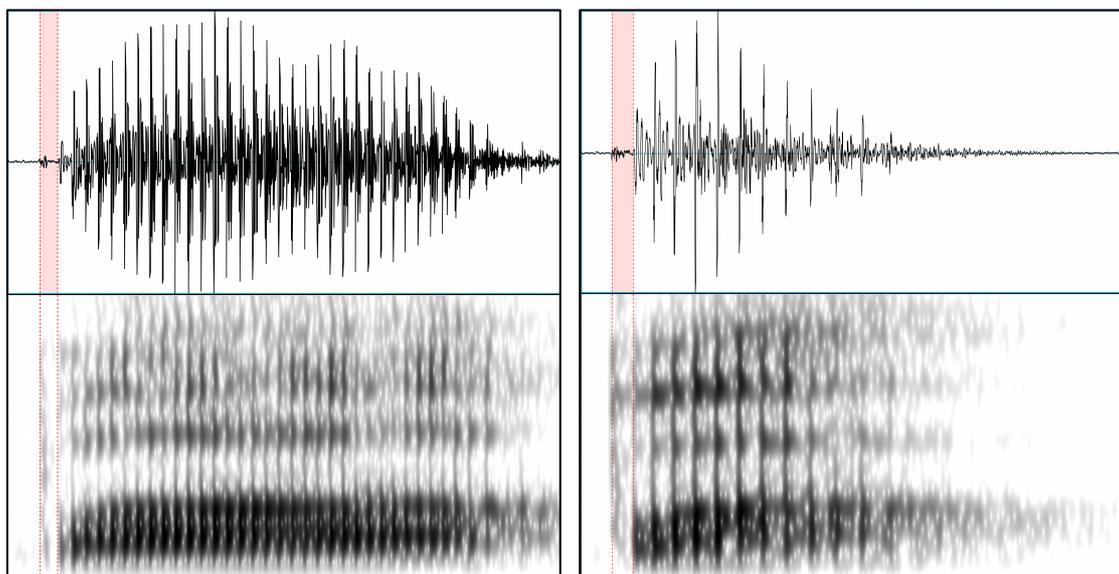


図 20 高ピッチ「パ」の具体例
(VOT= 11.3 msec)

図 21 低ピッチ「パ」の具体例
(VOT= 10.4 msec)

刺激音のピッチの平均は、高ピッチ (H) が 120.5 Hz, 低ピッチ (L) が 84.16 Hz, その差は 6.2 st であり, およそ音楽上の「ソ→レ」くらいの差にあたる。

3.3.1 刺激音における音響的特徴の比較

日本語における有声・無声の破裂音の弁別にかかわる可能性がある音響的特徴として, ①VOT, ②閉鎖区間の長さ (いわゆる「内破」), ③閉鎖区間と VOT を合計した全長, ④後続母音の長さ, ⑤後続母音の f_0 , ⑥高周波数帯域の噪音成分のパワーなどが挙げられている (杉藤 1996, 朱 2010, 山崎 2014, 邊 2019)。しかし, 2.2.1 で報告したように, ⑥破裂の強さについて, 李 (2021) では有声・無声破裂音の間に顕著な差が見られなかったと報告されていることを踏まえ, 本研究では破裂の強さを考慮しない。そこで, あらかじめ用意した低ピッチと高ピッチの 2 種類の刺激音を本実験の前に比較しながら, 5 つの音響的パラメータについてそれぞれの特徴を観察する。

知覚実験 I で使用した 30 個の高ピッチ刺激音と 30 個の低ピッチ刺激音, それぞれの VOT, 後続母音長, 後続母音の f_0 をまとめたものが, 表 8 である。なお, 刺激音は録音語から第 2 音節を切り出した 1 音節のものであるため, 先行研究で挙げられた音響的特徴「②閉鎖区間の長さ」は存在しない。

表 8 刺激音の音響的特徴

	ピッチ H	ピッチ L
VOT	12.3 msec	7.4 msec
後続母音の長さ	220.0 msec	85.5 msec
後続母音の f_0	120.5 Hz	84.16 Hz

次に, 閉鎖区間を除き, VOT と後続母音長において, 高ピッチと低ピッチの間に差があるかどうかを確認するため, 対応なしの t 検定をそれぞれ行った。その結果, VOT と後続母音の長さの両方で 2 種類(高ピッチと低ピッチ)の刺激音の間に有意な差が検出された。

VOTの結果は $t(58)=-5.09, p=.002$ で、後続母音の長さの結果は $t(58)=-25.18, p=.003$ 、高ピッチと低ピッチ2種類の刺激音の間に有意差があることがわかった。この結果から、以下の知覚実験 I では、刺激音を「高ピッチ」と「低ピッチ」との2つのグループに分けて、それぞれの結果を考察する。

3.4 知覚実験の方法

音声知覚実験には、日本語母語話者（近畿圏出身）20名に協力してもらった。実験を実施した時実験協力者の平均年齢は28.7歳（SD=1.93）である。実験には、ノートパソコン（Lenovo G500 15.6型）、ヘッドフォン（SONY MDR-10RNC, Audio-technica ATH-ANC9）を用いて、刺激音ごとの聴き取りによる正誤判断課題を実施した。実験制御ソフトとして、Hot Soup Processor（HSP）Ver.3.4⁸を使用し、コンピュータからランダムに音声提示した。実験は、図書館などの静かな場所で個別に行った。

実験の流れは図22の通りである。コンピュータの画面中央に凝視点「*」を600 msec 視覚提示し（図22の①）、凝視点が消えると同時に、平仮名の視覚提示と刺激音の聴覚提示が同時に現れる（図22の②）。聴覚提示は1 secの間隔で2回提示される。被験者にはヘッドフォンを通して聴覚提示された刺激音が画面に現れた仮名表記の語彙であるかどうかを判断するよう求めた。当該の刺激音が提示された視覚情報と一致している場合はキーボード上の「○」のボタンを、一致していない場合は「×」のボタンをクリックするよう指示した。ボタンをクリックすると、図22の③に示すように、1問の判定が終了し、次の問題提示のサイクルが始まる。判断の正誤を無声音の知覚率として測定した。なお、被験者に本番の実験を始める前に操作に慣れてもらうため、10問の練習問題を行った。刺激音はコンピュータからランダム順に提示された。この実験のトータル所要時間は約8分であった。

⁸ 実験に用いたこのソフトはインターネットから無料でダウンロードできるようになっている（このソフトの詳細 <http://hsp.tv/idman/download.html> を参照）。

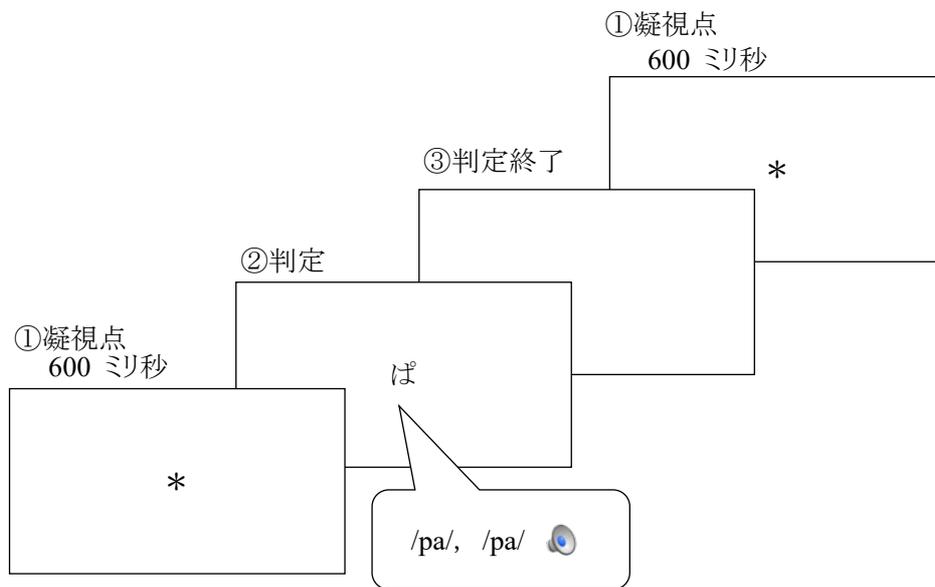


図 22 実験の流れ

3.5 結果と考察

図 23 は、ピッチ別に録音語から切り出した第 2 音節の「パ」を日本語母語話者が無声音「パ」と知覚した割合を表したものである。高ピッチと低ピッチの知覚率はそれぞれ 98.0% と 8.3%であった。

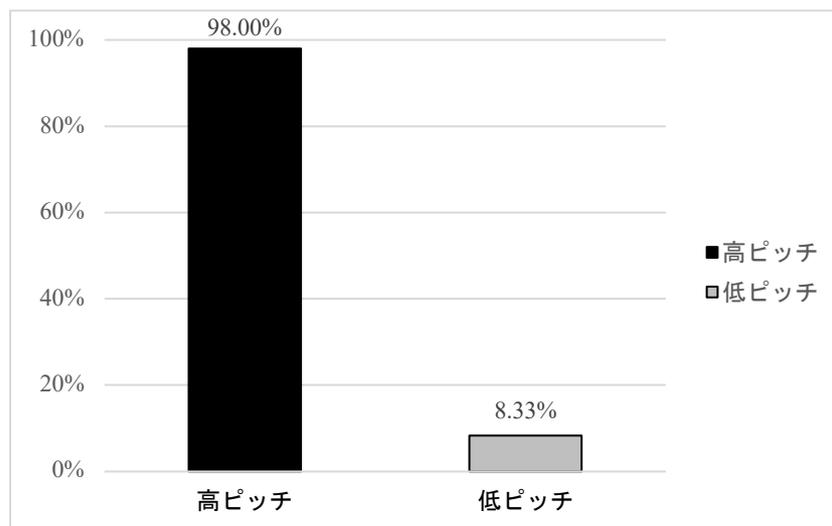


図 23 ピッチ別の「パ」を無声音とした知覚率 (%)

次に、高ピッチと低ピッチの知覚率の差を、対応ありの t 検定で検討した。その結果は、 $t(98)=-46.48, p<.001, d=1.26$ で、高ピッチと低ピッチの間に有意な差が確認された。また、効果量 (Cohen's d) も大きく、これにより、録音語 (「アパ」など) から切り出した無声音「パ」を日本語母語話者が有声音と知覚する割合が、低ピッチの方が有意に高いことが分かった。

先行研究によれば、語中破裂音の有声性の弁別においては閉鎖区間 (内破) の持続時間が関係しており、その持続時間が短くなると有声音と知覚されやすくなることが報告されている (杉藤・神田 1987, 朱 1994, 2010)。これを踏まえると、知覚実験 I で使用した刺激音においては閉鎖区間が存在しないため、日本語母語話者は切り出した無声音を有声音として知覚すると予測できる。しかし、図 23 の結果を見ると、高ピッチの場合、日本語母語話者が有声音と知覚することはほとんどなかった。すなわち、知覚実験 I で使用した刺激音は、語中位置を変えたもの (語中から語頭に移動した) なので、従来の先行研究で指摘された語中無声音の有声性の弁別と同じように解釈することはできない。

以上の結果を踏まえて考えると、既に先行研究で指摘された他の有声・無声音の弁別要素、つまり VOT・後続母音の持続時間、と後続母音の f_0 などにも検討されなければならない。「低ピッチ」グループの刺激音の有声音と知覚された理由は、VOT が短いのか、後続母音の持続時間が短いのか、またはピッチが低いからなのかどうか、このような複数の音響的パラメータが関与している可能性が出た以上、低ピッチの刺激音と高ピッチの刺激音の音響的特徴の違いを考慮し、さらに議論を進める必要があると考える。

3.6 まとめ

知覚実験 I では、日本語母語話者を対象として、日本語母語話者が発話された 2 音節語の語中から切り出した無声音が、単独では有声音と知覚するかどうかを検証するために、正誤判断課題による知覚実験を行った。その結果、録音語から第 2 音節を切り出して単独で聴かせると、音節頭の無声音が有声音と知覚することが確認できた。すなわち、筆者の指導教員の朱春躍教授が指摘した「文末の無声破裂音を含む音節を切り出して単独で聴かせると、日本語母語話者 (発話者本人を含む) は、無声破裂音を有声破裂音と知覚する」現象は、確実に存在していることが確認できた。また、この現象から「無声音として発声した無声無気音」が有声音に聴こえる結果に関しては、従来の先行研究で指摘されていた中国語系日本語学習者にとって母語干渉による知覚混同だけではなく、日本語母語話者に

とって普遍的な現象であることが確認できた。

さらに、知覚実験の結果をまとめると、切り出した第2音節の音節頭の無声音を無声音と聴く知覚率の平均値は、低ピッチの場合は7.8%で、高ピッチのほうは98.2%であった。また、どの刺激音も「アパ」、「イパ」、「ウパ」、「エパ」、「オパ」から切り出した第2音節の「パ」であったが、「低ピッチ」と「高ピッチ」グループの間に有意差が確認された。しかし、3.5.1にも述べたように、本実験に使用した刺激音には複数の異なる音響的パラメータが存在している。音響的に無声の（声帯振動が破裂後に始まる、+VOT値を持つ）破裂音が果たして後続母音の f_0 の影響で、有声音または無声音と知覚することであるかどうかについてさらに検討する余地がある。

日本語母語話者が頭高型アクセントの録音語（例：「アパ（HL）」）から切り出した低ピッチの刺激音（「パ（L）」）の音節頭無声音を有声音として知覚したことに対して、平板型アクセントの録音語（例：「アパ（LH）」）から切り出した高ピッチの刺激音（「パ（H）」）を無声音と知覚する。知覚実験Iの結果を踏まえて、ピッチの高・低が単音節の頭子音の有声・無声の知覚に大きく関わっていることを意味すると考えられる。

しかし、本実験に用いられた高ピッチと低ピッチの刺激音「パ」は、3.3.1節に示した通り、その内部にある音響的特徴として、後続母音の f_0 のほかにも、音響的パラメータに秘められた要因があるとの可能性は否定できない。低ピッチの刺激音と高ピッチの刺激音の音響的パラメータをさらに比較検討し、 f_0 の差異だけで有聲・無声の知覚が決まるのか、それとも他の音響的パラメータがより重要なのかをさらに検証する必要がある。

以上のように知覚実験Iで得られた結果と刺激音の各音声特徴を表9にまとめた。

表 9 知覚実験 I の結果と刺激音の音声特徴

	「パ」	「パ」
切り出す前の知覚	無声音「パ」	無声音「パ」
切り出した後の知覚	無声音「パ」	有声音「パ」
ピッチ	高い	低い
VOT	長い	短い
後続母音長	長い	短い

知覚実験 I の結果から、切り出した音節中の無声音を有声音に知覚する要因として、まずピッチの影響が挙げられるが、表 9 に示したように、他の音響的特徴が関与していることも排除できない。現段階で考えるのは、VOT と後続母音長である。すると、ピッチの高・低が有聲・無声判定の決定的要因であるかどうかについては、追加実験を用いて検証する必要がある。そこで、次のステップとして、VOT の影響と後続母音長の影響、さらには後続母音の f_0 について検証することにした。それぞれ第 4 章知覚実験 II、第 5 章知覚実験 III、第 6 章知覚実験 IV として報告する。

第4章 知覚実験 II : VOT の影響

4.1 実験の背景と目的

第3章の知覚実験 I の結果から、語中から切り出した無声音「パ」を、日本語母語話者に無声音と知覚されるものがあるが、有声音と知覚されるものもある。この結果から、刺激音の VOT の平均がグループにより違いがあり、無声音と知覚された「高ピッチ」グループの方が、有声音と知覚された「低ピッチ」グループより有意に長いことがわかった。そこで、知覚実験 II では、語中から切り出した無声子音を含む音節を知覚する際、VOT の長さの変化が知覚にどのような影響を与えるかについて検証する。この知覚実験 II を通して、主に以下の2点を観察していく。

- (1) 「高ピッチ (H)」グループの無声音が無声音と知覚したのは、VOT が「低ピッチ (L)」グループの無声子音より長いのが原因であるかどうか、VOT を短くして検証する。
- (2) (1) との対照として、有声音と知覚した「低ピッチ (L)」グループの無声音は、VOT を長くすると無声音に知覚するかどうかを確認。

4.2 刺激音の作成

第3章の知覚実験 I で使用した両グループの刺激音から無作為に5つずつを選択した。選択した「パ」について、まずは音響分析ソフトである Praat 6.0.1 用い、音声波形と広帯域スペクトログラムを参考にして VOT (破裂から声帯振動開始時間までの時間) を測定した。

VOT を測定した上、Audacity 2.2.3⁹を利用して、オリジナル音声の VOT の持続時間を3倍、6倍と増減加工をした。音声を加工する際に、Audacity 2.2.3 のエフェクト効果から「change tempo」の機能を利用し、選択した VOT 区間を加工した。Audacity 2.2.3 の説明によると、この「change tempo」機能は選択された音声のピッチが変更されずにテンポ (tempo) を編集できる。一方、もしピッチと速度と同時に変更する場合、「change speed」の機能もあるが、本実験では、加工対象である VOT がピッチと関係のない無音区間であるが、速度を制御するため、「change tempo」を利用した。知覚実験 I で無声音と知覚された「高ピッチ

⁹ 加工音声に用いたこのソフトはインターネットから無料でダウンロードできるようになっている (このソフトの詳細 <https://audacity.onl/download-audacity-2-2-3/>を参照)。

チ」グループの刺激音「パ」の加工した結果と、有声音と知覚された「低ピッチ」グループの刺激音「パ」の加工結果を表 10 と表 11 に示す。

表 10 無声音と知覚した「パ」(H) の VOT 加工結果 (単位 : msec)

NO.	刺激音	-6 倍	-3 倍	無加工	+3 倍	+6 倍
1	パ	2.2	4.5	13.6	40.0	81.6
2	パ	2.4	4.8	14.3	42.9	85.8
3	パ	2.4	4.8	14.4	43.2	86.4
4	パ	2	4	12.1	36.3	72.6
5	パ	2.1	4.2	12.7	38.1	76.2

表 11 有声音と知覚した「パ」(L) の VOT 加工結果 (単位 : msec)

NO.	刺激音	-6 倍	-3 倍	無加工	+3 倍	+6 倍
6	パ	1.9	3.8	11.5	34.5	69
7	パ	2	4	12.1	36.3	72.6
8	パ	1.8	3.7	11	33	66
9	パ	1.8	3.7	11	33	66
10	パ	1.6	3.3	10	30	60

また、それぞれ加工した刺激音の具体的な音声波形図は図 24～図 28 に示す。

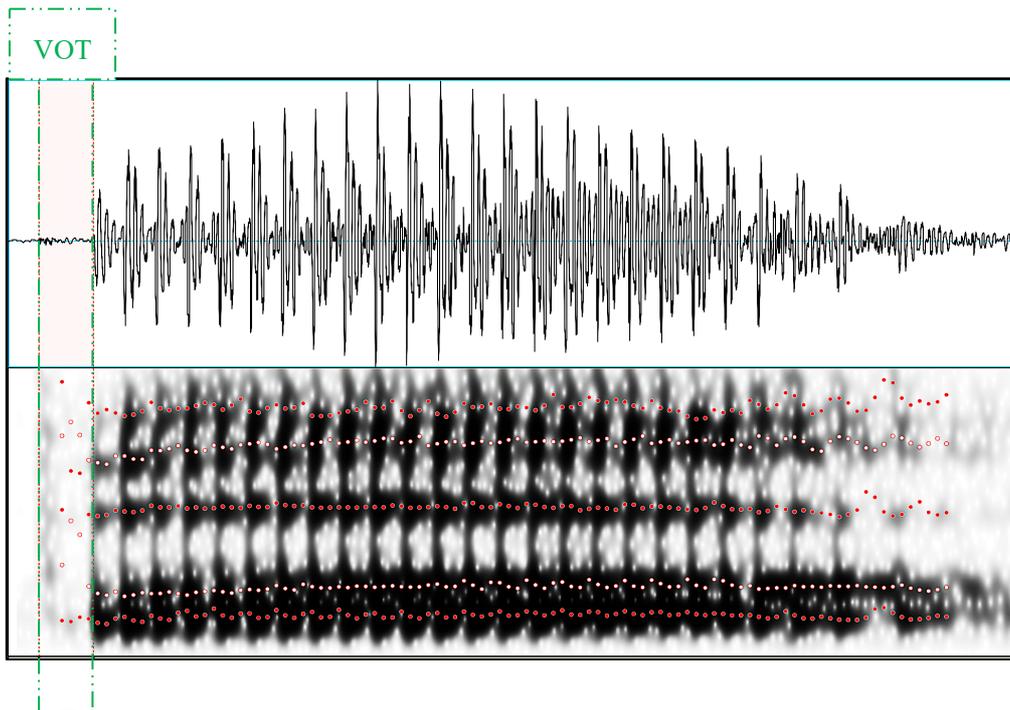


図 24 オリジナル音声の音声波形・スペクトログラム (VOT=13.7 msec)

[表 10 より No.1 無加工の刺激音を参照]

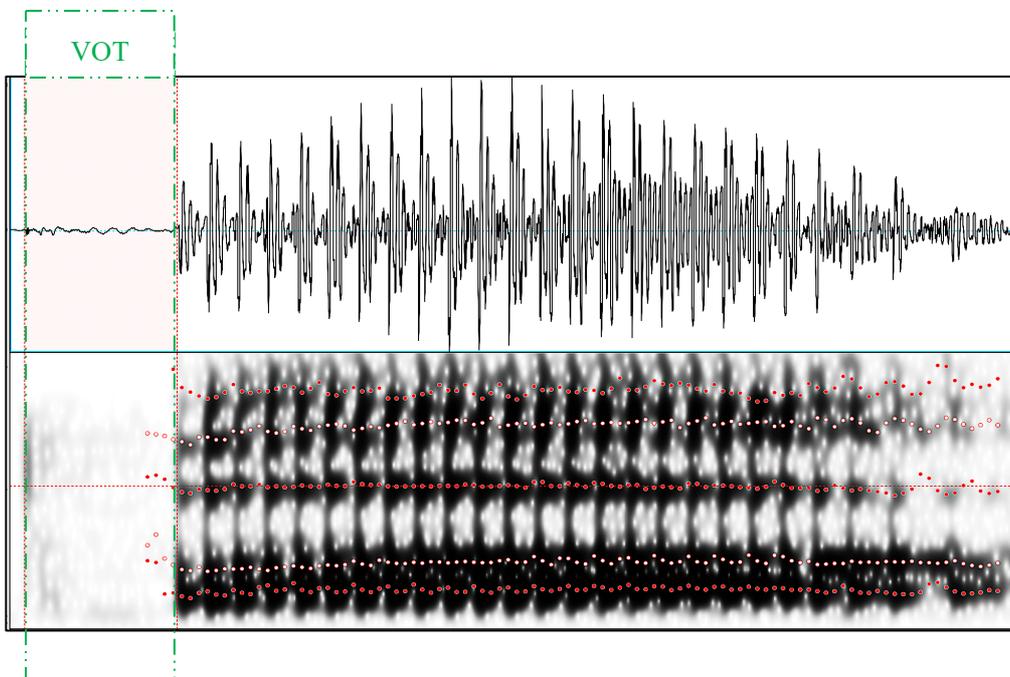


図 25 VOT+3 倍の音声波形・スペクトログラム (VOT=40.0 msec)

[表 10 より No.1+3 倍の刺激音を参照]

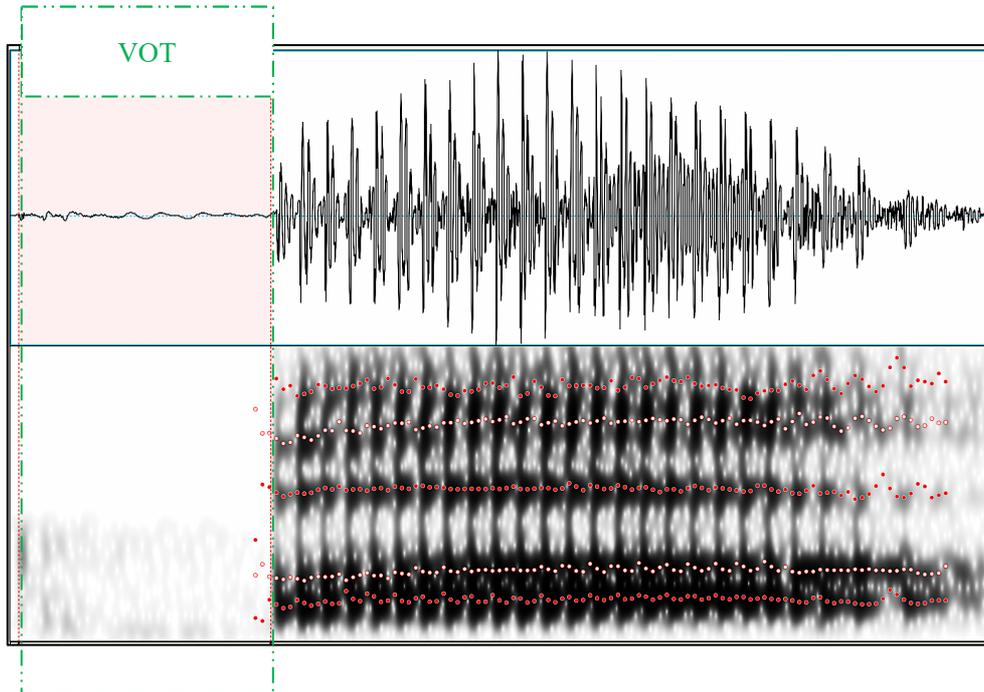


図 26 VOT+6 倍の音声波形・スペクトログラム (VOT=81.3 msec)

[表 10 より No.1+6 倍の刺激音を参照]

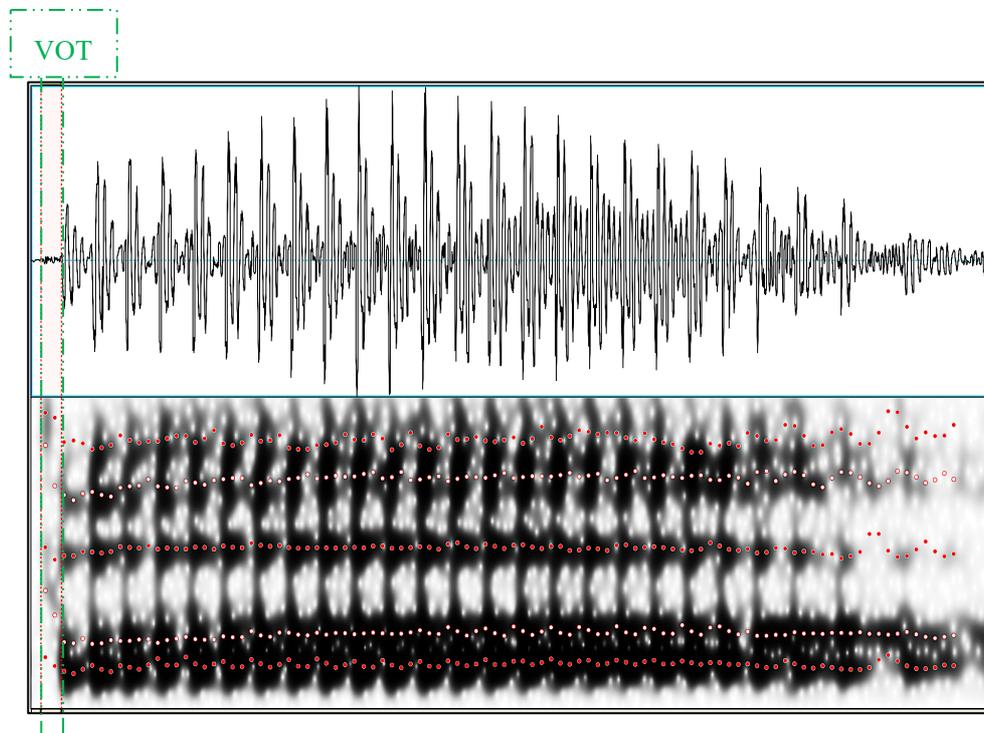


図 27 VOT-3 倍の音声波形・スペクトログラム (VOT=4.5 msec)

[表 10 より No.1-3 倍の刺激音を参照]

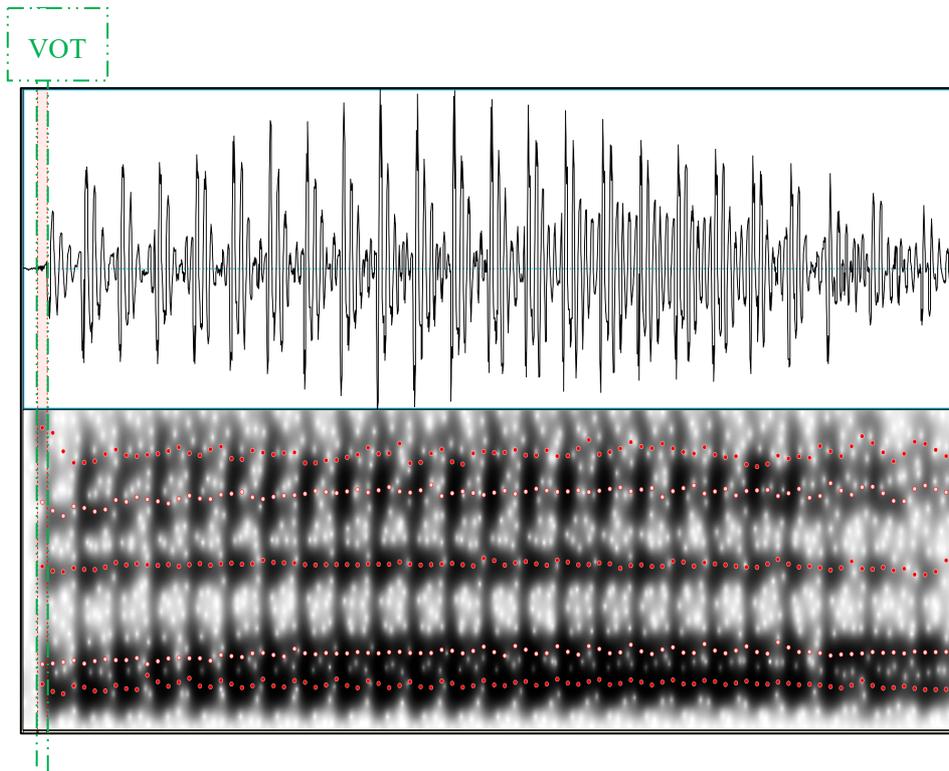


図 28 VOT-6 倍の音声波形・スペクトログラム (VOT=2.2 msec)

[表 10 より No.1-6 倍の刺激音を参照]

刺激音の VOT 加工倍率について、田村・森田・伊藤・廣瀬・森 (2016) では、VOT を連続的に変化させた合成音声/da/-/ta/を用いて知覚実験を行った結果、日本語母語話者が VOT 15~20 msec を境に知覚が/da/から/ta/へと急激に切り替わると指摘されている。この結果を踏まえ、知覚実験 II で使用する刺激音の VOT が+6 倍と+3 倍に伸長すると、先行研究で指摘された知覚境界以上の数値になるため、この実験の目的である VOT の影響があるかどうかの確認には適切であると思われる。

以上の作成方法で得た刺激音は、加工材料の「パ」(5 個) × 高・低ピッチ (2 グループ) × VOT の加工タイプ (5 種) の計 50 個となるが、さらに、実験に無関係の 50 個のダミー (例:「ア」,「ナ」など) を混ぜ込み、計 100 個の刺激音を知覚実験 II の実験材料とした。

4.3 知覚実験の方法

この知覚実験には、近畿圏出身の日本語母語話者 10 名に協力してもらった。実験協力者の内、女性 6 名、男性 4 名であった。実験を実施した時の平均年齢は 17.9 歳 (SD=0.74)

であった。

実験には、ノートパソコン (Lenovo G500 15.6 型)、ヘッドフォン (SONY MDR-10RNC, Audio-technica ATH-ANC9) を用いて、刺激音ごとの聴き取りによる正誤判断課題を実施した。

実験方法と実験制御ソフトは、第 3 章知覚実験 I の 3.4 節で紹介したと同様に、実験制御ソフトとして、Hot Soup Processor (HSP) Ver.3.4 を使用し、コンピュータからランダムに音声呈示した。この実験のトータル所要時間は約 7 分であった。

4.4 実験の結果

VOT の持続時間を伸縮加工した 2 グループ (高ピッチ・低ピッチ) の刺激音「パ」を実験協力者に判定させた結果、頭子音が無声音の[pa]と有声音の[ba]の両方に知覚された。グループ毎に無声音に知覚された割合についての結果を、図 29 と図 30 に示している。縦軸は刺激音が無声音としての知覚率 (%), 横軸はそれぞれのトークンの VOT 伸縮加工の倍率を示す。

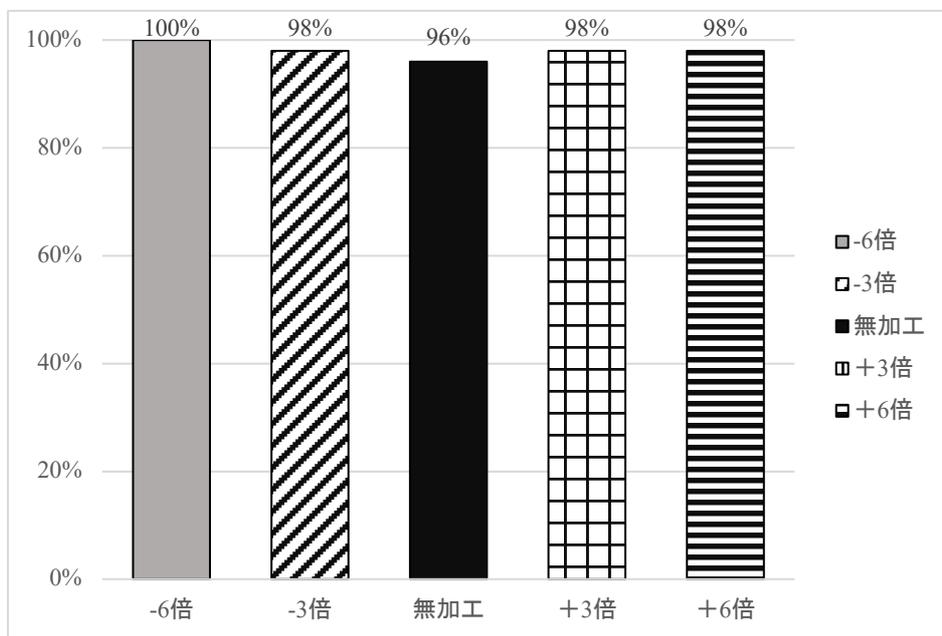


図 29 高ピッチ「パ」の無声音知覚率 (%)

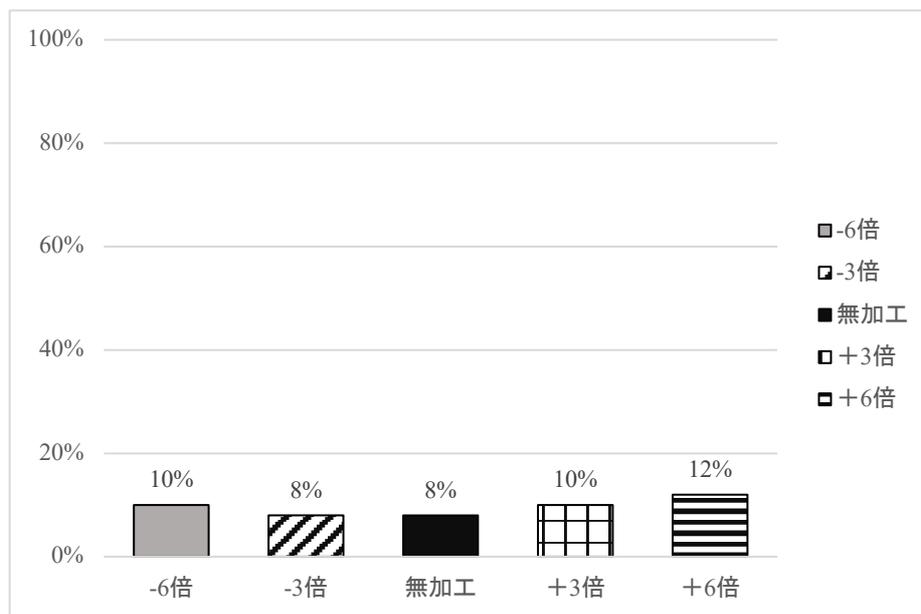


図 30 低ピッチ「パ」の無声音知覚率 (%)

VOTの伸縮加工による「高ピッチ」と「低ピッチ」グループの「パ」の知覚に差があるかどうかについて、それぞれ対応ありの一元配置分散分析で検討した。

「高ピッチ」グループに対する分析結果では、有意差が検出されなかった ($F(4,196) = .495, p = .739 > 0.05$)。一方、「低ピッチ」グループの刺激音に対する知覚実験の結果についても、「高ピッチ」グループと同じように一元配置分散分析を行ったが、有意な差は見られなかった ($F(4,196) = .696, p = .508 > 0.05$)。

以上の結果から、日本語母語話者の音響的無声音「パ」の知覚は、VOTの持続時間長とは直接関係がないと考えられる。

4.5 考察とまとめ

知覚実験 II では、無声音を有声音に知覚するケースの VOT 持続時間の要因を考慮し、VOT を伸縮加工した刺激音を用いて実施した知覚実験の結果、無声音・有声音の判断は VOT の持続時間長と関係ないことが明らかとなった。

従来、破裂音における有声・無声音は、VOT の長さによって識別されると言われてきた (Linsker & Abramson 1964)。日本語の破裂音においても、杉藤・神田 (1987) は有声音の VOT が無声音より短いと指摘されている。さらに、日本語母語話者が VOT による日本語の有声・無声音の識別境界について、語頭の場合は、6 msec 以下の無気音が有声音に知覚

されることに対して、語中の場合は、有声音の VOT が 20 msec 以上を超えると、有声音と知覚する正聴率が低下することがある（朱 1994, 2010）。また、田村・森田・伊藤・廣瀬・森（2016）にも、VOT が 15～20 msec を境に、日本語母語話者の知覚が有声音の/da/から無声音の/ta/へと急激に切り替わると報告されている。以上のいくつかの先行研究では、日本語の有声・無声音の弁別において、VOT の持続時間は少なくとも 1 つの重要な音響的パラメータと考えられると指摘している。しかし、この章で紹介した知覚実験 II においては、VOT 持続時間の伸縮加工をしても、知覚の結果が加工前と変わらないことが明らかになった。すなわち、本聴取実験の結果から、単音節語において、VOT の長さは日本語の有声性の弁別に積極的に関わっていないことが確認された。この結果については、杉藤・神田（1987: 16）で指摘されたように、日本語の有声・無声破裂音の知覚に関する特徴は「必ずしも VOT が重要とは言えない」との結論と一致した。

さらに、本実験で使用した加工前のオリジナル音声は無声無気音であり、先行研究で指摘されたように、中国語を母語とする日本語学習者が有声音に間違えて聴きやすいタイプである。このことから考えると、日本語母語話者でも学習者と同様に、VOT の値が中国語の無気音と類似する日本語の無声音を、一定の条件で必ずしも無声音と知覚するとも限らない。

以上の結果から、本研究の目的である「無声音として発声した無声無気音」を有声音に聴こえる現象の要因として VOT の持続時間と直接に関係がないと考えられる。さらに、高田（2004）と朱（2010）が報告した「日本語母語話者は無声音化した語頭有声音を有声音と知覚する」、すなわち「有声音として発声した無声無気音」の知覚結果を踏まえると、日本語母語話者の有聲・無声の弁別特徴に関しては、声帯振動の有無だけではなく、VOT の持続時間長も手がかりとなっていない可能性が示唆された。破裂音のカテゴリー知覚にかかわる最も重要な音響的特徴はおそらく +VOT の持続時間ではなく、ほかにあると推測される。それについては、後続母音長を中心に次章で考察する。

第5章 知覚実験 III：後続母音長の影響

5.1 実験の背景と目的

第3章の知覚実験 I の結果から、語中から切り出した2グループの刺激音（高ピッチ・低ピッチ）の母音における持続時間が異なり、無声音と知覚された「高ピッチ」グループの刺激音の方が、有声音と知覚された「低ピッチ」グループより長いとわかった。そこで、知覚実験 III では、語中から切り出した無声子音を含む音節を、後続母音の持続時間の要因で、無声音を有声音に知覚することがあるかどうかについて検証する。

この知覚実験 III では、知覚実験 II と同じ無加工音声/materialを材料に、後続母音の持続時間を伸縮して作成した刺激音を日本語母語話者に聴かせ、有声音に聴くかどうかの知覚特徴を考察する。観察する内容は以下の2点である。

- (1) 「高ピッチ (H)」グループの無声音が無声音と知覚したのは、後続母音が「低ピッチ (L)」グループの無声子音より長いのが原因であるかどうか、後続母音を短くして検証する。
- (2) (1) との対照として、有声音と知覚した「低ピッチ (L)」グループの無声音は、後続母音を長くすると無声音に知覚するかどうかを確認。

5.2 刺激音の作成

本調査では、日本語母語話者を対象とし、後続母音の長・短が、音節頭無声音[p]の知覚にどの影響を与えるかについて、知覚実験により確認する。刺激音の作成については、第4章の知覚実験 II と同じ無加工音声、つまり知覚実験 I の2グループ（高ピッチ・低ピッチ）から5個ずつ選択した無加工音声/voiceをベースとして作成する。刺激音の作成手順、利用した音声編集ソフト、音響分析の方法は4.2節の刺激音作成と同様である。

刺激音の加工手順については、音響分析ソフト Praat 6.0.16 で音声波形と広帯域スペクトログラムを参考にして、2.2.1 の図7で紹介した後続母音区間を測定した。具体例は図31に示したように、後続母音区間の測定については、母音の声帯振動の開始時点を始点とし、第3フォルマント周波数 (F_3) が消える時を終点と判定する。

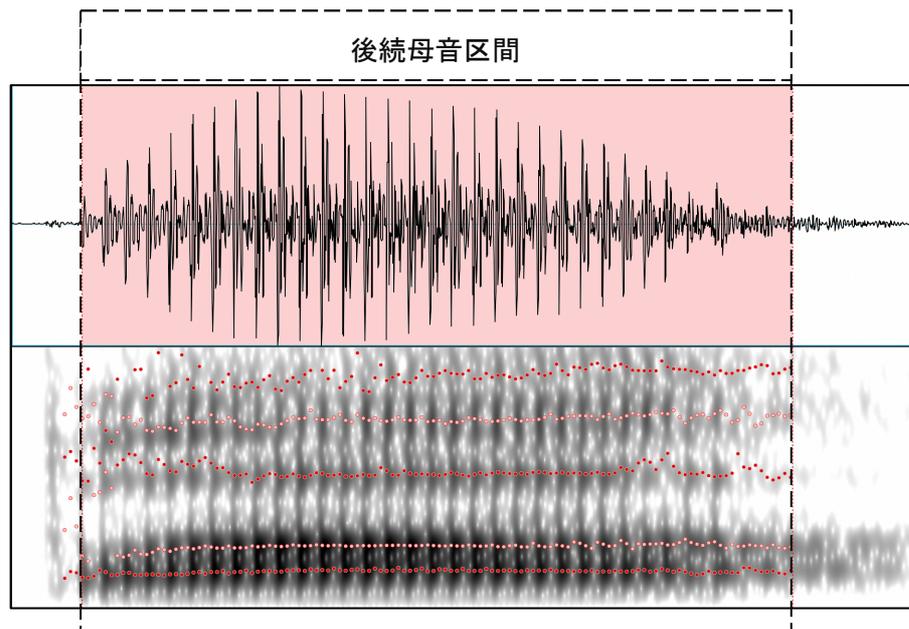


図 31 後続母音区間の測定

Praat で測定した後続母音の持続時間を、音声編集ソフト Audacity2.2.3 で 2 倍、3 倍と増減加工をした。音声の加工については、4.2 節の説明の通り、Audacity2.2.3 の「change tempo」機能を利用し、ピッチを変えずに後続母音の区間のみに対して持続時間を編集した。また、前章の知覚実験 II では、VOT は無音区間であったためピッチは存在しないが、本章はピッチがある後続母音であるため、編集する際に、ピッチが持続時間の変動により変わることがないように、これまで以上に音響データの編集方法に注意しなければならない。その理由については、音声のエフェクト効果を利用する際に、通常スピードを遅くするとピッチが低くなり、スピードを速くするとピッチも高くなるからである。本論の知覚実験は、「持続時間の伸縮」の効果を観察するのが目的であるため、ピッチを変えてはならない。刺激音を作成する際に、ピッチの変動が起らないように特に注意しながら、後続母音の持続時間のみを加工した。

無加工音声の後続母音長とそれぞれの倍率で伸縮操作をした加工結果を表 12 と表 13 にまとめた。表 12 は無声音と知覚された「高ピッチ」グループの加工結果で、表 13 は有声音と知覚された「低ピッチ」グループの加工結果を示す。

表 12 無声音と知覚した「パ」(H)の母音持続時間の加工結果(単位:msec)

NO.	刺激音	+3倍	+2倍	無加工	-2倍	-3倍
1	パ	618.3	435.5	217.5	132.4	94.2
2	パ	759.9	535.2	267.3	162.7	115.7
3	パ	725.0	510.5	255.0	155.2	110.4
4	パ	774.1	545.2	272.3	165.8	117.9
5	パ	774.1	545.2	272.3	165.8	117.9

表 13 有声音と知覚した「パ」(L)の母音持続時間の加工結果(単位:msec)

NO.	刺激音	+3倍	+2倍	無加工	-2倍	-3倍
6	パ	413.1	290.9	145.3	88.5	62.9
7	パ	477.0	335.9	167.8	102.1	72.6
8	パ	436.4	307.3	153.5	93.4	66.5
9	パ	359.6	253.3	126.5	77.0	54.8
10	パ	409.1	288.1	143.9	87.6	62.3

加工した具体的な刺激音それぞれの音声波形とスペクトログラムは図32～図36に示す。

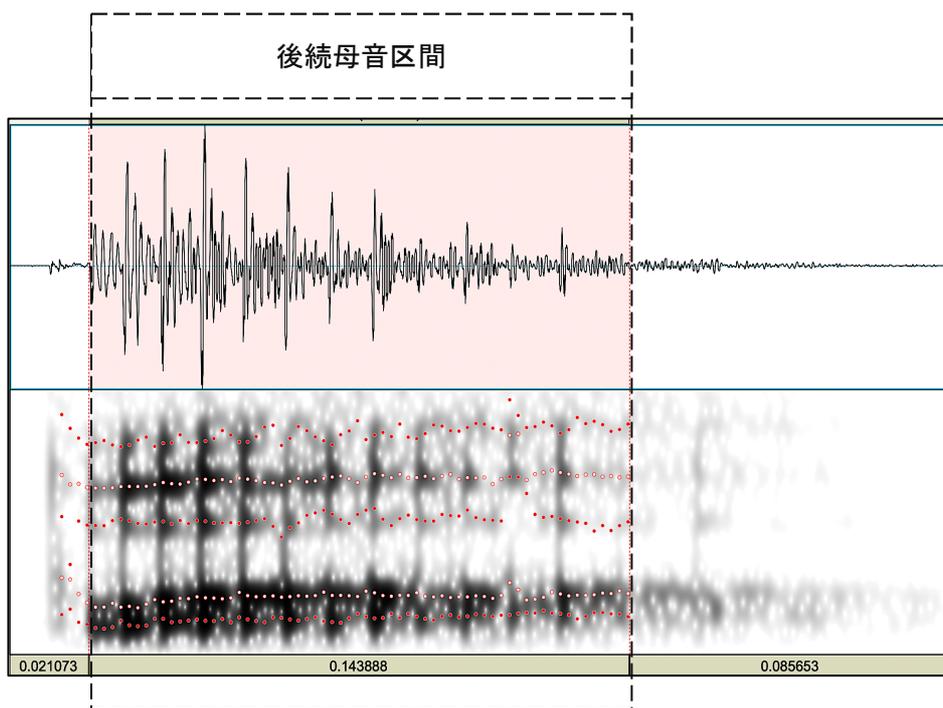


図 32 オリジナル音声の音声波形・スペクトログラム（後続母音長：143.888 msec）

[表 13 より No.10 無加工の刺激音を参照]

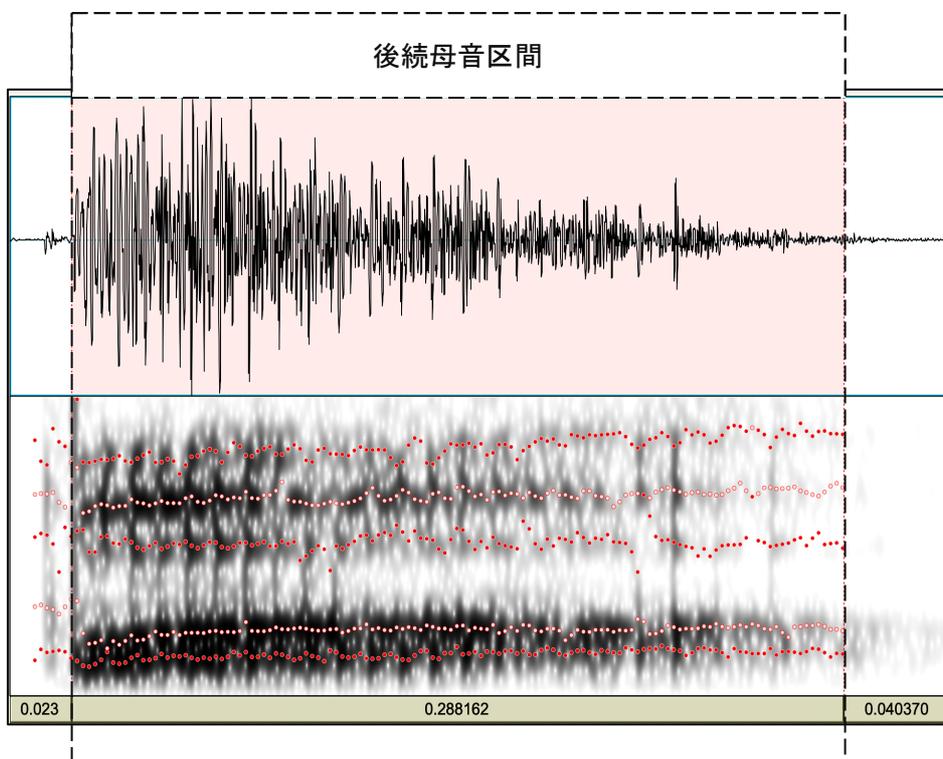


図 33 後続母音+2 倍の音声波形・スペクトログラム（後続母音長：288.162 msec）

[表 13 より No.10+2 倍の刺激音を参照]

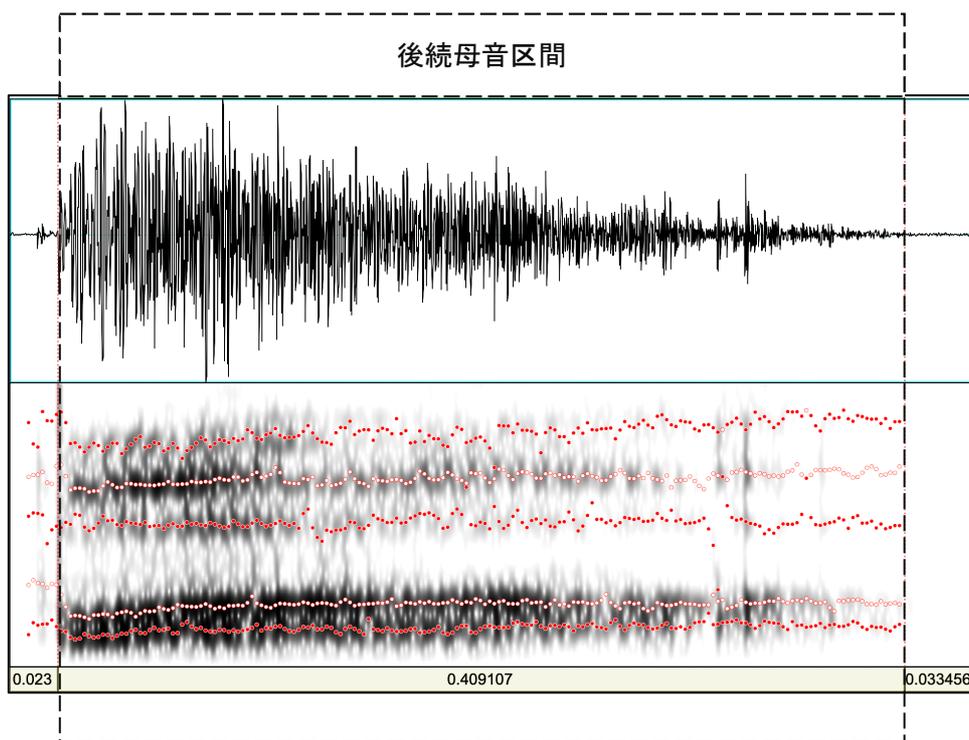


図 34 後続母音+3 倍の音声波形・スペクトログラム（後続母音長：409.107 msec）

[表 13 より No.10+3 倍の刺激音を参照]

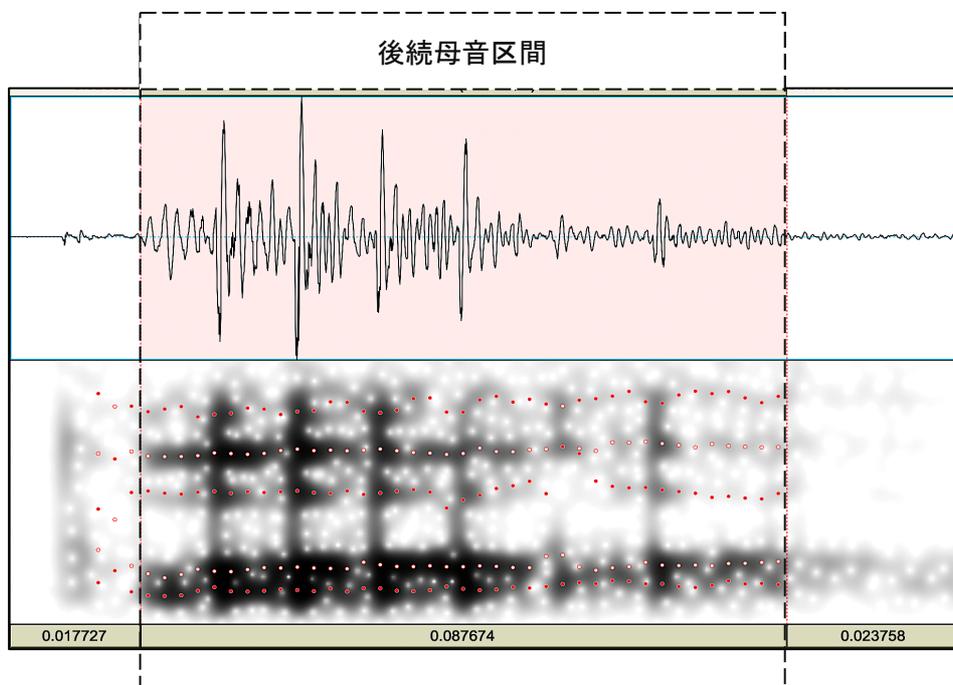


図 35 後続母音-2 倍の音声波形・スペクトログラム（後続母音長：87.674 msec）

[表 13 より No.10-2 倍の刺激音を参照]

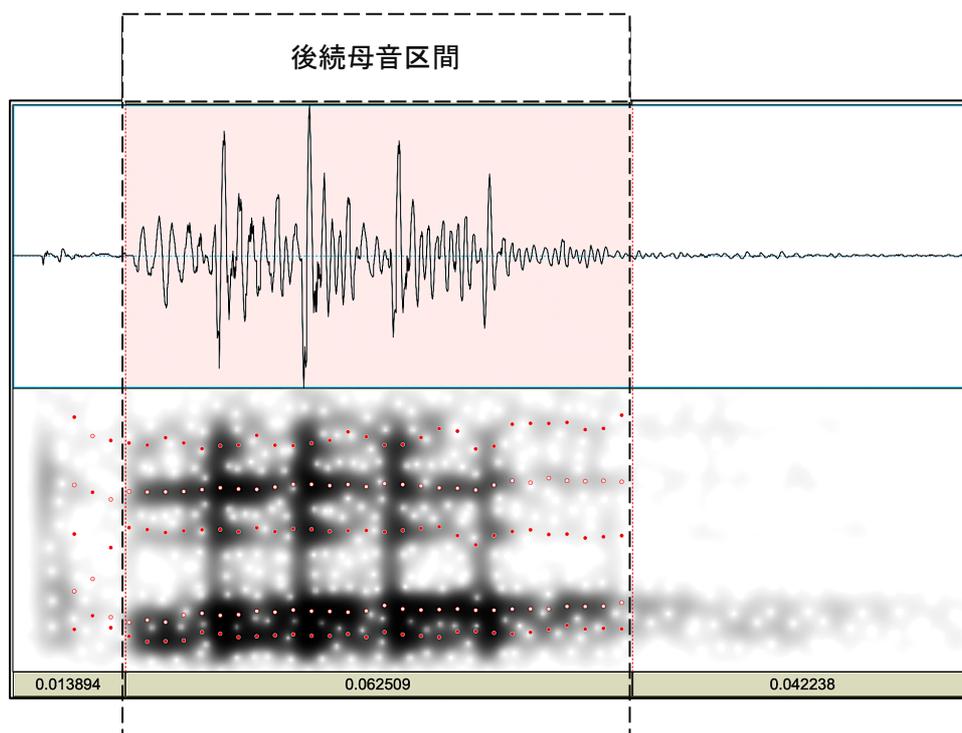


図 36 後続母音-3 倍の音声波形・スペクトログラム（後続母音長：62.3msec）

[表 16 より No.10-3 倍の刺激音を参照]

以上の作成方法で得た刺激音は、加工材料の「パ」（5 個）×高・低ピッチ（2 グループ）×後続母音の加工タイプ（5 種），計 50 個となるが，さらに，実験に無関係の 50 個のダミー（例：「ア」，「ナ」など）を混ぜ込み，最終的に計 100 個の刺激音を知覚実験 III の実験材料とした。

5.3 実験方法

この知覚実験には，第 4 章の知覚実験 II と同じ日本語母語話者 10 名に協力してもらった。実験協力者は女性 6 名，男性 4 名であった。実験を実施した時の平均年齢は 17.9 歳（SD=0.74）であった。

実験方法と実験制御ソフトは，第 3 章の知覚実験 I の 3.4 節で紹介したとおり，Hot Soup Processor（HSP）Ver.3.4 を使用し，コンピュータからランダムに音声呈示した。この実験のトータル所要時間は約 7 分であった。

5.4 実験の結果

後続母音の持続時間を伸縮加工した2グループの無声音「パ」(高ピッチ・低ピッチ)を、実験協力者に判定させた結果、無声音「パ」/有声音「バ」の両方に知覚された。グループ毎に無声音「パ」に知覚された割合は次の図37と図38に示す。縦軸は無声音としての知覚率(%), 横軸はそれぞれのトークンの後続母音の伸縮加工の倍率を示す。

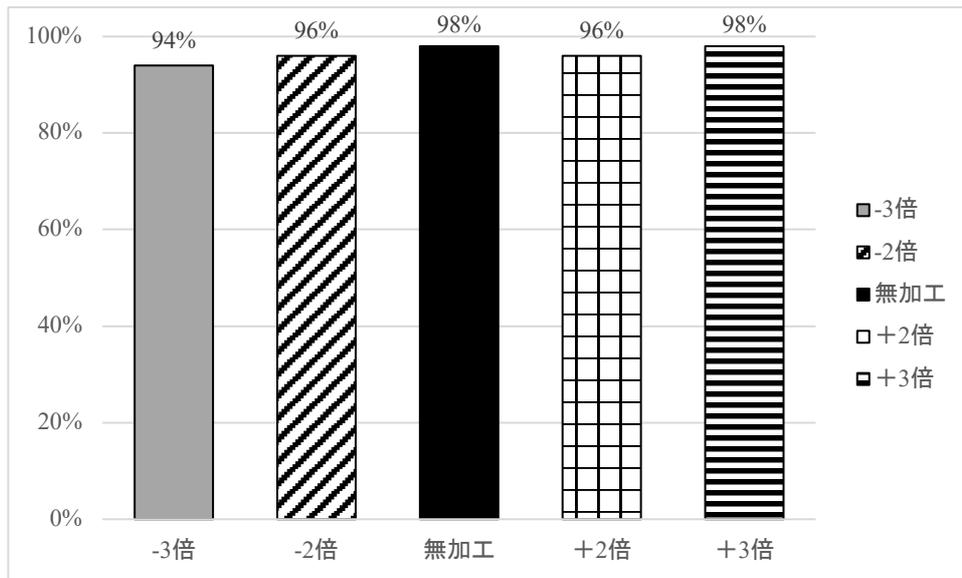


図37 高ピッチ「パ」の無声音知覚率 (%)

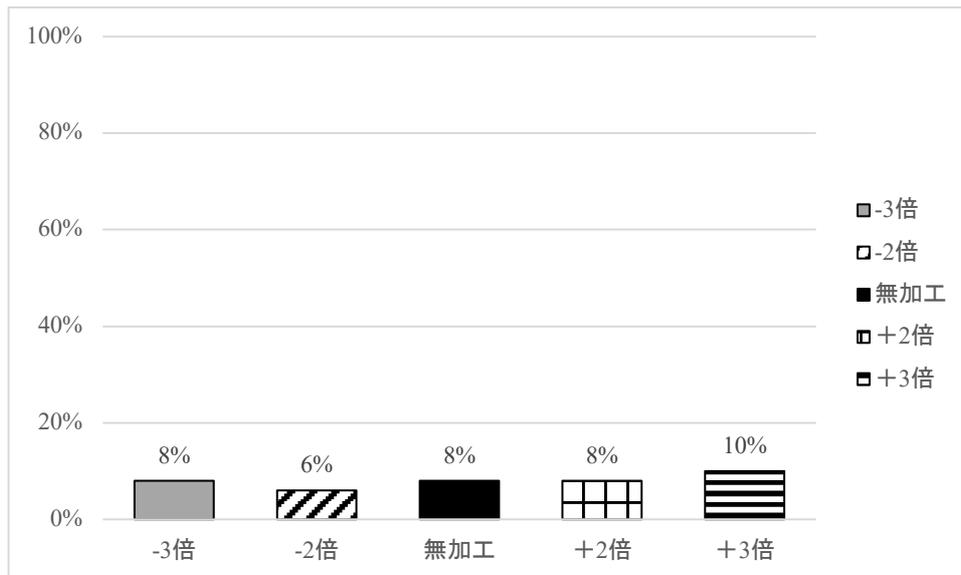


図38 低ピッチ「パ」の無声音知覚率 (%)

次に、後続母音の持続時間の伸縮加工による「高ピッチ」と「低ピッチ」グループの「パ」の知覚に差があるかどうかについて、それぞれ対応ありの一元配置分散分析を実行した。「高ピッチ」グループに対する分析結果では、 $F(4, 196) = 3.384$, $p = .820 > 0.05$ となっており、後続母音の伸縮加工倍率の間に有意の差が検出されなかった。一方、「低ピッチ」グループの刺激音に対する知覚実験の結果についても、「高ピッチ」グループと同じように分散分析を行なったが、その結果も同じく後続母音長の伸縮倍率の間に有意な差は見られなかった ($F(4, 196) = 3.53$, $p = .764 > 0.05$)。

以上のような結果から、日本語母語話者は、録音語（「アパ」など）から切り出した2種類（高ピッチ・低ピッチ）の無声音「パ」を無声音に聴く知覚率において、後続母音の長さが顕著に異なるにもかかわらず、結果は変わらなかった。すなわち、音響的無声音「パ」の知覚は、後続母音長とは直接関係がないと言えそうである。

5.5 考察とまとめ

知覚実験 III では、後続母音の持続時間の要因を考慮し、切り出した音節中の無声音を有声音に聴こえるかどうかについて加工音声を用いて検証した。その結果、後続母音の持続時間の伸縮と関係がないことが明らかとなった。

日本語破裂音の後続母音長について、Homma (1973) では、[piti] や [titi] などの刺激語を用いて、後続母音の実態を検証した結果、日本語の有声音の方が無声音より長いと指摘した。また、後続母音の持続時間はアクセント型と関係なく、語中にくる音節の場合は、破裂音の後続母音が語頭より若干短くなると指摘されている (Homma 1981:280)。本実験で用いた刺激音は語中から切り出した無声子音を含む音節、すなわち、Homma (1973, 1981) で指摘された後続母音が短いものである。

しかし、第3章の知覚実験 I において、平板型 (LH) 語から切り出した高ピッチ (H) と頭高型 (HL) 語から切り出した低ピッチ (L) の2種類の無声音を用意したが、高ピッチ音節の母音長が低ピッチ音節より長かった (3.3.1 節に表 8 を参照)。また、加工音声を用いた第4章の知覚実験 II と本章の知覚実験 III の結果、ピッチ (高・低) と後続母音の持続時間 (長・短) は関係がないことが明らかになった。すなわち、日本語母語話者は、知覚実験 I で無声音と判断された平板型 (LH) の録音語から切り出した語中にくる高ピッチの「パ」の母音を短くしても有声音と知覚されないほか、頭高型 (HL) の録音語から切り出した語中にくる低ピッチの「パ」も、母音を長くしても無声音と知覚されないと、本

章の知覚実験で検証できた。

以上の結果から、2種類の刺激音（高ピッチ・低ピッチ）の後続母音の持続時間に差異があっても、無声音を有声音に聴こえる決定的要因にはならないと言える。すると、ピッチの高・低が破裂音の知覚の際の範疇知覚にかかわる重要な判断材料となるかどうかは、後続母音の f_0 を中心に考察する必要があると考えられる。次章でその点について検討する。

第6章 知覚実験 IV : ピッチの影響

6.1 実験の背景と目的

第3章の知覚実験 I で残された課題は3つある。それは、語中から切り出した音節中の無声子音を有声音と知覚する判断材料が、①VOT の持続時間、②後続母音の持続時間、③ピッチの高低と関係があるかどうかである。筆者は切り出した音節中の無声音の①VOT と②後続母音の持続時間長を有聲・無声判断の条件として設定し、それぞれ第4章の知覚実験 II と第5章の知覚実験 III で確認した。その結果、語中から切り出した無声音の種類が「高ピッチ」と「低ピッチ」のどちらも、VOT と後続母音の持続時間の長短を操作しても、知覚結果は第3章の知覚実験 I と変わらなかった。すなわち、VOT と後続母音長は無声の刺激音を有声音と知覚する理由にはならないことが明らかとなった。

そこで、次の知覚実験 IV では、切り出した2種類（高ピッチと低ピッチ）の無声音を含む音節を加工音声に用いて、後続母音の f_0 の高・低による知覚特徴を主に以下の2点について観察していく。

- (1) 「高ピッチ (H)」グループの無声音が無声音と知覚したのは、その後続母音の f_0 が「低ピッチ (L)」グループのより高いのが原因であるかどうか、後続母音を低くして検証する。
- (2) (1) との対照として、有声音と知覚した「低ピッチ (L)」グループの無声音は、後続母音の f_0 を高くすると無声音に知覚するかどうかを確認する。

6.2 刺激音の作成

第3章の知覚実験 I で使用した高ピッチの60個、低ピッチの60個の刺激音「パ」をそれぞれのグループから6個ずつ選択した。選択の基準は、知覚実験 I で実験を協力してもらった20名日本語母語話者の知覚率と反応時間、両方の結果を参考にした。高ピッチの刺激音「パ」の場合、反応時間が短く、無声音としての知覚率が高いものを上位から6個選び、低ピッチの場合も、反応時間が短く、無声音の知覚率が低いものを上位から6個を選択した。選択した12個の「パ」の反応時間はすべて1000 msec 以内のものである。両グループの平均反応時間は、高ピッチの「パ」は551.2 msec で、低ピッチの「パ」は621.0 msec である。

なお、橋本・北川・樋口 (1998:177) では、「ATR 連続音声データベース」中の日本人

男性が発声した日本語文を対象に、個人性知覚に影響を与える音響特徴について ABX 方式¹⁰による聴取実験の結果、 f_0 とスペクトルに顕著な寄与が認められると指摘している。すると、本実験に使用する語中から切り出した高ピッチと低ピッチのオリジナル音声「パ」を加工する前に、種類毎に同じ音域に存在しているかどうかを確認することが必要である。刺激音を作成する前に、選択した 2 種類（高ピッチ・低ピッチ）の総数 12 個の「パ」の f_0 に対し、対応なしの一元配置分散分析で比較した。その結果、種類毎にそれぞれの「パ」は同じ音域にある（高ピッチグループの「パ」： $(F(5,36)=.116, p=.988 > 0.05)$ 、低ピッチグループの「パ」： $(F(5,36)=.016, p=1.000 > 0.05)$ ）。すなわち、後続母音の f_0 を同じ変更率で変えてもグループ内の刺激音同士はほぼ同じ音域にあることがわかり、実験の目的である「加工音声の高さ」が唯一の要因であることが確認できた。

次は、刺激音の作成において、Praat 6.0.1 の分析画面上で後続母音の波形が周期的になった後で、第 1・第 2 フォルマントの双方が観察され始める時点を測定した上、音声解析ソフト Adobe Audition 13.0.11¹¹を利用して加工音を作成した。音全体の f_0 を -6 半音（semitones；以降「st」と表示）から +6 半音まで 2 半音ずつシフトさせ、無加工のものも含め、合計 7 種類の刺激音を作成した。また、後続母音の f_0 を加工する際に、持続時間などの音響的パラメータが変わらないよう、十分注意した上、後続母音の f_0 のみ加工した。無声音と知覚された高ピッチの「パ」、有声音と知覚された低ピッチの「パ」に対して、後続母音の f_0 を加工した結果は表 14 と表 15 に示す通りである。

¹⁰ 橋本・北川・樋口（1998:170）によれば、A と B はそれぞれ話者 A、話者 B の分析合成音とし、X は話者 A、B 間で基本周波数、スペクトル、音素継続時間の各音響特徴を入れ換えた合成音声とした実験内容である。聴取実験では、被験者に「X の話者は、A、B のどちらの話者だと思えるか」を強制判断させた実験であった。

¹¹ Adobe audition は Adobe 社が開発した有料デジタルオーディオ編集ソフトである。（このソフトの詳細 <https://www.adobe.com/jp/products/audition.html> を参照）。

表 14 無声音と知覚した「パ」(H)の後続母音の f_0 の加工結果

No.	-6 st	-4 st	-2 st	無加工	+2 st	+4 st	+6 st
1	78.0 st	80.0 st	82.0 st	84.5 st	86.4 st	88.5 st	90.4 st
	90.1 Hz	101.1 Hz	113.8 Hz	127.9 Hz	143.7 Hz	160.7 Hz	180.3 Hz
2	77.3 st	79.3 st	81.3 st	83.3 st	85.3 st	87.4 st	89.3 st
	87.0 Hz	100.7 Hz	109.3 Hz	123.2 Hz	137.6 Hz	155.4 Hz	173.8 Hz
3	77.2 st	79.1 st	81.1 st	83.1 st	85.1 st	87.1 st	89.1 st
	86.3 Hz	96.6 Hz	109.4 Hz	121.6 Hz	136.4 Hz	153.0 Hz	171.6 Hz
4	76.4 st	78.4 st	80.5 st	82.5 st	84.4 st	86.4 st	88.3 st
	82.3 Hz	92.3 Hz	105.5 Hz	117.3 Hz	130.1 Hz	146.6 Hz	164.2 Hz
5	77.0 st	79.1 st	81.0 st	83.0 st	85.0 st	87.0 st	89.0 st
	85.4 Hz	101.5 Hz	107.3 Hz	120.4 Hz	135.1 Hz	151.6 Hz	170.2 Hz
6	76.8 st	78.8 st	80.8 st	82.8 st	84.8 st	86.8 st	88.8 st
	85.3 Hz	94.7 Hz	106.2 Hz	119.2 Hz	133.8 Hz	150.1 Hz	164.8 Hz

*1 st=100 Hz

表 15 有声音と知覚した「パ」(L) の後続母音の f_0 の加工結果

No.	-6 st	-4 st	-2 st	無加工	+2 st	+4 st	+6 st
7	71.7 st	73.7 st	75.7 st	77.6 st	79.6 st	81.4 st	81.4 st
	67.3 Hz	75.6 Hz	82.4 Hz	88.6 Hz	93.7 Hz	105.9 Hz	118.2 Hz
8	70.6 st	72.7 st	74.7 st	76.7 st	78.5 st	80.8 st	82.6 st
	69.3 Hz	73.4 Hz	76.0 Hz	84.0 Hz	92.6 Hz	106.5 Hz	117.7 Hz
9	71.1 st	73.1 st	75.1 st	77.1 st	79.0 st	81.1 st	83.0 st
	73.4 Hz	75.6 Hz	77.6 Hz	86.1 Hz	93.6 Hz	108.1 Hz	117.7 Hz
10	70.0 st	72.0 st	74.0 st	76.0 st	78.0 st	80.0 st	82.0 st
	73.4 Hz	77.8 Hz	82.4 Hz	87.3 Hz	90.1 Hz	101.6 Hz	106.9 Hz
11	71.0 st	73.0 st	75.1 st	77.0 st	79.1 st	81.2 st	83.0 st
	65.9 Hz	74.0 Hz	82.4 Hz	86.3 Hz	96.8 Hz	108.5 Hz	114.7 Hz
12	71.0 st	73.0 st	75.0 st	77.0 st	79.0 st	81.2 st	83.3 st
	69.3 Hz	73.4 Hz	76.0 Hz	86.8 Hz	96.0 Hz	107.6 Hz	120.0 Hz

*1 st=100 Hz

また、7種類の刺激音の具体的なピッチ曲線を重ね合わせて図 39 と図 40 に示す。

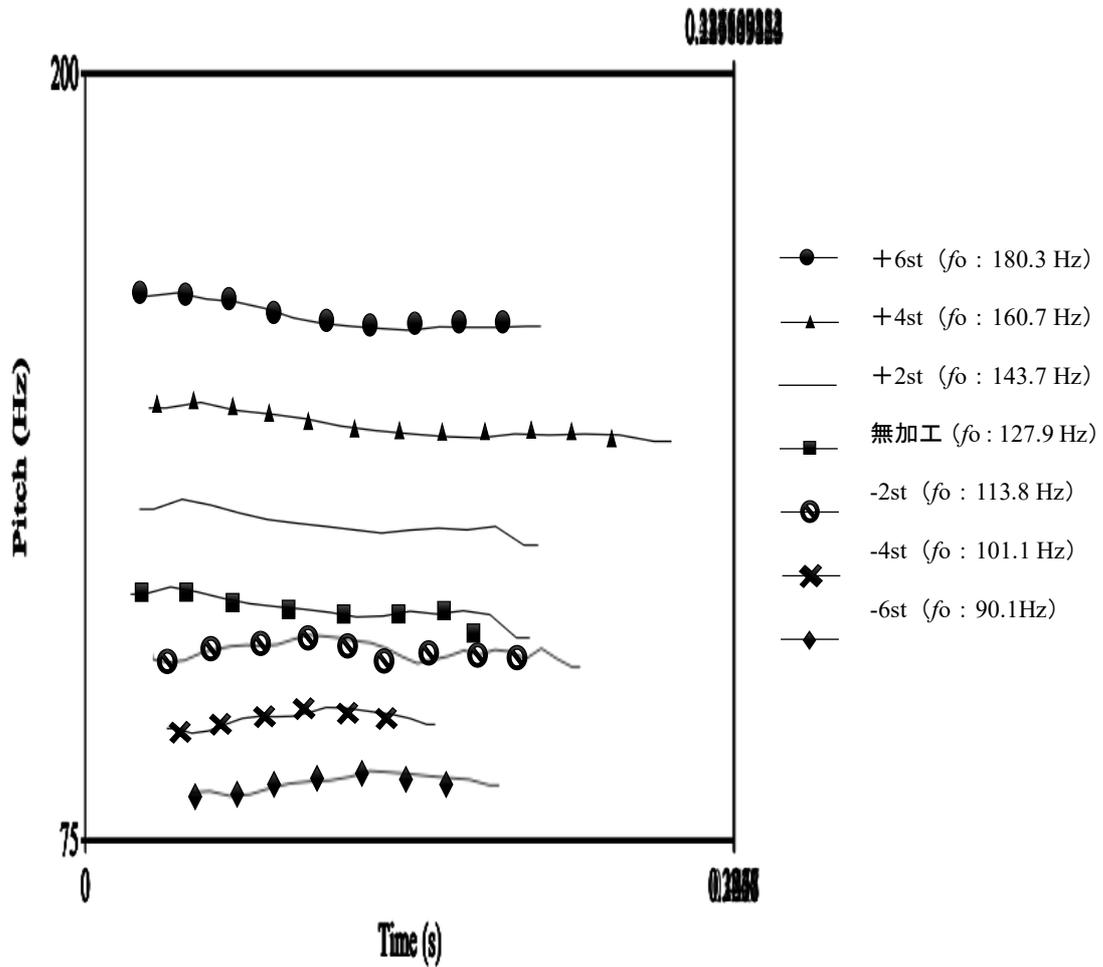


図 39 各加工タイプのピッチ曲線の具体例

[表 14 より No.1 の刺激音を参照]

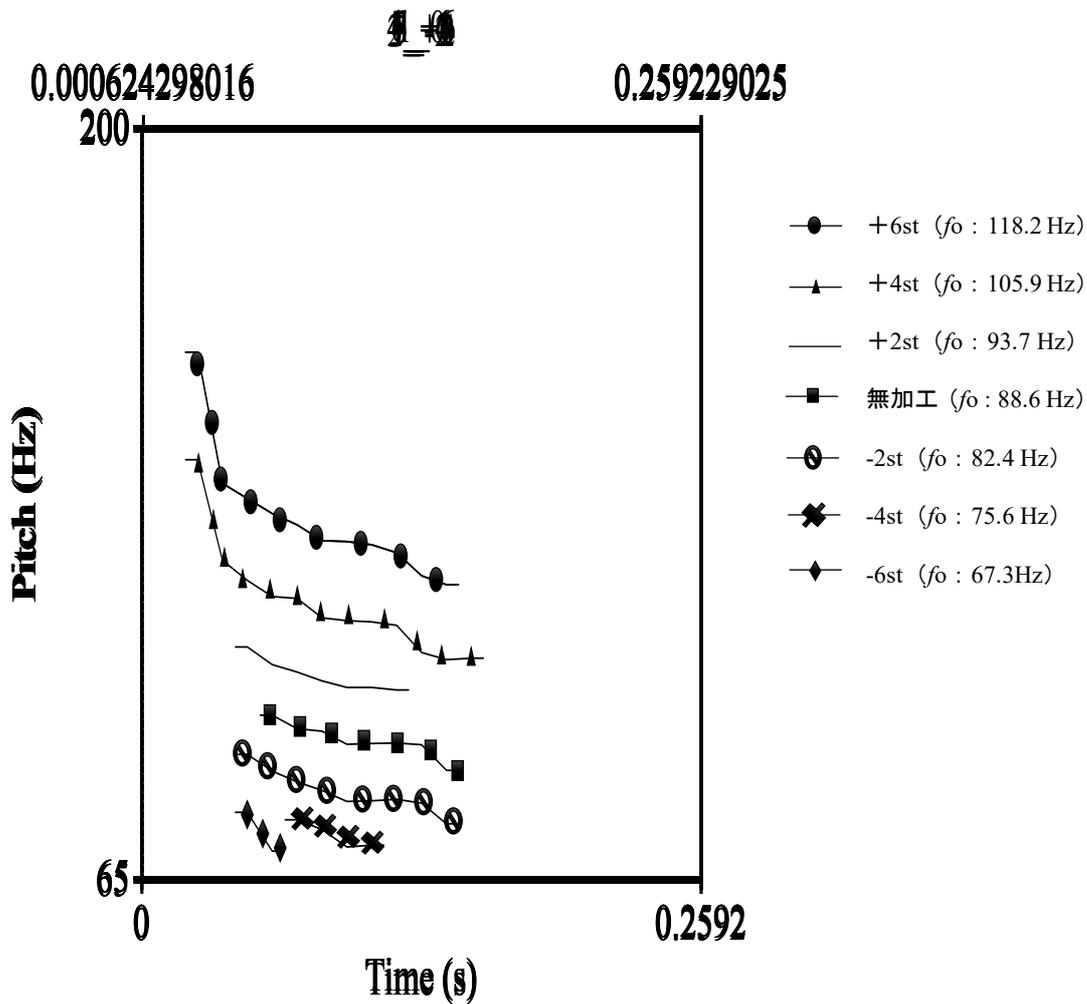


図 40 各加工タイプのピッチ曲線の具体例

[表 15 より No.7 の刺激音を参照]

ピッチと基本周波数の違いについて、ピッチとは、知覚される音の高さ、つまり心理的な感覚の程度であり、それに対応する物理量は基本周波数 (f_0) の対数值 (st) である (日本音響学会 online)。後続母音の f_0 が破裂音の有声・無声のカテゴリライズにかかわる音響特徴であると報告する先行研究が散見される。第 2 章の先行研究で既に述べたように、鄭・桐谷 (1998) においては、ピッチパターンを入れ替えた一部の音声に対し、日本語母語話者でも低ピッチの誤聴率が高ピッチよりも高くなる傾向が確認された。これは、日本語母語話者の有声・無声の弁別に後続母音のピッチが関与している可能性を示唆している。また、邊 (2019) においては、日本語の語頭有声・無声音の区別に関与する音響的特徴について調査した結果、後続母音の f_0 が有声・無声のカテゴリライズに関わるパラメータの可能

性を示唆している。

さらに、北村・川元 (2013) では、後続母音の f_0 のシフトが個人性知覚に及ぼす影響について調査した結果、 f_0 から 2st の上昇・下降が話者識別への影響が小さく、4st 以上の上昇・下降シフトが知覚への影響が大きいと報告している。知覚実験 IV の刺激音の作成は、先行研究の実験方法を参考にし、-6st, -4st, -2st, 無加工, +2st, +4st, +6st と計 7 種類の刺激音を作成した。

以上の作成方法で得た刺激音は、加工材料の「パ」(6 種) × ピッチの高・低 (2 種) × VOT 伸縮加工の結果 (7 種) の計 84 個となるが、さらに、実験に無関係な 16 個のダミー (例: 「ア」, 「ナ」など) を混ぜ込み、最終的に計 100 個の刺激音の実験材料とした。実験方法と実験制御ソフトは、第 3 章の知覚実験 I と同じく、Hot Soup Processor (HSP) Ver.3.4 を使用し、コンピュータからランダムに音声呈示した。また、この知覚実験には、近畿圏出身の日本語母語話 10 名に協力してもらった。実験協力者の内、男性 7 名、女性 3 名であった。実験を実施した時の平均年齢は 28.7 歳 (SD=1.89) であった。実験の所要時間は事前説明、練習問題、本調査を合わせると、約 8 分であった。

6.3 実験の結果

後続母音の f_0 を加工した 2 種類 (H・L) の「パ」を実験協力者に判定させた結果、無声の「パ」/有聲の「バ」の両方に知覚された。無声音に知覚された割合は図 41 に示す。縦軸は知覚率 (%), 横軸は刺激音の種類 (高ピッチ・低ピッチ)。総合的知覚率は、高ピッチの場合が 96.19% に対して、低ピッチの場合は 9.29% であった。

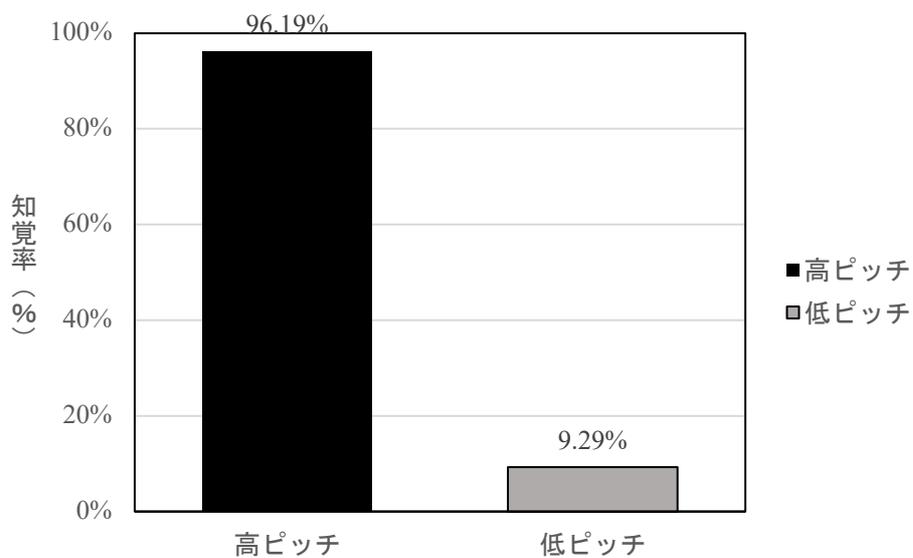


図 41 無声音の知覚率

種類毎の加工タイプに対する知覚率は図 42 と図 43 に示したとおりである。縦軸は知覚率 (%), 横軸はそれぞれのトークンの後続母音の f_0 加工倍率を示す。

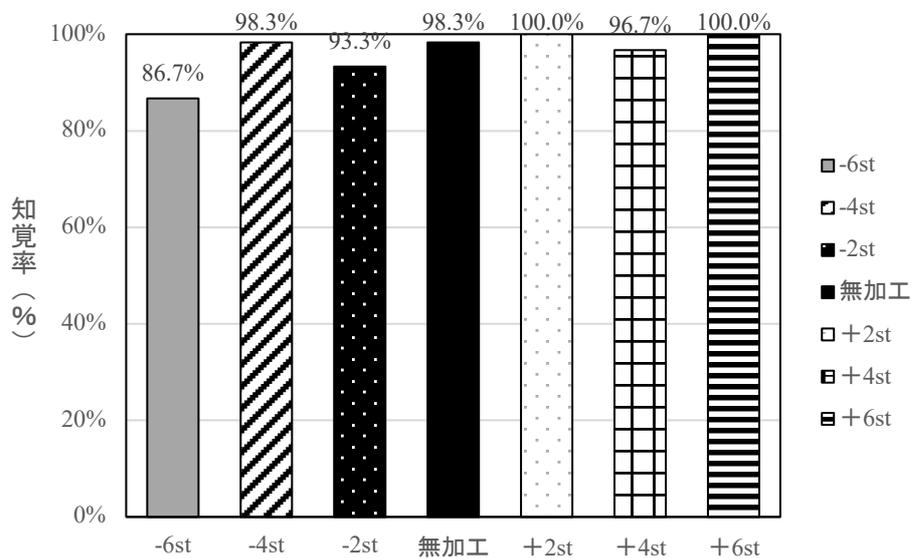


図 42 高ピッチ「パ」の無声音知覚率

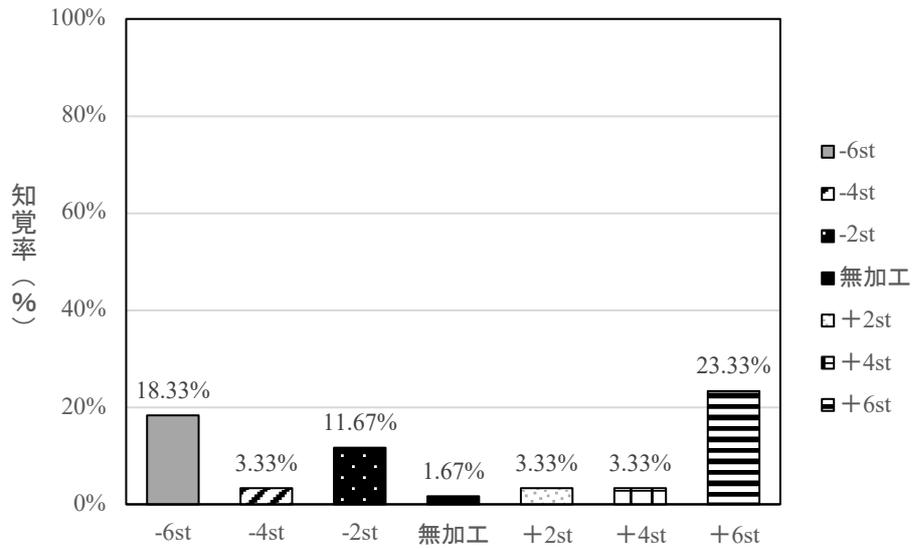


図 43 低ピッチ「パ」の無声音知覚率

次に、後続母音の f_0 の 7 種類の加工タイプによる 2 種類の組み合わせの間に「パ」の知覚に差があるかどうかを確認するため、対応ありの一元配置分散分析で検討した。

高ピッチ「パ」に対する分析結果では、 $F(6, 54) = 2.795$, $p = .069 > 0.05$ でその差は有意ではなかった。

一方、低ピッチ「パ」の f_0 加工された刺激音に対する知覚実験の結果についても、高ピッチ「パ」と同じように分散分析を行った。分析結果では、 $(F(6, 54) = 3.709, p = .044 < 0.05)$ で有意であった。しかし、ボンフェローニ補正を用いて多重比較を行なった結果、有意な差は確認されなかった (表 16)。

表 16 加工倍率が異なる刺激音の知覚結果の多重比較（低ピッチ）

	-6 st	-4 st	-2 st	無加工	+2 st	+4 st	+6 st
-6 st							
-4 st	0.314						
-2 st	1.000	1.000					
無加工	0.366	1.000	0.501				
+2 st	.878	1.000	1.000	1.000			
+4 st	.481	1.000	1.000	1.000	1.000		
+6 st	1.000	1.000	1.000	1.000	1.000	1.000	

有意水準<.05

以上のような結果から、刺激音「パ」の知覚と後続母音の f_0 の関係が明らかとなった。まず、刺激音を加工する前に無声音と知覚される高ピッチ「パ」の場合では、後続母音の f_0 を低くしても有声音「バ」に知覚することはなかった。 f_0 を高くすると、結果が変わらず、依然として無声音に知覚することが分かった。一方、刺激音を加工する前に有声音と知覚された低ピッチ「パ」の場合でも、知覚率に対しては有意な差が確認されたが、加工のタイプ毎に対しては有意差がなかった。

6.4 考察とまとめ

知覚実験 IV では、有声・無声知覚に関する後続母音の f_0 の要因を考慮し、切り出した音節中の無声子音を有声音に聴こえるかどうかについて加工音声を用いて検証した。その結果、後続母音の f_0 の高・低と関係があることが明らかとなった。そこで、実験結果に基づいて以下のように考察する。

まず、 f_0 の高・低を取り上げて検討した結果、語中から切り出した無声音が有声音に知覚する要因は、 f_0 が関与しているとわかった。日本語はピッチアクセントを有する言語で

あり、破裂音の有声性区別に関しては、語頭の音節のみ後続母音の f_0 を利用することができる。なぜなら、2.2.1 で説明した通り、語頭以外の音節がピッチアクセントの機能を担う f_0 、すなわちピッチの変動に影響があることに対して、語頭の音節が後続拍以降のピッチの変動に影響がなく、破裂音を含む音節中の母音に由来する f_0 を利用することが可能である (Gao&Arai 2018)。この特徴について先行研究では、語頭無声音の場合において後続母音の f_0 を高め、語頭有声音の場合において後続母音の f_0 を低める傾向があると指摘された (Shimizu 1996, 鄭・桐谷 1998)。しかし、邊 (2019) では、日本語母語話者を対象とし、日本語破裂音の有声性区別についての聴取実験が行われた結果、母語話者の出身地域や使用方言のより、後続母音の f_0 が有声性弁別に有効なパラメータである地域もあるが、そうでもないものもあると指摘している。それは、東北地方では、後続母音の f_0 が VOT と同様に重要なパラメータであるが、中部地方では、VOT に比べて f_0 の効果が相対的に小さいと報告されていた。以上の先行研究の結果をまとめると、日本語母語話者の有声・無声の知覚に、後続母音の f_0 が一つの弁別パラメータとして寄与していると言えるが、本実験の結果では、先行研究の見解と一致しなかった。後続母音の f_0 が弁別パラメータとして寄与している場合、日本語母語話者は両グループの刺激音とも有声/無声に知覚するはずだが、実際高ピッチグループの刺激音が無声音のみ知覚した。つまり、日本語母語話者の有声・無声の区別に、後続母音の f_0 が関与している可能性は示唆されていない。

次は、後続母音の f_0 の加工タイプについて考察する。低ピッチの場合、 f_0 の高・低の結果で有意差が検出されたため、さらに事後比較 (ボンフェローニ) を用いて全ての水準間 (加工タイプ) の組み合わせの差を検証した。その結果、有意差が検出されなかった。すなわち、高ピッチの刺激音は後続母音の f_0 の変動により、依然として無声音と知覚された。一方、低ピッチの刺激音は後続母音の f_0 の変動により、有声音ではなく無声音と知覚される傾向が見られた。このような結果について、日本語母語話者の有声・無声の区別に後続母音の f_0 が関与している可能性は示唆されていない。

最後に、図 41 の実験結果を表 14 と表 15 の加工結果と比べて考察する。高ピッチの刺激音の加工タイプ (表 14) と低ピッチの刺激音の加工タイプ (表 15) それぞれの後続母音の f_0 の平均値 (Hz・st) を表 17 にまとめて表す。

表 17 高・低ピッチの刺激音の加工タイプ毎の後続母音の f_0 平均値

*1 st=100 Hz

刺激音	-6 st	-4 st	-2 st	無加工	+2 st	+4 st	+6 st
高ピッチ	77.1 st	79.1 st	81.1 st	83.2 st	85.2 st	87.2 st	89.2 st
	86.1 Hz	97.8 Hz	108.6 Hz	121.6 Hz	136.1 Hz	152.9 Hz	170.8 Hz
低ピッチ	70.9 st	72.9 st	74.9 st	76.9 st	78.9 st	81.0 st	82.6 st
	69.8 Hz	75.0 Hz	79.5 Hz	86.5 Hz	93.8 Hz	106.4 Hz	115.9 Hz

表 17 の各加工タイプの f_0 平均値を見ると、高ピッチ刺激音の f_0 を 6 st 下げると ($f_0=76.9$ st・86.5 Hz)、低ピッチ刺激音の無加工タイプ ($f_0=77.1$ st・86.1 Hz) に近い高さとなる。しかしながら、知覚実験の結果では、高ピッチ刺激音の f_0 が -6 st の場合、86.7% の割合で無声音と知覚されたが、低ピッチ刺激音の f_0 が無加工の場合、無声音と知覚された比率が僅か 1.7% であった。また、低ピッチ刺激音の f_0 を 6 st 上げると、平均値が 82.6 st (115.9 Hz) となる。この結果を高ピッチの刺激音と比べれば、-2 st タイプ ($f_0=81.1$ st・108.6 Hz) と無加工タイプ ($f_0=83.2$ st・121.6 Hz) の間にある。無声音の知覚率を比べれば、低ピッチ刺激音の -2 st タイプが 23.3% であるのに対して、高ピッチ刺激音の -2 st タイプが 93.3%、無加工タイプが 98.3% である。以上の結果を踏まえると、高ピッチの刺激音の後続母音の f_0 を下げても、無声音の知覚率が低ピッチの場合と同じくらいに低くなることはなかった。一方、低ピッチの刺激音の f_0 を上げてても、無声音の知覚率はいくらか高くなるものの、高ピッチの場合ほど高くない。つまり、後続母音の f_0 以外に、なんらかの音響的特徴があるだろう、と推測される。

第 3 章の知覚実験 I から切り出した無声音を「高ピッチ」と「低ピッチ」2 グループに分けて一連の知覚実験で検証を行った結果、日本語母語話者は高ピッチの刺激音が無声音に、低ピッチの刺激音を有声音に知覚した。そのため、有声・無声の弁別的パラメータについては、一見「ピッチの高低」と関係があるにみえるが、本章ではより厳しい手法を用い、切り出した無声音の後続母音の f_0 を加工し、各タイプで比較を行った。その結果、ピッチ

が決定的パラメータとして寄与していない時もあることが判明した。それは、高ピッチのオリジナル音声の f_0 を上げてても下げてても、8割以上の確率で無声音として知覚されると確認できた。この結果は、知覚実験 I と同様である。一方、低ピッチのオリジナル音声の f_0 を変更すると、無声音の知覚率も上がった。すなわち、低ピッチの場合では、後続母音の f_0 が有声・無声判断の1つの判断材料となりうるが、高ピッチの場合を含めて考慮すると、必ずしも絶対的な条件にはならないことがわかった。語中にある音節を「切り出した」、すなわち「語中から語頭への移動」が無声音を有声音に知覚した要因ではないかと思われる。それで、次章では語中位置の影響について検討することにする。

第7章 知覚実験 V : 語中位置の影響

7.1 実験の背景と目的

第3章の知覚実験 I で残された3つの課題、つまり VOT の長さ、後続母音長、後続母音の f_0 が、切り出した音節中の2種類の無声音（高ピッチと低ピッチ）の有声・無声の知覚との関連性については、加工音声を用いた知覚実験を用いて、第4章、第5章と第6章で確認した。第4章と第5章の結果では、2種類の刺激音とも、VOT と後続母音の持続時間に対して短縮・伸長の操作をしても知覚実験 I と同じ結果であった。つまり、VOT が ± 6 倍、後続母音が ± 3 倍の条件では、VOT と後続母音の一定の限度内における長さの変化は日本語母語話者の知覚に影響を与えなかった。言い換えれば、有声・無声の弁別に VOT と後続母音長という要因は関わっていないのである。

一方、後続母音の f_0 では違う結果が見られた。知覚実験 I で無声音に知覚された高ピッチの無声音は、後続母音の f_0 の変動に関わらず無声音と知覚されたのに対し、有声音に知覚された低ピッチの無声音は、後続母音の f_0 の変動により有声音/無声音の両方に知覚されたことが判明した。この結果から、日本語母語話者の有声・無声の弁別において、低ピッチの場合は後続母音の f_0 が関与していることが確認できたが、高ピッチの場合は確認できなかった。すなわち、語中から切り出した音節頭の無声子音を有声音と知覚する判断材料として、後続母音の f_0 がその一つである可能性が示唆された。しかし、無声音に知覚された高ピッチの場合は後続母音の f_0 の変動からの影響を受けなかったため、 f_0 以外に他の音響的特徴が影響を与えている可能性があると考えられる。1つの可能性としては、無声音として発声した「無声子音+母音」を語中から「切り出す」、つまり語中位置を語中から語頭に移動したため、日本語母語話者がその単独の音節を有声音と知覚したと推知できる。

そこで、知覚実験 V では、語中の「無声音として発声した無声無気音」と音響的特徴が類似している「有声音として発声した無声無気音」、すなわち「無声音化した語頭有声音」を用いて、語中位置の変更（語頭から語中へ移動する）操作をすることによって、知覚にどのような影響を与えるかについて明らかにする。具体的には、知覚実験 I から IV までの実験と同様に、高・低ピッチによる2つのグループに分け、加工音声を作成し、それを刺激音として、日本語母語話者の知覚結果を検証する。この知覚実験 V を通して、主に以下の2点を観察していく。

- (1) 日本語母語話者は、日本語の語頭無声音化した有声音を語頭から語中へ移動すると、有声子音を無声子音として知覚するかどうか。
- (2) もし (1) が確認できた場合、その要因は知覚実験 I の結果のように、ピッチの高低の影響なのかどうか。

7.2 実験材料の作成方法

7.2.1 録音協力者

音声の録音には、日本語母語話者 1 名に協力してもらった。実験協力者は関東圏出身の 20 代男性で、高等教育機関で 6 年以上の教育歴を有するものである。また、録音の前に、録音協力者の発音特徴を確認するため、語頭有声音を有する 2 モーラの無意味語 3 個と有意意味語 3 個を 5 回ずつ発音してもらい、無声音化の割合が 100%であることを確認した。つまり、録音協力者は高田 (2011) で指摘された語頭有声音の無声音化が起きやすい若年層と関東圏出身、2 つの条件と一致することが確認できた。

7.2.2 録音語

実験対象である語頭有声音の子音は /b/ と /d/ とした。なお、実験語彙数と聴取実験の参加者の負担を考慮した結果、/b/ の後続母音は /a//i//u/、/d/ の後続母音は /a//e//o/、計 6 種とした。また、本研究の知覚実験 IV の調査結果により、ピッチの影響があると考え、録音音声のピッチを高 (H) と低 (L) の 2 種類とした。以上の組み合わせで、発話してもらった語頭無声音化した有声音を含む単音節が 6 種類と、ピッチが H と L で 2 種類、刺激音用の音声は計 12 種である。

7.2.3 音声の録音

音声の録音の詳細は以下のとおりである。

- (1) 実験用音声の録音：上記の 12 種の無意味語を図 44、図 45、図 46 のように配列して、録音協力者に発音してもらい、録音を行った。音声の録音は、大学の無響室で単一指向性マイクを使って、メモリレコーダーに収録した。
- (2) 実験用音声の録音機材：ポータブルリニア PCM レコーダー (SONY PCM-D50)。
- (3) 実験用音声の録音条件：録音のサンプリング周波数は 44100Hz、量子化ビット数は 16 bit。
- (4) 実験用音声の日本語母語話者による知覚確認：得られた発話音声につて、音声知覚実験に参加していない日本語母語話者 (近畿圏出身) 2 名に、録音内容が発話語

読み上げリスト（図 44，図 45，図 46）のとおりに正確に発音されたかを評定してもらい，紙面から聴いた第 1 音節の破裂音の部分を仮名表記で書くようお願いした。その結果，5 名の母語話者はすべて録音語の第 1 音節を有声音と判断した。

7.2.4 録音手順

録音手順と発話者への指示については，本番の録音前に，録音協力者に録音リストを渡し（図 44，図 45，図 46），録音の仕方を説明した。また録音協力者が十分指示を従えるよう 10 分間の練習時間を用意した。

図 44，図 45，図 46 に示すように，読み上げリストは，平仮名で書かれた無意味語の文字を 4 行ずつ並べたものである。なお，日本語のアクセントは高さアクセント (pitch accent) で，音声の相対的な高さが意味の区別をする（木村 1967）との事実があるため，録音語の種類により，それぞれ違う読み上げリストと指示を用意した。その理由は，本実験は自然会話に近いピッチを求めているため，録音協力者本人が慣れている音域で負担なく発音できるのを注意した。また，発話の面において，1 拍のみの単独発話の場合，録音協力者が専門的なボイストレーニングを受ける経験がなければ，指示通りのピッチを発声できない可能性もあるほか，慣れない音域でピッチがずれやすくなる可能性もあると考えられる。また，ピッチが低 (L) の場合（図 45，図 46），語末にくる「バ」のピッチが同じ L であっても，2 拍目と 3 拍目が異なるため，最もピッチが低い音を録音できるよう 2 種類の読み上げリストを用意した。

* () 内の「○回目」は発音しないでください。

な ば

(1回目) ば

(2回目) ば

(3回目) ば

(4回目) ば

な ば

(5回目) ば

(6回目) ば

(7回目) ば

(8回目) ば

図 44 発話語読み上げリスト (高ピッチの場合)

* () 内の「〇回目」は発音しないでください。

な ば

(1回目) ば

(2回目) ば

(3回目) ば

(4回目) ば

な ば

(5回目) ば

(6回目) ば

(7回目) ば

(8回目) ば

図 45 発話語読み上げリスト（低ピッチの場合）①

* () 内の「○回目」は発音しないでください。

な な ば

(1回目) ば

(2回目) ば

(3回目) ば

(4回目) ば

な な ば

(5回目) ば

(6回目) ば

(7回目) ば

(8回目) ば

図 46 発話語読み上げリスト（低ピッチの場合）②

以上の録音材料を用意し、読み上げリストについて、高ピッチの録音を図 44 のように、低ピッチの録音を図 45 と図 46 のように示した。

録音の手順については、録音を開始する前に発話者に、収録の手順、発話内容、発話スピードに関する説明を行った後に読み上げリストを渡した。また、発話内容については、実験用の対象音を発声する前に、高ピッチの場合が 2 拍の無意味語 (LH) を、低ピッチの場合が 3 拍の無意味語 (HLL) を発音してもらい、その後語末のピッチをキープしたまま 4 回ずつ発話をしてもらう。4 回目と 5 回目の間に 30 秒の休憩を挟んで、1 語につき計 8 回の発話をしてもらった。発話の留意点と指示については第 3 章知覚実験 I と同様で、必ず 1 語ずつ区切って、「ななば (HL₁L₂), ば (L₂), ば (L₂) ...」と話者にとって普通の発

話速度で読み上げるようにと指示した。なお、刺激音に選んだのは2回目か3回目の1発話で、計1発話である。

7.3 刺激音の作成

本調査では、語中位置（語頭・語中）の影響を確認するため、知覚実験 I に使用した語中無声音の音響的特徴とほぼ一致している「無声音化した語頭有声音」を語中に移動した加工音声を用いて、日本語母語話者を対象に知覚実験を行う。また、刺激音を作成する前にも、日本語母語話者に無声音化した語頭有声音を有声音と判断してもらった結果、高田（2004）と朱（2010）の指摘通り、日本語母語話者が語頭有声音において無声音化が起きても無声音ではなく有声音と知覚することを確認できた。

刺激語は、2 拍の無意味語を用意する。作成の内容については、録音協力者にとって最も発音しやすい音域で発声した単独の「ナ」を第 1 音節にして、無声音化した有声音を第 2 音節にする。しかし、閉鎖区間（内破）の持続時間については、杉藤・神田（1987）と朱（2010）によれば、閉鎖区間（内破）の持続時間が日本語の有声・無声音の知覚と重要な関係であり、内破が 100 msec を超えると無気音が無声音に聴こえ、短くすると有声音に聴こえると指摘している。そこで、本知覚実験に使用する刺激語の第 1 音節と第 2 音節の間の閉鎖区間（内破）の持続時間が 100 msec となるように、音響分析ソフト Praat 6.0.16 で音声波形と広帯域スペクトログラムを参考にして加工した。その詳細については、第 1 音節の「ナ」の第 3 フォルマントの最終抽出点を閉鎖区間（内破）の始点とし、そこから 70 msec の無音区間を第 1 音節の終点とする。第 2 音節は、破裂の瞬間の 30 msec 前を第 2 音節の始点とし、2 音節間に 100ms の無音区間を確保した。具体例は図 47 に示す。

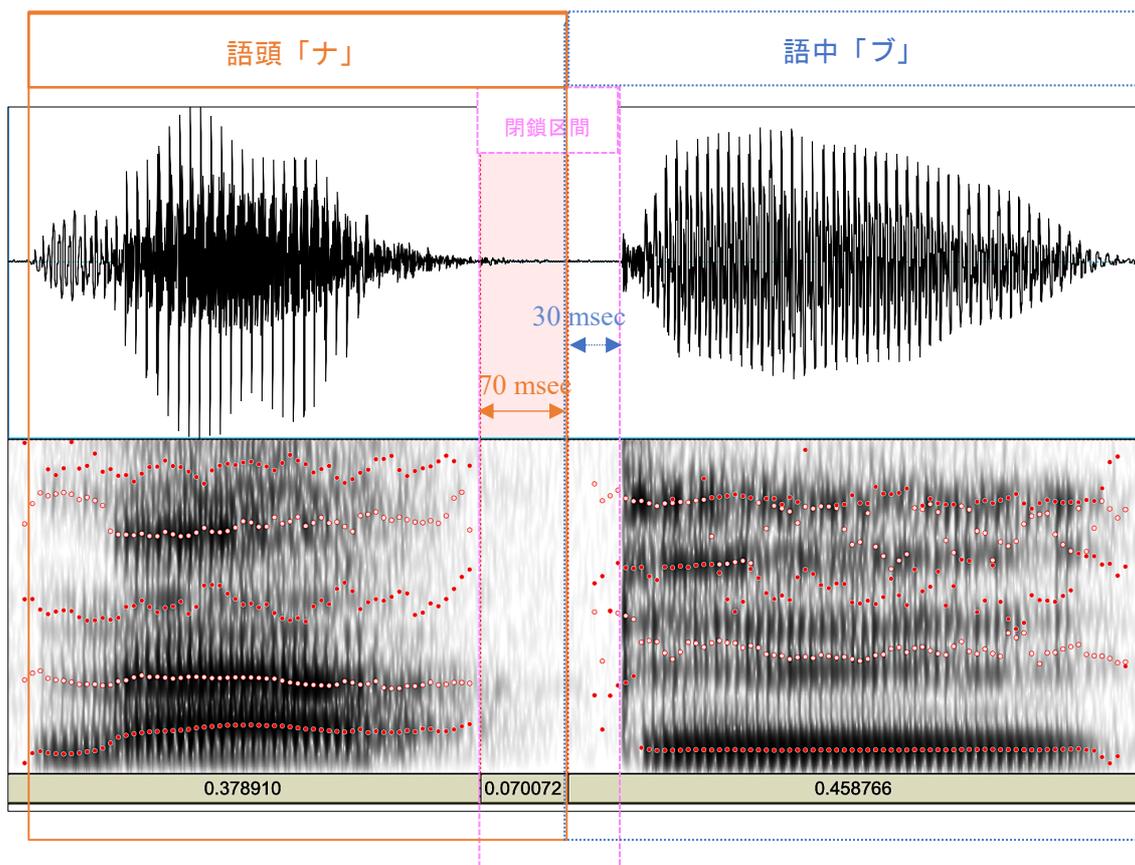


図 47 刺激語の音声波形・スペクトログラム

以上の作成方法に従い、第 2 音節の対象音が「バ、ビ、ブ、ダ、デ、ド」6 種類、ピッチによる高 (H) と低 (L) 2 種類のを合わせて、計 12 個の刺激音を「ナ」と組み合わせて加工音声を用意した。12 個の加工音声の VOT・後続母音長・ピッチなどの詳細を表 18 と表 19 にまとめた。ピッチは 1 st=100 Hz を基準に測定したものであるため、高ピッチの場合において、プラスの値を示す刺激音の周波数 (Hz) は 100 Hz 以上のものとなっている。低ピッチの場合は、周波数 (Hz) が 100 Hz 以下のものもあるため、半音値 (st) がマイナスとなる場合もある。

表 18 高ピッチ刺激音の音響特徴

	[ba]	[bi]	[bu]	[da]	[de]	[do]
VOT	6.0 ms	8.2 ms	10.3 ms	10.4 ms	9.5 ms	10.0 ms
後続母音長	367 ms	400 ms	398.9 ms	432 ms	443.7 ms	400.8 ms
ピッチ	5.24 st	4.49 st	3.34 st	3.98 st	3.74 st	4.02 st

表 19 低ピッチ刺激音の音響特徴

	[ba]	[bi]	[bu]	[da]	[de]	[do]
VOT	9.6 ms	10.0 ms	9.6 ms	7.8 ms	8.9 ms	10.8 ms
後続母音長	401.4 ms	316.1 ms	387.3 ms	328.3 ms	353.6 ms	381.6 ms
ピッチ	0.02 st	-1.35 st	-0.18 st	0.46 st	-0.34 st	-0.59 st

さらに、12個のダミー（例：「ザ」、「ガ」など）も刺激語と同じ方法で作成した上、12個の刺激語と混ぜ込み、最終的に24個の刺激語がタイプに関わらず、コンピュータからランダムに音声提示された。

反応時間の測定は、音声提示から始まるので、刺激語の音声継続時間が反応時間に含まれる。それぞれの刺激語の音声継続時間は表 20 にまとめた通りである。

表 20 刺激語の音声継続時間と平均（単位：msec）

	[ba]	[bi]	[bu]	[da]	[de]	[do]	平均 (SD)
高ピッチ	1052.7	1038.8	1071.6	1053.1	1041.4	1010	1044.60 (20.53)
低ピッチ	1036.7	1042.7	1037.5	1021.5	1014.2	1022.5	1029.18 (11.28)

各刺激語を作成する際、どの刺激音にも第1音節の前に3 msecの空白時間を挿入した。その結果、高ピッチの刺激語の音声継続時間の平均は1044.60 msec (SD=20.53 msec)、低ピッチの場合は1029.18 msec (SD=11.28 msec)である。また、独立したサンプルの*t*検定を行なった結果、お互いに有意な差が見られなかった ($t(10)=1.612, p=.138$)。以上のように音声持続時間は刺激語の聴覚提示による反応時間に影響しないことが推測される。

7.4 知覚実験の方法

音声知覚実験には、日本語母語話者（東海地方・近畿地方出身）15名に協力してもらった。実験を実施した時の平均年齢は20.57歳 (SD=2.01)である。

実験には、ノートパソコン（HP ProBook 450G8, 15.6型）、ヘッドフォン（Audio-technica ATH-ANC9）を用いて、刺激語ごとの反応時間を計測する聴き取りによる即時的正誤判断課題を実施した。

実験は、図書館や教室などの静かな場所で個別に行った。実験制御ソフトと流れは知覚実験Iと同様である。コンピュータの画面中央に凝視点「*」を600 msec視覚提示し、凝視点が消えると同時に、平仮名の視覚提示と刺激語の聴覚提示が同時に現れる。聴覚提示は1回のみ提示される。被験者には、ヘッドフォンを通して聴覚提示された音声画面に現れた仮名表記の語彙であるかどうかをできるだけ速く、かつ正確に判断するよう求めた。当該の刺激音が提示された視覚情報と一致している場合は「○」のボタンを、一致していない場合は「×」のボタンをクリックするよう指示した。ボタンをクリックすると、1問の

判定が終了し、次の問題提示のサイクルが始まる。各語が聴覚提示されてからボタンがクリックされるまでの時間を反応時間とする。判断の正誤を有声音の知覚率として測定した。被験者に本番の実験を始める前に操作に慣れてもらうため、10問の練習問題を行った。この実験のトータル所要時間は約7分であった。

7.5 実験の結果

本実験は、刺激語の第2音節にある有声音と発した音響的無声音の音声において、日本語母語話者が有声音に知覚された場合の割合を知覚率とする。実験の結果、刺激語に関する正誤判断課題の平均知覚率は、高ピッチが89%、低ピッチが92%であった。

反応時間に関しては、条件毎の反応時間の平均値から標準偏差3倍を超える反応についても外れ値¹² (outlier) として、分析から除外した。Miller (1991) は、平均値から標準偏差の3倍では条件間のサンプルサイズの違いが、2倍の時と比べると、大きくなることを報告している。しかし、本実験は有声音を有声音または無声音どちらの音として知覚するのかが確認することが目的である。そのため全ての反応時間が分析の対象となり、条件間のサンプルサイズを統一することができる。以上のことを踏まえて、外れ値除去の影響を小さくすることができる平均値から3倍を基準に用いた。この操作によって除外されたデータは2.2%であった。その結果、平均反応時間は、高ピッチが1502.80 msec、低ピッチが1644.53 msecであった。

7.5.1 知覚率の結果

図48は実験協力者の知覚率を表したものである。第2音節のピッチが高(H)と低(L)の間に、有声音に知覚される差を確かめるために、対応ありの t 検定を行った結果、その間に有意な差が見られなかった ($t(14)=-1.00, p=.334$)。

¹² ここでの外れ値とは、実験中のコンピューターの操作ミス、被験者自身の集中力低下、周りの突発的な騒音などが原因で生じる、極端に短い、あるいは極端に長い反応時間のことである(橋本2010)。

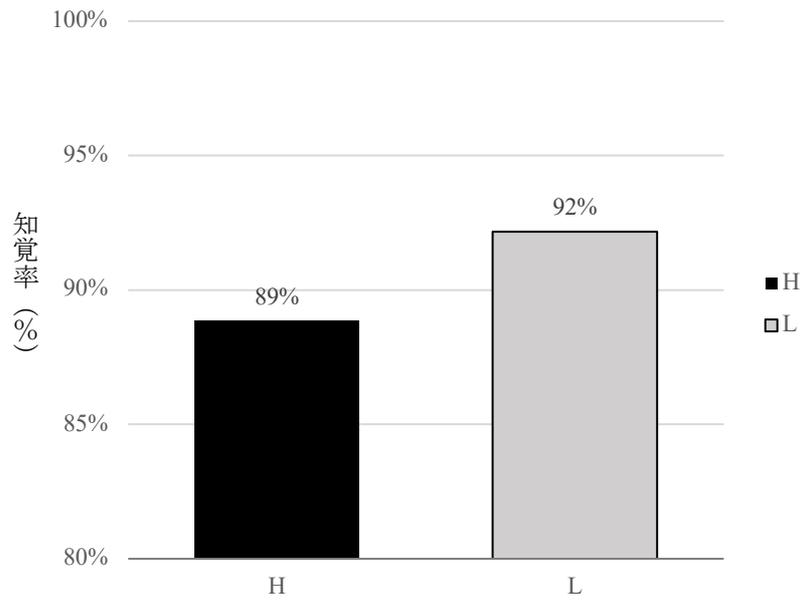


図 48 日本語母語話者の知覚率 (%)

7.5.2 反応時間の結果

平均反応時間について、第 2 音節のピッチが高 (H) と低 (L) の刺激語に対する結果を図 49 にまとめた。両タイプの刺激語の反応時間の差を確かめるために、対応ありの t 検定を行った結果、有意な差が見られた ($t(14)=-2.676, p=.018$)。低ピッチの方の反応時間が有意に長いことがわかった。また、Cohen の効果量を算出した結果、 $d=9.378$ となり、効果が大きいことがわかった。

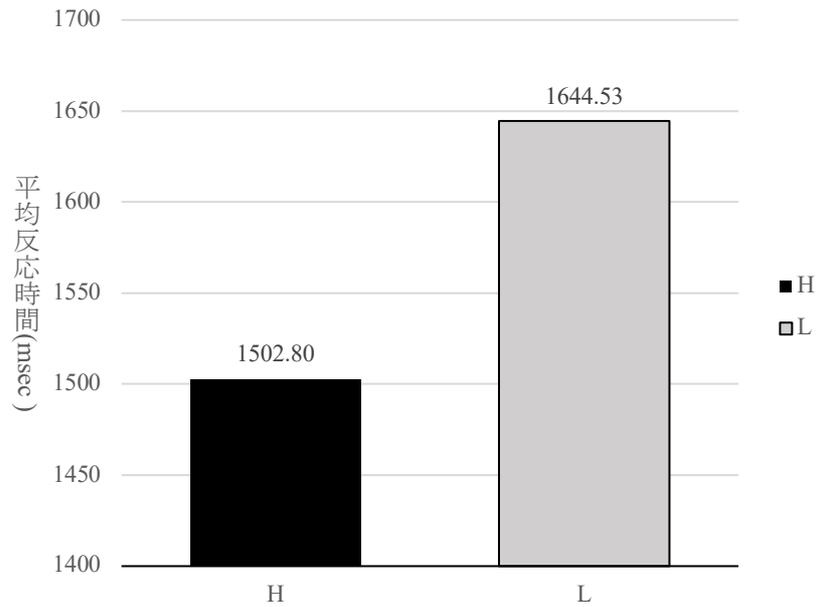


図 49 日本語母語話者の平均反応時間 (msec)

以上のように無声音化した語頭有声音を語頭から語中に移動して、日本語母語話者に聴かせる結果、有声音と知覚する知覚率においては、高ピッチと低ピッチグループの間に差異は確認されなかった。しかし、反応時間においては、高ピッチと低ピッチグループの間に差異が確認され、低ピッチの反応時間の方が長かった。

7.6 考察

前節の結果は表 21 のようにまとめられる。日本語母語話者が発する音響的無声音の語頭無声音化した有声音を、母語話者が有声音と知覚する知覚率と反応時間について、ピッチの高低別に語中位置の影響の有無を示す。

表 21 知覚率と反応時間における語中位置の影響

	正聴率	反応時間
加工前	n=2 高/低ピッチ：有声音	未測定
加工後	n=15 高/低ピッチ：有声音	n=15 高ピッチ：反応が短い 低ピッチ：反応が長い

研究課題に対して検証した結果をまとめると以下のようなになる。

- (1) 日本語母語話者は日本語の語頭無声音化した有声音を語頭から語中へ移動しても有声音として知覚する。
- (2) 語頭の無声音化した有声音を語中に移動しても、ピッチの高低が知覚結果に影響しなかった。しかし、反応時間の結果において、低ピッチの場合、反応時間が高ピッチより長かった。

次にこれらの結果を、語中位置（語頭・語中）の影響、閉鎖区間（内破）の持続時間、ピッチの影響の3点を中心に、知覚実験 I の結果と比べながら考察する。

7.6.1 語中位置（語頭・語中）の影響

本調査では、日本語の語頭有声音の語中位置（語頭・語中）に関して、従来の先行研究では十分取り上げられてこなかった無声音化の現象に着目し、語中位置が日本語母語話者の知覚に及ぼす影響を検証した。その結果、有声音の知覚率において語中位置による影響は見られなかった（表 21）。高田（2004）と朱（2010）では、語頭の無声音化した有声音について知覚実験を行なった結果、日本語母語話者は無声音化によって破裂前に声帯振動を伴わなくても、有声音として知覚できると報告していた一方、本調査では無声音化した有声音を語頭から語中に移動しても同じ結果であった。すなわち、日本語の無声音化した有声音は生起環境が語頭でも語中でも、日本語母語話者の無声・有聲の弁別に影響がないことを示唆している。

次は、2種類の無声無気音の音響的特徴と語中位置の関係による知覚結果について考察する。本実験で用いた「有声音として発声した無声無気音」は、破裂前に声帯振動を伴わない、すなわち無声音化したもので、朱（2010：65）によれば、破裂の弱い場合が有声音として知覚されるが、破裂の強い場合が無声音として知覚される場合もあると報告されている。また、知覚実験Ⅰで用いた「無声音として発声した無声無気音」は、朱（2010：65）では「中国語の無気音的なもの」と報告され、その特徴は、呼気流率が無声音より半減とし、VOT（外破）が短く aspiration も不明瞭であるという。

知覚実験Ⅰと知覚実験Ⅴの結果から、「無声音として発声した無声無気音」と「有声音として発声した無声無気音」、両タイプの無声無気音とも有声音と知覚されることがあると確認できた。しかし、知覚実験Ⅰの場合は、低ピッチの刺激音のみ有声音と知覚されることがあるため、「低ピッチ」との条件付きで、語中位置の影響が認められた。一方、知覚実験Ⅴにおける反応時間の結果から考えると、母語話者はピッチの高低と関係なく有声音として知覚したとしても、低ピッチの場合の方は判定するのに時間がかかった。すなわち、知覚実験Ⅰと知覚実験Ⅴとも、低ピッチの場合において、語中位置（語頭・語中）はある程度の影響があることを示唆している。

7.6.2 閉鎖区間（内破）の持続時間

有声音として知覚される音の音響的特徴とは、声帯振動の有無のみならず、閉鎖区間（内破）の持続時間、VOT（外破）の持続時間、破裂の強さが関与していると朱（2010）と杉藤・神田（1987:15）が報告している。

閉鎖区間（内破）の持続時間について、杉藤・神田（1987）では、内破の持続時間の短い無気音は有声音に聴こえるが、長い場合は無声音に聴こえる。特に日本語の内破の持続時間が先行母音の4割近くまで短縮されると有声音として知覚されると報告されている（杉藤・神田 1987:5）。朱（2010:53）でも、閉鎖区間（内破）が100 msecを超える日本語母語話者が発した無気音は、無声音に知覚される比率が70%を上回ると報告している。

ところで、知覚実験Ⅴで用いた刺激語の内破は、先行母音の終了時点から70 msecの区間と後続の「無声音として発音された有声子音」の破裂前30 msecの区間、あわせて100 msecの無音区間で、先行研究で指摘された有声・無声弁別の境界となる長さである。実験結果では、90.5%の比率で無声音化した有声音が有声音に知覚された。しかし、知覚実験Ⅰで使用した刺激音は語中から切り出した無声音の音節であり、語中の場合のように有声・無声の弁別の手掛かりとなる閉鎖区間がないため、さらに無声音を有声音に知覚されやす

いはずだが、実際に有声音に知覚された比率は 47%どまりであった。

以上のことから、知覚実験 V の閉鎖区間がちょうど有声・無声弁別の境界にあたる長さであったから有声音と知覚したのだらうとの考えも、知覚実験 I の結果から考えると、必ずしも成立しない。おそらく、閉鎖区間（内破）の持続時間以外にも、何らかの要因があったのかもしれない。今後さらなる実験で検証する必要があるだろう。

7.6.3 ピッチの影響

これまでの実験において、それぞれの知覚実験に用いた刺激音声が高ピッチと低ピッチ計 2 種類である。ピッチ別で有声音として知覚する割合をみると、知覚実験 V の 90.5%のうち、高ピッチの場合が 88.8%で、低ピッチの場合が 92.2%。知覚実験 I の結果において、有声音の知覚率 47%のうち、高ピッチの場合が 1.8%で、低ピッチの場合が 85.1%である。この結果によれば、有声音として発声した音響的無声音はピッチの影響を受けないのに対して、無声音として発声した音響的無気音はピッチの影響を受けると見ることができる。両タイプとも無声無気音であるが、ピッチの影響による結果から、「有声音として発声した無声無気音」と「無声音として発声した無声無気音」の間に、現段階ではまだ分からない何らかの違いがあることが示唆されている。

さらに、日本語の語頭有声音の音響的特徴に関して、邊 (2019) では VOT と後続母音の *fo* について調査を行った結果、破裂前に声帯振動が必ずしも伴わないこともあるため、VOT の +/-のみでは有声・無声の区別が難しく、後続母音の *fo* も区別にかかわると報告している。しかし、後続母音の *fo* については地域差があり、東北地方では後続母音の *fo* が VOT と同等に有声・無声の弁別に主要な手掛かりとなるのに対し、中部地方では後続母音の *fo* 効果が VOT に比べて相対的に非常に小さいとも指摘されている。本研究の知覚実験 V で協力してもらった日本語母語話者の内訳は東海地方出身 8 名と近畿地方出身 7 名であり、半分くらいの協力者が邊 (2019) で指摘された後続母音の *fo* 効果が小さい方言話者である。今後、語頭無声音化した有声音が語中に移動しても有声音と知覚するかどうか、実験協力者の出身地や邊 (2019) で指摘された地域差の現象を考慮した上、さらに検証する必要があると考える。

7.7 まとめ

本実験では、語頭に発話された日本語の無声音化した有声音を語中に移動すると、日本語母語話者の知覚にどのような影響を及ぼすかについて、加工音声を用いた知覚実験を通

して検証した。その結果、日本語話者はピッチの高・低にかかわらず、無声音化した語頭の有声音が第1音節から第2音節に移動された場合、無声音と知覚する傾向は見られなかった。一方、知覚実験Ⅰでは、語中の無声音を第2音節から第1音節に移動した場合は有声音と知覚する傾向が確認され、低ピッチが高ピッチより有声音の知覚率が高かった。この結果を表22に示す。

表22 知覚実験Ⅰと知覚実験Ⅴの結果のまとめ

		加工前位置・知覚結果	加工後位置・知覚結果
知覚実験Ⅰ： 無声音として発声 した無声無気音	高ピッチ	語中・無声音	語頭・無声音
	低ピッチ	語中・無声音	語頭・有声音
知覚実験Ⅴ： 有声音として発声 した無声無気音	高ピッチ	語頭・有声音	語中・有声音
	低ピッチ	語頭・有声音	語中・有声音（無声音と躊躇）

表22に示したとおり、有声音として発声した無声無気音（すなわち「無声音化した語頭有声音」）は、ピッチの種類（高・低）が異なっても、知覚結果は変わらなかったが、語中位置（語頭・語中）が変わると、少なからず影響があることが確認された。一方、無声音として発声した無声無気音（すなわち「語中無声音」）は、ピッチの影響が確認されたほか、語中位置（語頭・語中）の影響もあることが確認された。

日本語の有声音と無声音は音韻論的には2つの範疇となっているため、従来多くの先行研究でも2つの範疇に分けて知覚実験が行われた。本研究では、音声的パラメータであるVOTや後続母音の*f₀*が類似している無声無気音を、語頭から語中へ、または語中から語頭へ移動したうえ知覚実験を行い、日本語母語話者が語中の場合は無声音に、語頭の場合には有声音に知覚されるかどうかについて検証した。ただし、語頭の無声音化した有声音を

語中に移動しても無声音ではなく有声音に聴こえるかの検証は、知覚実験 I のように一連の追実験（知覚実験 II, III, IV）を行なっていないため、語中位置（語頭・語中）が有声・無声の知覚に影響を与えなかったかどうかの確認は今後の課題としてさらに検討したい。

第8章 結論

8.1 本研究のまとめ

本研究では、「無声音として発声した無声無気音」と「有声音として発声した無声無気音」の音声の特徴に焦点を当て、日本語母語話者の有声・無声破裂音の知覚にどのような音響的パラメータが影響を与えているのかを明らかにするため、5つの知覚実験を行なった。

知覚実験Iでは、「日本語母語話者（発話者本人を含む）は、文脈から語末の「無声子音+母音」の音節を切り出して単独で聴くと、無声子音が有声音と知覚する現象について検証するため、日本語母語話者がピッチの異なる（高ピッチと低ピッチ）無声音として発声した無声無気音を語中から切り出して単独で聴かせる知覚実験を行なった。その結果、日本語母語話者は切り出した低ピッチの音節中の無声子音を有声音と知覚することが確認できた。しかし、知覚実験Iで用いた2種類（高ピッチ・低ピッチ）の刺激音の間で音響的特徴（VOT・後続母音長・後続母音の*f₀*）が異なっていたため、どの音響的パラメータがこの現象に影響を与えているのかは分からない。

そこで、知覚実験IIではVOTの長さを、知覚実験IIIでは後続母音長を、知覚実験IVでは後続母音の*f₀*をそれぞれ操作した加工音声を用いて追実験を行なった。なお、知覚実験IIからIVでは、知覚実験Iで用いた日本語母語話者が無声音として発音した無声無気音、すなわち語中から切り出した2種類（高ピッチ・低ピッチ）の無声音をベースとして加工した。

知覚実験IIでは、高ピッチと低ピッチの刺激音に対し子音部のVOTを5段階（0倍、±3倍、±6倍）に加工し、VOTの長短が日本語母語話者の知覚に影響を及ぼすのかを検証した。その結果、高ピッチの刺激音はVOTの長さに関わらず無声音と知覚され、低ピッチの刺激音もVOTの長さに関わらず依然として有声音と知覚された。この結果は知覚実験Iと同様であり、すなわち、VOTが+/-6倍の伸長・短縮では、日本語母語話者の破裂音の知覚に変化が見られなかった。そのため、切り出した無声音を有声音に聴こえる要因として、VOTが関与していないことが示唆された。

次の追実験である知覚実験IIIでは、知覚実験IIと同様の高ピッチと低ピッチの刺激音に対し子音部の後続母音の持続時間を5段階（0倍、±2倍、±3倍）に加工し、後続母音の持続時間が頭子音に対する日本語母語話者の知覚に及ぼす影響を検証した。その結果、知

覚実験IとIIと同様に、高ピッチの刺激音は母音の持続時間に関わらず無声音と知覚され、低ピッチの刺激音も母音の持続時間に関わらず有声音と知覚された。すなわち、後続母音が ± 3 倍の条件では、後続母音長の短縮・伸長が日本語母語話者の知覚に影響を与えなかった。この知覚実験の結果によれば、有声・無声の区別に後続母音の長さという要因が関わっていないことが明らかとなった。

さらに、知覚実験 II と III を同様の高ピッチと低ピッチの刺激音を利用した知覚実験IVでは、刺激音に対し子音部の後続母音の *fo* を 2st 間隔で 7 段階 (0st, ± 2 st, ± 4 st, ± 6 st) に加工し、後続母音の *fo* が日本語母語話者の知覚に影響を及ぼすのかを検証した。その結果、知覚実験 I で無声音に知覚された高ピッチの刺激音は、後続母音の *fo* の変動に関わらず無声音と知覚されたのに対し、有声音に知覚された低ピッチの刺激音は後続母音の *fo* の変動により、有声音ではなく無声音と知覚される傾向が見られた。この結果から、低ピッチの場合は後続母音の *fo* が日本語母語話者の有声・無声の弁別に関与していることが確認できたが、高ピッチの場合は確認できなかった。すなわち、低ピッチの場合のみ後続母音の *fo* は無声音として発声した無声無気音の有声・無声の 1 つの弁別的パラメータである可能性があることが明らかとなったため、その他に何らかの音響的特徴がある可能性が示唆された。

知覚実験IからIVでの「日本語母語話者は、語中の無声子音+母音の音節を切り出して単独で聴くと有声音と知覚する」現象についての検証を通じて、日本語母語話者の有声性の弁別に無声子音の VOT やその後続母音の持続時間は積極的に関わっていないことが確認され、後続母音の *fo* に関しては「条件付き」で有声性の弁別に寄与することが明らかとなった。しかし、知覚実験IVでは、高ピッチの刺激音において後続母音の *fo* の影響が確認できなかったため、後続母音の *fo* 以外にも他の要因があると考えられる。1 つの可能性としては、無声音として発声した「無声子音+母音」を語中から「切り出す」、つまり語中位置を語中から語頭に移動したため、日本語母語話者がその単独の音節を有声音と知覚したのであろう。

そこで、次の知覚実験Vでは、知覚実験I~IVで用いた語中の無声音と同じ音響的特徴を持つ有声音として発音された語頭の無声無気音、すなわち無声音化した有声音を、語頭から語中に移動し、日本語母語話者の知覚にどのような影響を与えるのかについて検証した。その結果、高ピッチと低ピッチの刺激語とも語中位置の移動による影響は見られず、日本語母語話者は無声音化した語頭有声音を有声音と知覚した。すなわち、「有声音として発

声した無声音」は、知覚実験IからIVで確認された「無声音として発声した無声音」と音響的に類似していても、母語話者に異なる音声として確認されることがわかった。この結果を踏まえると、日本語母語話者の有声・無声の弁別には、まだ言及されていないパラメータがあるとの可能性は否定できない。

知覚実験Vでは、反応時間も併せて測定することで、ピッチの違いにより処理時間に差があるかどうかを検証した。その結果、低ピッチの刺激語の反応時間は高ピッチの刺激音より有意に長いことが確認された。このことは、低ピッチの刺激語の知覚処理が難しいことを意味すると思われる。

8.2 教育現場への提言

上述したように、語中における低ピッチの「無声子音+母音」が、語中から切り出すと無声子音が有声音に聴こえる現象が起こる。本研究の結果により、無声音として発声した無声無気音が一定の条件でピッチが低くなると、日本語母語話者でも有声音と知覚する事実は、これまであまり表立って議論されてこなかった日本語の破裂音の音響的特徴を示すものである。一方、有声音として発声した無声無気音の場合、ピッチの高低と関わらず、母語話者は有声音として知覚することがわかった。

従来、日本語破裂音の知覚の問題では、中国語系話者は日本語の語中無声音を有声音に聞き間違えると多くの先行研究に報告されている。さらに、日本語の語頭有声音が無声音化する現象も珍しくないこと。これらのことを踏まえると、日本語の語中無声音と無声音化した語頭有声音と中国語の無気無声音は VOT 値がプラスの無声領域にあり、音響的に類似している。中国語を母語とする日本語学習者への教育現場では、朱 (2010:63) が「日本語の無声子音は中国語の有気音より呼気が弱い。それが語中に来るときは、呼気がさらに弱くなる」という事実を踏まえた指導法を検討すると報告している。そこで、筆者は朱 (2010:63) が指摘した指導法を踏まえ、本論で5つの知覚実験を通した結果から、さらに「語中に来るピッチが低い日本語の無声子音は、後続母音の *f₀* の高/低により、無声音/有声音に聴こえる」という事実を授業の中で与える必要があると考える。特に初級日本語学習者に破裂音を指導する際、アクセント型によるピッチの高低がさらに有声・無声の区別に影響を与えることも可能になると考えられる。

8.3 今後の課題

本研究では、日本語母語話者が日本語の語中から切り出した無声音を含む音節を有声音に聴こえる現象に対して、「無声音として発声した無声無気音」と「有声音として発声した無声無気音」を用いて5つの知覚実験を行った。これらの結果を総合的にみると、後続母音の f_0 の条件付きで、日本語母語話者でも無声音を有声音に知覚することがある。すなわち、日本語母語話者は語中から切り出した低ピッチの音節頭の無声子音を有声音と知覚することが明らかとなった。しかし、この結果にはなお疑問が残る。それは、日本語母語話者はオリジナルの語中低ピッチの「パ」の音節頭子音を無声音[p]と知覚するのに、語中から切り出して単独で聴くと無声音[p]ではなく、有声音[b]と知覚することである。この現象は、確かに後続母音の f_0 が1つの要因であることが実験で証明できたが、切り出す前と切り出した後の位置変化のほかに、なんらかの要因、もしくは音声的特徴の変化が生じたことで、母語話者が無声から有声に知覚するなど、課題もいくつか残している。そこで、今後の課題としては、以下の3点が挙げられる。

まず、第7章知覚実験Vの語中位置（語頭・語中）に関する疑問と課題である。破裂音の知覚について今まで数多くの研究で有声音・無声音の弁別に関わる音響的特徴が挙げられたが、語中位置（語頭・語中）を問わず有声・無声破裂音の共通する音声特徴であるかについて本研究の第7章で検証した。その結果、語頭の無声音化した有声音が語中に移動しても知覚には影響がないことが明らかとなった。しかし、この実験に対して2つの疑問が残されている。1つ目は、邊（2019）では日本語の語頭有声音の音響的特徴である VOT と後続母音の f_0 について調査した結果、有声・無声の弁別に方言差があると指摘している。特に東北方言話者の場合は後続母音の f_0 と VOT が有声・無声弁別の重要な手がかりになるが、それに対して、中部地方の方言話者の場合は後続母音の f_0 の効果は極めて小さいと指摘された。第7章の結果と邊（2019）の結果を比較して考察すると、知覚実験Vに参加した日本語母語話者は東海と近畿地方出身の方言話者であり、邊（2019）で示された後続母音の f_0 の効果が小さい方言話者である。つまり、本調査は有声・無声の弁別特徴に後続母音の f_0 の効果が極めて小さい方言話者が協力者であったため、無声音化した有声音が語頭から語中に移動しても、語中位置（語頭・語中）とピッチ（高・低）の間に差異が見られなかったのかもしれない。本当にそうであるかどうかについて、今後の課題としてさらに追実験で検証する必要がある。

2つ目は、知覚実験II, III, IVの一連の実験では、「無声音として発声した無声無気音」

を用い、その無声無気音の VOT・後続母音長・後続母音の *f₀* を加工したうえ、母語話者の知覚に影響を及ぼす要因を探った。しかし、「有声音として発声した無声無気音」を用いて、知覚実験 II, III, IV の一連の研究を行なっていないので、語中位置（語頭・語中）が影響を与えなかったかどうかについて検証する必要があると考える。

最後に、今後は本論文の結果を踏まえ、破裂音の有声・無声の弁別にかかわる音響的パラメータについて、先行研究で指摘された音響的特徴のほかには何かあるのか、引き続き検討していく。具体的には、EGG（電気声門図記録）を利用し、破裂音生成時の声門動作における緊張性発声コントラストについて 2 種類の実験を計画している。山中（2000）では、EGG の有用性について検討した結果、声門閉鎖時の EGG 波形が声帯振動パタンの推測に実用的・有用で指摘されている。日本語母語話者が「無声音として発声した無声無気音」と「有声音として発声した無声無気音」を発声する際、声門閉鎖時の EGG 波形が同様であるか確認することで、この 2 つの破裂音が音声的に類似しているかどうか確認できるのではないかと考えている。

本論文で取り扱った「日本語母語話者が日本語の語中から切り出した無声音を含む音節を有声音に聴こえる」現象が日本語のみか、通言語的な現象であるかどうかについてさらに研究を進めたいと思う。今後の課題として、英語の「無声音として発声した無声無気音」と「有声音として発声した無声無気音」、中国語の「無気無声音」と「調音的有声音化の無気音」を取り込んで、それぞれの母語話者の知覚について調査を行いたい。

参考文献

<英語文献>

- Gao, J., & Arai, T. (2018). F0 perturbation in a “pitch-accent” language. 6th International Symposium on Tonal Aspects of Languages, Berlin, Germany, 18–20.
- Homma, Y. (1973). An Acoustic Study of Japanese Vowels – Their Quality, Pitch, Amplitude, and Duration. *Study of Sounds*, 16, 347–368.
- Homma, Y. (1980). Voice-Onset Time in Japanese Stops. *Bulletin of the Phonetic Society of Japan*, 163, 7–9.
- Homma, Y. (1981). Durational relationship between Japanese stops and vowels. *Journal of Phonetics*, 9, 273–281.
- Lisker, L. & Abramson, A.S. (1964). A cross-language study of voicing in initial stops: acoustical Measurements. *Word*, 20, 384–422.
- Miller, J. (1991). Reaction time analysis with outlier exclusion: Bias varies with sample size. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*, 43A(4), 907–912.
- Shimizu, K. (1996). A Cross-language Study of Voicing Contrasts of Stop Consonants in Asian Languages. Seibido Publishing Co.

<中国語文献>

- 何萬順 (2009) 「語言與族群認同:從臺灣外省族群的母語與台灣華語談起」『語言暨語言學』 10:2, 375–419.
- 李淑鳳 (2010) 「台、華語接觸所引起的台語語音的變化趨勢」『台語研究』 2:1, 56–71.

<日本語文献>

- 王育徳 (2013) 『王育徳の台湾語講座』東方書店.
- 大友信一・佐藤利男・高倉稔 (1957) 「日本語の無声破裂音と有声破裂音の相違」『音声学会会報』 95, 4–7.
- 木村和夫 (1967) 「日本語・英語におけるストレスとピッチ」『天理大学学報』 19:2, 143–156.

- 瞿琦 (2017) 「日本語の語頭有声破裂音の知覚に+VOT化が及ぼす影響-日本語話者・中国語話者・台湾語話者との比較を通して-」 東呉大学日本語文學系研究所修士論文.
- 瞿琦 (2021) 「日本語の語頭有声破裂音の無声音化が知覚に及ぼす影響-日本語・台湾華語・台湾閩南語話者の比較を通して-」 『実験音声学・言語学研究』 13, 1-18.
- 斎藤純男 (2009) 『日本語音声学入門 改訂版』 三省堂.
- 蔡茂豊 (1976) 『中国人に対する日本語教育の理論と実践-音声教育篇』 東呉大学日本文化研究所.
- 清水克正 (1993) 「閉鎖子音の音声的特徴-有声性・無声性の言語間比較について-」 『アジア・アフリカ言語文化研究』 45, 163-175.
- 清水克正 (1999) 「日英語における閉鎖子音の有声性・無声性の音声的特徴」 『音声研究』 3:2, 4-10.
- 清水克正 (2018) 「閉鎖子音の有声性・無声性の音声的特徴に関する考察」 『音声研究』 22:2, 69-80.
- 朱春躍 (1994) 「中国語の有気・無気子音と日本語の無声・有声子音の生理的・音響的・知覚的特徴と教育」 『音声学会会報』 205, 34-62.
- 朱春躍・杉藤美代子 (1998) 「中国語無気音の有声化-その母音環境及び声調との関連」 『1998年度日本音声学全国大会予稿集』, 115-220.
- 朱春躍 (2010) 『中国語・日本語音声の実験的研究』 くろしお出版.
- 鄭恩禎・桐谷滋 (1998) 「ピッチパターンが日本語の有声・無声の弁別に与える影響-韓国語母語話者と日本語母語話者の比較-」 『音声研究』 2:2, 64-70.
- 杉藤美代子 (1996) 「中国語話者による日本語の無声子音・有声子音と有気音・無気音」 杉藤美代子著 『日本語の音』 (日本語音声の研究 3) (pp.264-285) 和泉書院.
- 杉藤美代子・神田靖子 (1987) 「日本語と中国語話者の発話による日本語の無声および有声破裂子音の音響的特徴」 『大阪樟蔭女子大学論集』 24, 1-17.
- 鈴木義昭 (1985) 「中国語教育と日本語教育-中国語教育の視点から-」 『日本語教育』 55, 59-70.
- 田村俊介・森田大樹・伊藤一仁・廣瀬信之・森周司 (2016) 「有声・無声子音のカテゴリー境界と無音検出閾値の関連性」 『日本音響学会春季研究会講演論文集』, 485-488.
- 高田三枝子 (2004) 「日本語の語頭の有声歯茎破裂音/d/における+VOT化と世代差」 『音声研究』 8:3, 57-66.

- 高田三枝子 (2006) 「語頭有声破裂音における VOT の地域差と世代差:東北から関東の分析」『日本語の研究』 2:2, 34-45.
- 高田三枝子 (2008) 「日本語語頭有声閉鎖音の VOT に関する全国的分布パターン」『日本語の研究』 4:4, 48-62.
- 高田三枝子 (2011) 『日本語の語頭閉鎖音の研究-VOT の共時的分布と通時的変化』 くろしお出版.
- 日本音響学会「音のなんでもコーナー」<https://acoustics.jp/qanda/answer/101.html> (2021年10月25日)
- 橋本健一 (2010) 「反応時間計測実験における外れ値の取扱い: L2 心理言語実験の場合」『より良い外国語教育研究のための方法 外国語教育メディア学会 (LET) 関西支部メソドロジー研究部会 2010 年度報告論集』, 133-145.
- 橋本誠・北川敏・樋口宜男 (1998) 「音声の個人性知覚に影響を及ぼす音響的特徴の定量的分析」『日本音響学会誌』 54:3, 169-178.
- 平井明代 (2012) 『教育・心理系研究のためのデータ分析入門』 東京図書出版.
- 福岡昌子 (1995) 「北京語・上海語を母語とする日本語学習者の有声・無声破裂音の横断的および縦断的習得研究」『日本語教育』 87, 40-53.
- 福岡昌子 (2011) 「中国 (北京・上海方言) および韓国 (ソウル方言) 日本語学習者の破裂音習得-知覚と生成における共通性と相違性-」『三重大学国際交流センター紀要』 6, 11-29.
- 邊姫京 (2017) 「韓国語ソウル方言における語中閉鎖音の音響特徴」『音声研究』 21:2, 61-79.
- 邊姫京 (2019) 「日本語における語頭閉鎖音の音響特徴-VOT と後続母音の *fo*-」『音声研究』 23, 174-197.
- 本間弥生 (1985) 「日英語破裂音の Voice-Onset Time」『梅花女子大学開学二十周年記念論文集』, 383-401.
- 山崎亜希子 (2014) 「ソウル方言における語中母音間破裂音の音響音声学的特徴-三項対立を支える音響特徴に関する考察-」『言語・地域文化研究』 20, 121-133.
- 山中盾 (2000) 「声帯振動の解析における電気声門図 (EGG) 及び光電声門図 (PGG) の有用性-声帯超高速デジタル撮影法による検討-」『日耳鼻』 103, 905-915.
- 山本富美子 (2004) 「日本語談話の聴解率と破裂音の知覚との関係-中国語北方方言話者と上海方言話者に対する比較調査より-」『音声研究』 8:3, 67-79.

山本富美子 (2009) 『第2 言語の音韻習得と音声言語理解に關与する言語的・社会的要因』
ひつじ書房.

横山和子 (2000) 「中国語話者の日本語閉鎖音習得における困難点-有標性と無標性の観点
から」『多摩留学生センター教育研究論集』 2, 1-11.

李通 (2021) 「中国北方方言話者の日本語有声破裂音に対する日本語母語話者の知覚」神戸
大学国際文化学研究科修士研究レポート (未公刊) .

劉佳琦 (2005) 「中国 (北方・上海) 方言話者による日本語有声・無声破裂音の知覚に關す
る—考察—初級学習者を対象として」『早稲田大学日本語教育研究』 6, 79-90.

劉佳琦 (2008) 「中国語母語話者における日本語の有声・無声破裂音の混同」戸田貴子編著
『日本語教育と音声』 (pp.141-162) くろしお出版.

謝辞

本研究は筆者が大学時代からずっと関心を持った日本語破裂音の音声的特徴について考察したものである。現在筆者の出身である台湾で使用されている言語については、「国語」とされる台湾華語のほか、台湾閩南語、客家語、先住民諸語などの言語がある。多くの台湾人は、普段の生活から台湾華語が第1言語で、台湾閩南語など他の言語が第2言語として、2つ以上の言語を併用している。このような多言語社会で生まれ育った私は、高校時代が英米文学科、大学時代が日本語文学科で、第二外国語と第三外国語を勉強した。母語の中国語以外の言語を学習していく中で、「どうすれば言語習得が上手くいくのか」を真剣に考えるようになった。その後、台湾の東呉大学大学院では「音声学」、名古屋大学大学院では「第二言語習得」という学問分野に出会った。これを機に、日本語の音声学に興味を持つようになり、博士課程で神戸大学大学院へ留学する決意をした。

これまで日本語の破裂音に関する研究を進めている中で、先生方から沢山のご指導を、ご鞭撻をいただいた。神戸大学に留学してきた以来、ずっとお世話になっていた恩師である朱春躍教授には日々ご指導いただいた。実験音声学の世界に引導してくださり、研究者の道を選んだ私をずっと励ましてくださった先生には、この場を借りて深く感謝の意を表す。

年に5回も行う本コース最大の特徴である集団指導では、専門性が異なる教員の方々から貴重なアドバイスが得られ、自身の研究を深化させることができ感謝を申し上げたい。本博士論文の副査を引き受けくださった大和知史教授と林良子教授、TAをやらせて頂いた柏木治美教授、そして石川慎一郎教授、グリア・ティモシー教授、木原恵美子准教授、芹澤円助教、懇切なご指導に対して深く感謝している。

なお、研究科の先輩・後輩、実験の録音や調査をご協力してくださった友達や学生さん、数多くの方々のご支援により、この論文をまとめることができた。ここで全ての方々の名前を挙げて謝意を表すことができないことをお詫びし、これまでお世話になった方々に改めて深く御礼を申し上げたい。

2022年1月

瞿琦