



感情音声の生成と知覚-日本語母語話者および中国語を母語とする日本語学習者を対象に-

李, 歆玥

(Degree)

博士 (学術)

(Date of Degree)

2020-03-25

(Date of Publication)

2022-03-25

(Resource Type)

doctoral thesis

(Report Number)

甲第7646号

(URL)

<https://hdl.handle.net/20.500.14094/D1007646>

※ 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

感情音声の生成と知覚

—日本語母語話者および中国語を母語とする

日本語学習者を対象に—

2020年3月

神戸大学大学院国際文化学研究科

李 歆玥

感情音声の生成と知覚

— 日本語母語話者および中国語を母語とする日本語学習者を対象に —

所属専攻・コース：グローバル文化専攻・感性コミュニケーションコース

氏名：李 歆玥

指導教員氏名：林 良子 教授

要約

本研究では、中国語を母語とする日本語学習者が日本語音声によって生成した感情はどの程度伝わるか、母語話者が生成した感情音声とどう違うか、背景言語である中国語が学習者の感情音声の表現に影響するについて検討、実験音声学的手法を用いて検討を行った。具体的には、発話された音声の音響分析および声質に関する分析を行い、生成と知覚の関係について明らかにすることを目的とした。学習者がどのように感情音声を知覚するかについては多くの先行研究で議論されてきたが、なぜ学習者の感情伝達の誤解が起りやすいかについて知るためには、学習者がどのように感情音声を生成するかについて検討する必要がある。本論文の第1章で背景と先行研究に言及した上で、第2章では、日本語母語話者が生成した日本語感情音声と、学習者が生成した日本語感情音声の知覚についての分析を行った。日本語母語話者8名および中国人学習者8名により、「喜び」「悲しみ」「怒り」「驚き」「恐れ」「嫌悪」「中立」という7つの感情について日本語で生成した音声を収録し、発話者とは別の日本語母語話者および中国人学習者を対象として、音声が表示感情を判定させる知覚実験を行った。その結果、発話者と聴者の母語が一致している場合の感情同定率は、母語が一致していない場合の感情同定率より高いことが観察された。さらに、日本語母語話者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」および「悲しみ」と「中立」の4ペアの混同が顕著であり、中国人学習者による感情音声では、「悲しみ」と「恐れ」、「喜び」と「驚き」の2ペアが混同されやすいという結果を得た。これらの結果は、「内集団優位性 (in-group advantage)」の考えを支持しており、日本語母語話者の発話の意図する感情は中国人学習者の発話より混同が多いことを示したものである。

第3章では日本語母語話者および中国人学習者による日本語感情音声にあらわれる音響特徴量の違いについて検討した。音響分析に用いた音声資料は、知覚実験で収録した7

感情の発話であった。全てのデータの発話において F0Level, F0Max, F0Span, F0Mean, Duration, Median Intensity の 6 つのパラメータを対象として分析を行なった。F0 については、日本語母語話者と中国人学習者の発話の間では、F0Level, F0Max, F0Span および F0Mean の 4 つのパラメータに有意差が観察された。このことは、日本語母語話者と中国人学習者が感情音声を生成する際、「怒り」「悲しみ」「嫌悪」および「中立」に関して異なった F0 のパラメータを用いている可能性を示唆する。Duration は、中国人学習者の「嫌悪」と「怒り」で差が見られた。この結果は、「怒り」と「嫌悪」は混同されやすい感情ペアであり、中国人学習者は Duration を用いてそれらを区別して発話する可能性があることを示すと考えられるさらに Intensity でも、中国人学習者では「喜び」と「悲しみ」で差が見られ、日本語母語話者とは異なる音響特性を用いていることが明らかになった。

第 4 章では日本語母語話者および中国人学習者が発話した感情音声について生理学的検討を行うために、Electroglottography (EGG) により Open Quotient の違いについて検討した。日本語母語話者による発話と中国人学習者による発話に Open Quotient の有意差が見られた感情は「怒り」と「悲しみ」であった。また、日本語母語話者による発話では、感情間の Open Quotient の差が見られなかった一方、中国人学習者による発話では、「恐れ」と「嫌悪」は他の感情と有意な差が観察された。特に中国人学習者による「怒り」と「悲しみ」は日本語母語話者より tense voice に近いことが考えられる。これらのことから、中国人学習者は「恐れ」と「嫌悪」を他の感情と区別して生成するために Open Quotient を用いている可能性がある。

第 5 章では中国人学習者による中国語感情音声も分析対象に加え、日中両言語に現れる声質的特徴（スペクトル特性の H1-A1, F1F3syn）および音響特徴量の複合的な分析である Oq-valued VRP を行なった。その結果、発話者の母語によって感情表出様式が異なった。中国人学習者が発話した「押し殺した怒り」と「嫌悪」は日本語母語話者より声帯が緊張した発声が見られた。一方、日本語母語話者が発話した「恐れ」「押し殺した怒り」は中国人学習者による日本語と中国語より氣息性の強い発声が見られた。さらに、中国人学習者による日本語と中国語の「激しい怒り」ともに高音域で Open Quotient が顕著に下がった。この特徴は日本語母語話者による発話では見られなかった。

以上の実験結果を踏まえ、第 6 章では感情音声の生成と知覚の関係ならびに感情音声の発声様式における母語影響について考察を行った。まず、中国人学習者は日本語の感情音声の伝達にあたって、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唆された。また、中国人学習者の感情音声は日本語母語話者より緊張した発声で表出されていると言え、このような発声様式は母語である中国語の感情表出方法に影響されていることが明らかとなった。パラ言語情報に関わる音響・生理学的特徴を検討した本研究の知見は、

感情音声表出における言語差を明らかにし、日本語をはじめとした外国語教育においても有益であると考えられる。

目次

第1章 背景と目的	3
1.1. 感情音声の研究概要	3
1.1.1. 音声による感情情報伝達	3
1.1.2. 感情音声の通言語的共通性と文化差	6
1.2. 日本語感情音声研究	9
1.2.1. 日本語母語話者による感情音声研究	9
1.2.2. 非日本語母語話者による日本語感情音声研究	10
1.3. 感情音声に現れる音響特徴量および声質情報	11
1.4. 本論文の目的と構成	15
第2章 感情音声の知覚	19
2.1. 実験の目的	19
2.2. 知覚実験	19
2.2.1. 方法	19
2.2.2. 結果と考察	21
第3章 感情音声の生成	32
3.1. 実験の目的	32
3.2. 音響分析	32
3.2.1. 方法	32
3.2.2. 結果と考察	33
第4章 EGGによる音声生理学的分析	51
4.1. 実験の目的	51
4.2. EGG分析	51
4.2.1. 方法	51
4.2.2. 結果と考察	52
第5章 声質分析	54
5.1. 実験背景と目的	54

5.2. スペクトル分析	54
5.2.1. 音声資料	55
5.2.2. 声質の分析方法	56
5.2.3. 結果と考察	56
5.3. 声質と音響の複合的分析	61
5.3.1. 感情ごとの F0 と OQ の傾向	61
5.3.2. Oq-valued VRP	63
第 6 章 総合考察	66
6.1. 実験結果のまとめ	66
6.2. 生成と知覚の関係	68
6.3. 感情音声の発声様式における母語影響	70
6.4. 今後の展望	70
参考文献	73
付表 1	86
付表 2	89
付録 1	91
謝辞	100

第1章 背景と目的

1.1. 感情音声の研究概要

音声によるコミュニケーションには言語情報だけではなく、発話者の性別、年齢、健康状態、感情などの情報も多数含まれている。これらの言語情報以外の情報も発話意図をうまく伝えるために重要な役割を果たしている。言語情報以外の情報が多く含まれている音声は Expressive Speech と呼ばれる (Erickson, 2005)。ヒトが音声による円滑なコミュニケーションを行なうには、言語情報とともに、発話意図などを表現する情報の表出と理解が不可欠であり、これ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。感情音声の研究は音声工学・心理学・認知科学・神経科学・日本語教育の各分野において論じられている。

本研究では、音声学の立場から、知覚において中国語を母語とする日本語学習者が日本語によって生成した感情音声と日本語母語話者が生成した感情音声はどの程度伝わるのか、そして伝わらない原因は何かについて検討する。また、生成において、これらの音響特徴量と声質の違いは何か、知覚にどのように影響するかを検討する。さらに、学習者の背景言語である中国語は感情音声の表出に影響するかについて検討する。

第1章ではまず音声が伝える情報の一つに発話者の感情があることを説明する。次に音声と感情の関係について、特に音響特徴量と声質にあらわれる感情音声の通言語的共通性と発話者の言語による文化差について概観する。

1.1.1. 音声による感情情報伝達

人間の言語に用いられる音声は、国際音声学会 (International Phonetic Association) が提案した国際音声記号 (the International Phonetic Alphabet: IPA) を用いておおよそ記述することが可能とされている。しかし、音声言語コミュニケーションの過程で発される音声信号には、話し手の性別、年齢、意図、態度、感情などの情報が含まれており、また同じ言語音を用いてもイントネーションや表現の違いによって、その意味は様々に変化する。例えば、「ナニヤッテンノ」という発話の音声信号から音声列を正しく認識できたとしても、

それを発した話し手の意図が「質問」にあるのか「叱責」にあるのか「揶揄」にあるのかを知ることはできない（森・前川・粕谷 2014）。このような言語外とも言える情報は、パラ言語（paralanguage）情報とも呼ばれる。『国際音声記号ガイドブック』（2003）によると、パラ言語について以下のように述べられている。

パラ言語に含まれるのは、声の質、高さの音域、話す速さといった現象を用いて、話者の感情的状態や話し相手に対する態度の様々な側面を伝えたり、伝達されている情報の種類（例えば秘密など）を示したり、相手の発話を促したり抑制したりして会話の進行を司ることなどである（p48）。

上記のように、パラ言語のような言語現象の記述手段を明らかにする必要があると言われており、音声によって伝達される情報を体系的に分類した研究には Trager（1958）の二分法、Fujisaki（1996）の三分法、および森・前川・粕谷（2014）の分類方法がある。

まず古典的な研究に Trager（1958）がある。Trager（1958）は、人間同士の対話音声（interpersonal speech communication）を、言語と周辺言語の二分法を基準とし、言語、声質要素、発声要素に分けた。話し手の感情情報や意図発話などは声質要素から知覚される。発声が伝える情報の中で、発声特徴（vocal characterizers）は笑い、泣き、つぶやきなどであり、発声上の限定性（vocal qualifiers）は声の高さ、速さ、強さなどから成り、発声上の遊離素（vocal segregates）、例えば肯定的、否定的な「うーん」、戸惑いの「うー」を設けている。

次に Fujisaki（1996）は発話者が情報伝達を意図的に制御できるかどうか、意図の強さに連続的な変化が認められるかどうかという新しい基準を取り入れ、言語情報（linguistic information）、非言語情報（non-linguistic information）およびパラ言語情報（paralinguistic information）の3つに分類した。パラ言語情報とは、ここでは書き言葉に転写すると推測不可能となる情報で、言語情報を補助ないし変容するために話者が意図的に生成する情報としている。言語記号に基づいた離散的な言語情報は、発話者の態度など心の意図と関わる量的かつ連続性のあるパラ言語情報と区別することができる。例えば前川・北川（2002）と Maekawa（2004）はパラ言語情報として用いられた「失望」の強さの度合いは連続的であると認められる。非言語情報は身体的状態、性別や個性が関わるため、制御不可能かつ離散的な特徴を持っており、言語情報とパラ言語情報を区別することができる。

森・前川・粕谷（2014）は Fujisaki（1996）の三分法では感情の位置付けが十分でないことを指摘した。Fujisaki（1996）の分類では、感情は性別や年齢と同様、制御のできない情報として扱われたが、森・前川・粕谷（2014）は現実の話し言葉では感情が意図的に伝達されることが確かであると述べ、メッセージとして意図的に表出される感情¹と不随意的に表出される感情を分け、前者を Fujisaki（1996）のパラ言語情報とまとめて「パラ言語メッセージ」としている。具体的に言うと、日常生活のコミュニケーションにおける人間の感情表出には、不随意的なもの—例えばクマにでくわして感じた恐怖など、外部からの刺激による身体の自動的プロセスによるものおよび楽しかった出来事を思い出すような内部から起こるものだけでなく、話し手が意図的に選択した「メッセージ（意図を符号化したもの）」としての性格を持つものもある。感情音声研究のために用いられる音声資料の多くは発話者が意図的に発話したものであり、感情について本研究は森・前川・粕谷（2014）と同様の立場をとる。つまり、メッセージ性の持ち意図的に生成される感情音声をパラ言語メッセージとする。

続いて音声研究における感情の定義を音声学的な視点からまとめる。まずは南（1958）によれば、「感情」「情動」「気分」「センチメント」のそれぞれの定義は次の通りである。感情とは、欲求満足が順調に進まなくなった時に生ずる自己評価であり、情動とは「感情」がかき乱され動揺している状態である。気分とは、「感情」が比較的長い間継続している状態である。センチメントとは、「感情」に特定の知的評価が加わったものとしている。

次に、『最新 心理学事典』（2013）では以下のように「感情」を説明している。情動の項目はなく、感情の項目で用語の関係が整理されている。

感情に *feeling* を、情動に *emotion* を当てることも多かった。*Feeling* は経験のもたらす感触という評価的側面を意味するから、感情に当てられるのは自然といえる。一方、*emotion* は身体的、生理的表出を伴う興奮状態を指しているが、現象学派のようにその主観的側面を強調すれば、当然感情と大きく重なり合い、喜怒哀楽などの激烈で一過性の感情といった解釈が生まれる。本書

¹ 例えば社会生活の中で、相手は悲しい話をしている時は、声のトーンを下げ一緒に悲しそうな声を出す。相手が怒っている話の時は一緒に怒って見せ、笑っている時は喜んで見せる。このような感情は意図的に表出されたものであり、一種の演技ともいえよう（森・前川・粕谷 2014）。

の「感情・意志」分野において、原則として感情は「文化に関連した人間の意識構造」を説明する場合、情動は「生理的動因に関連する心理的プロセス」を説明する場合の用語として使用する。なお、感情と情動は、意識の中でスペクトラム構造をなしており、厳密に区分できない場合には、「感情・情動」と併記する (p89)。

Ekman や Izard は心理学の観点から少数の基本感情を仮定し、それらの感情に特有の表情を研究した (Ekman 1992, Izard 1977)。彼らの研究手法は感情研究に大きな影響を与えたため、音声と感情の研究でも基本 6 感情 (“big six” emotions と呼ばれるが、必ずしも定まっていない) などの基本感情を仮定するものが多い (森・前川・粕谷 2014)。現代の感情音声研究を代表する 1 人である Scherer の研究もその例外ではない (Scherer 1986, 1999, 2003, Scherer et al. 1984, 1991, 2001, 2011, 2015)。Ibrakhim (2005) によると、感情と感情情報の関係は現象とその記述であり、コミュニケーションにおいては、感情情報はコミュニケーション過程と関わっているものであり、感情そのものではないと述べている。また、郡 (1997) はイントネーションを意味機能から、文法的機能、情緒的機能、社会的機能の 3 つに分類しており、情緒的機能は話し手の感情及び心理状態を表す機能としている。さらに、赤木 (2010) は多次元空間としての感情空間のモデルを提案し、感情知覚空間は連続であり、高い同定を得ている感情音声はカテゴリをなすように見えるが、他の感情音声は広く連続的に分布していることが見られたと報告している。

1.1.2. 感情音声の通言語的共通性と文化差

本節では、感情音声の通言語的共通性と文化差、および音響特徴について述べる。英語の音声研究においては、表情と同様に、喜び、怒り、悲しみなどの感情の知覚に文化を超えた共通性があるということが指摘されている (Ekman & Friesen 1969, Tickle 2000)。一方で、感情の知覚と生成は母語話者と非母語話者の間で社会文化的要因による異なる傾向があることも報告されている (Banse & Scherer 1996)。下記のようによく知られている感情音声の研究を例として挙げる。

Van Bezooijen et al. (1983) はオランダ語母語話者が発話した「喜び」「驚き」など 10 種類の感情音声を収録し、それをオランダ語母語話者、日本語母語話者、台湾中国語母語話者に聞かせ、感情を同定させたところ、全ての聴者がチャンスレベル以上の確率で感情を正しく判定することができた。しかし、「喜び」に対して、オランダ語母語話者の同定率が高い一方、日本語母語話者および台湾中国語母語話者の同定率が低いなど、感情によって母語話者と非母語話者の同定率に異なる傾向が認められた。

Scherer et al. (2001) はドイツ語母語話者が生成した「喜び」「怒り」「悲しみ」「恐れ」「嫌悪」の感情音声をを用い、ドイツ語母語話者、フランス語母語話者など 9 つの言語の母語話者に聞かせ、知覚実験を行なった。その結果、母語の言語的距離 (language dissimilarity) が小さいほど感情の同定率が高くなり、大きいほど感情の同定率が低くなることを明らかにしている。

Campbell & Erickson (2004) では日本語母語話者による様々な発話音声に対し、日本語母語話者、アメリカ英語母語話者、韓国語母語話者に聴覚印象に基づいて自由に分類させ、ラベリングさせるという実験を行ったところ、同じ母語話者間でラベリングの傾向が類似しており、一方、異なる母語話者間の頻繁に使われるラベリングとそうでないラベリングが異なっていた。たとえばアメリカ英語母語話者のラベリングで最も多かったのが「幸福」であり、韓国語母語話者のラベリングで最も多かったのが「飽き飽きした」であった。著者らは、このような知覚の相違が感情音声知覚の誤解を引き起こす可能性があるとは指摘している。

Pell et al. (2009) はイギリス英語、ドイツ語、スペイン語およびアラビア語の 4 カ国母語話者による「中立」「怒り」「嫌悪」「恐れ」「悲しみ」「喜び」の 6 種類の無意味語感情音声を収録し、各発話者グループと同じ母語の聴者に聞かせたところ、発話者と聴者の母語が一致している場合の感情同定率は、母語が一致していない場合の感情同定率より高いことを報告し、「内集団優位性 (in-group advantage)」を提案している。その原因として、各文化の社会規則により、同じ母語背景における感情音声の表出と知覚がよりスムーズに成り立つと解釈している。

さらに、Chronaki et al. (2018) は 8-10 歳、11-13 歳および成人を聴者とし、イギリス英語、スペイン語、中国語とアラビア語による「中立」「喜び」「怒り」「悲しみ」「恐れ」の感情音声をを用いて知覚実験を行なった。その結果、成人だけではなく、児童の感情音声知覚にも「内集団優位性」を観察できたと述べている。

李他（2018）は、日本語母語話者が「喜び」「怒り」「悲しみ」「平静」の感情を込めた発話に対して、中国人日本語学習者がどのように知覚するかについて検討した。その結果、学習歴が短い被験者では、日本語の感情知覚の正答率が高くはないが、学習歴の長い被験者では、正答率が高くなっていることが示された。また、4つの感情音声の中では「平静」の正答率が一番高く、「喜び」ではの知覚が最も困難であることが観察された。これらの結果から、非母語話者においても感情音声の学習が可能であることと、感情の種類によって知覚の難易度が異なることが示された。

また、研究方法として、宮島他（2013）は演技音声を「朗読型の演技音声」と「自発的演技音声」に分類している。「朗読型の演技音声」とは典型的な基本感情語などをタスクとして与えて、発話者に感情表現を統制して発してもらった音声である。一方、「自発的演技音声」とは日常生活の中で「店員」「教師」などの役割を与えて、発話者に発してもらった音声である。ここでは『SEN』（「発話時の背景」「話し手」「聞き手」など具体的な指示を示した台本による自発的演技音声）と『Typical』（比較用として基本感情語だけを提示して収録した音声）の音響的、心理的な分析を行っている。その結果、適切な演技指示である台本を用いた『SEN』のほうが「多様な感情表現」を獲得できたと報告している。Maekawa & Kikuchi（2005）では自発音声の定義は朗読音声の反対概念として考え、発話の開始時点において発話内容が完全には決定されていない発話としている。そして、自発音声の最大の問題点は、データとして組織的統制をかけていることであると指摘し、「自発音声データベース」を提案した。あらかじめ実験テキストを制約することは不可能であるが、データの量が十分であれば研究価値の高い質を備えた実験テキストが含まれる可能性が大きくなり、自発音声の音声学的研究の可能性を広げることができるとしている。著者らは、このような学術的背景によって開発された総語数 752 万以上にのぼる『日本語話し言葉コーパス』（CSJ）を用い、自発音声の分析結果を報告している。そこでは、発話速度の分布は自発音声と朗読音声の最も見逃すのできない相違点の一つであると述べる他、スペクトル特徴、音声変異およびデータベースの限界についても論述されている。

以上紹介してきた感情音声の研究では、感情が音声に与える影響は通言語的な共通性として心理的活性度の高さと声の高さ、強さと関連するとされる。文化差として社会と文化的要因による影響を受け、その感情伝達に誇張や抑制が見られ、このような生成と知覚の相違が感情音声知覚の誤解を引き起こす可能性があるという指摘がある。

1.2. 日本語感情音声研究

1.2.1. 日本語母語話者による感情音声研究

日本語母語話者による感情音声研究については、飯田（1997）がよく知られている。飯田（1997）は「喜び」「怒り」「驚き」「悲しみ」「嫌悪」を対象に、2名の発話者による「ありがとう」「こんにちは」「ちばたつや」「何言ってるの」の4つのタスク文を収録し、基本周波数レンジ、持続時間長とインテンシティ平均値などについて検討した。その結果、「喜び」と「驚き」の基本周波数レンジとインテンシティ平均値が他の感情種類より高く、「怒り」の持続時間長が他の感情種類より短かったことが報告されている。

重野（2001）では、2名の日本語母語話者が発話した「東京」「河原崎さん」「11時半」「さようなら」「そうですか」に含まれる「幸福」「驚き」「怒り」「嫌悪」「恐れ」「悲しみ」の6つの感情発話の生成について分析した。その結果、基本周波数とエネルギーの標準偏差が、「怒り」では他の感情より高く、「嫌悪」ではもっとも小さく、大きな差異が認められた。

田川（2001）は、第1拍目の後に下降が来る「見ないの」という音声の「平静の問いかけ」と「非難の問いかけ」について分析を行った。「非難の問いかけ」の冒頭部分から上昇部分までのピッチ曲線は「平静の問いかけ」より急な傾斜が見られ、「非難の問いかけ」の全体時間長と各音節の時間長は「平静の問いかけ」より長いと述べている。この2つの研究から日本語の感情音声生成には、発話のピッチパターンや持続時間、声の大きさが大きく関与していることが分かる。

さらに、前川・北川（2002）および Maekawa（2004）は「中立」「疑い」「落胆」「強調」「無関心」の6種類のパラ言語情報について、生成実験と同定実験を通して持続時間、ピッチ、振幅、母音の音質などを分析した。その結果、非言語情報における音響特徴が発話全体にわたって実現されるのに対し、パラ言語情報に関する音響特徴は局所的にも実現されうることを示した。

また、感情表現と調音運動の関連を調査した先行研究として赤井（2017）がある。この研究では感情音声発話時の調音運動に関して、演技感情音声発話時には、言語情報の特徴を保持したまま感情表現を行った。その結果、言語情報の生成に重要でない調音運動が感情表現に対して一定の値に変化すること、それに対して、自発感情音声発話時には、上記

のような調音運動の使い分けは行われず、生成する音声と言語的に異なる音であるという特徴が、強調するような調音運動に反映されることが示された。石井他（2006）は自然対話音声において、声質についてさらに分析を進め、喉頭が力んだときなどに起こる不規則な声帯振動による vocal fry 音が「感心」などに関連していることや、「驚き」など高ぶりやすい感情には harsh 発声が見られることなどが観察されたと報告した。

1.2.2. 非日本語母語話者による日本語感情音声研究

非日本語母語話者を対象とした日本語感情音声研究もいくつか行われている。櫻庭他（2001, 2004）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、音声が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と報告した。

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

中林（2009,2011）では、「韓国」「毎日」（頭高型）、「現金」「片仮名」（中高型）、「カラオケ」「温泉」（平板型）の6つの4拍語をタスク語とし、問い返し疑問文に含まれた「嫌」と「驚き」の感情について、ロシア人日本語学習者を対象に分析を行った。その結果、「嫌」という感情の認識が困難であることが示された。また、アクセント核の有無によって感情の聴取の難易度が異なる可能性、さらに正答率の低い学習者では特定のアクセント型からの感情の聴取が困難である可能性が示唆された。

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%以上の音声の数は、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。このように非母語話者が日本語音声から感情を知覚する際、母語話者と同様の傾向も認められるが、聞き手の母語が日本語感情音声の知覚に影響する可能性も示された。これらの研究のいずれにおいても、Pellらによる「内集団優位性」が妥当であることを裏付けてはいるものの、聴者の母語や感情ごとに異なる影響を及ぼす可能性を示唆していると考えられる。

1.3. 感情音声に現れる音響特徴量および声質情報

感情音声の音響的な分析もこれまで多く行われてきた。感情は、主に声の高さ（基本周波数 F0: Fundamental frequency）、声の強さ（インテンシティ Intensity）、声の長さ（持続時間 Duration）および声質に現れるとされる。先行研究において具体的には、音声の音響的特徴量として、基本周波数 F0 に関する諸パラメータ（F0 平均値、F0 最大値、F0 最小値、F0 レンジ、F0 標準偏差、F0 ゆらぎ、F0 曲線形状など）、持続時間長（Duration）、インテンシティ（Intensity）に関する音響特徴量パラメータ、声質の分析には、声門時間開放率（Oq: Open quotient）、ジッタ（声の高さのゆれ Jitter）、シマー（声の大きさのゆれ Shimmer）、スペクトル傾斜特性、フォルマント周波数、フォルマント帯域幅などのパラメータが用いられてきた（野田他 2005、藤江他 2005 他）。

例えば英語感情音声「喜び」の F0 平均値、F0 レンジ、F0 標準偏差は「中立」より大きく、持続時間長は「中立」より短い。これらは活性次元と緊張した声に関連するパラメータとして知られている（Scherer 1986）。Erickson et al. (2016) は「怒り」「悲しみ」「喜び」と「中立」について、日本語母語話者 2 名の発話者が生成した 3 つのタスク文を録音し、音響分析を行なった結果、「喜び」と「怒り」の F0 平均値とインテンシティは「中立」と「悲しみ」より大きかった。また、中国語感情音声の研究では、Yang et al. (2005) は「怒り」「恐れ」「喜び」「悲しみ」「驚き」を対象に、1 名による 4 つのタスク文を分析した

結果、「恐れ」のF0平均値は「悲しみ」より高く、「喜び」の持続時間長が「怒り」より長かったと報告している。

森・前川・粕谷（2014）によると、「怒り」「驚き」のような心理的活性度の高い感情においては、声の高さと強さは大きくなる一方、「悲しみ」など心理的活性度が低い感情においては、声の高さと強さは小さくなるとされる。このような変化は、呼吸筋や喉頭筋などの緊張亢進、減退といった感情の生理的反応を反映したもので、どの言語でも同じような傾向を示し、基本感情の音声による表出は、学習よりは生得的なものである傾向があると述べている。

また、音声によって他者への伝達は意図的に制御できる場合があり、さらに社会と文化的要因による影響を受け、その感情伝達に誇張や抑制が見られる（浅井，2017）。Ibrakhim（2005）では、異なる言語の話者は、同じ韻律特徴の変化について異なる解釈をする場合があり、「感情情報」とは、絶対的なものとして存在するのではなく、聞き手の側にどう受け止められるかによる側面もあると報告している。発話における「感情情報」は、発話者のその瞬間の心理状態、ムード、発話内容または聞き手への態度についての情報で、感情やムードと個人の性格の相互影響と働きの結果として捉えることができる立場をとっている。このことは、聞き手は音声を総合的な情報として知覚し、聞き手の言語文化により理解、解釈している可能性を示唆する。

声質²は話者の個人性に関する重要な情報として研究されてき、対象が音声に含まれる言語情報の範囲を超え、非言語情報とパラ言語情報にまで広がったことの反映である（Fujisaki 1996）。そしてその定義はLaver（1980）と前川・西川（2019）が見られる。Laver（1980）では、話し手が程度の差こそあれ、話している間中ずっと存在する特性であり、発声された音声にほぼ永久てきに存在する質は声質である。そして、前川・西川（2019）によると、文節音よりも広い範囲において観察される話し手に固有の音声的特徴であり、その中には、発声（Phonation）の特徴と調音（articulation）の特徴がともに含まれているとされている。

声質の音声学的記述には、Laver（1980）が提案した声質記述体系がよく知られている。Laver（1980）では音声器官の生理、物理的な個人差に起因する声質特徴を排除し、発話者の言語とともに獲得した経験的な点に焦点を当て、modal, breathy および whispery, vocal fry

² 狭義での声帯振動モードによる声の質“laryngeal voice quality”を指す。

および creaky, harsh および ventricular, またはこれらの組み合わせについての提案がなされた。国際音声学会が定めた IPA や拡張 IPA (ExtIPA) の記号あるいは補助記号としては、裏声 (falsetto), ささやき声 (whispery), 息漏れ声 (breathy), きしみ声 (creaky), 耳障り声 (harsh) などがあるが, これらは Laver (1980) の声質記述用語を基本にしたものである (粕谷・楊 1995)。また, 日本音声言語医学会では, 粗造性 (rough), 気息性 (breathy), 無力性 (asthenic) と努力性 (strained) の四つを嗄声の声質記述として提案している (今泉 1994)。さらに, 緊張—弛緩, 病的—健常などの尺度もよく用いられる (榊原 2014)。

声質は発声の特徴としては, きしみ声や息漏れ声での発話が挙げられる (前川・西川 2019)。本研究は, 発声様式に限定する意味で用いる。また, 声質は調音の例としては, 発話全体の鼻音化, 調音空間の前後へのシフトなどが挙げられる (前川・西川 2019)。言語によって, 声質的特徴は発音を区別する現象が多く見られ (Gordon & Ladefoged 2001), 例えばある言語の中で病理学的発音 (pathological voice) とみなされる声質は別の言語では有意味音素として使われる (Ladefoged 1983)。

石井他 (2006) は, 韻律および声質に関連する音響特徴とバラ言語情報 (「意図」「感情」「態度」を総称する) の関連について, 台本をもとにした演技対話音声と CREST/ESP の自然対話音声データベースにおける自然対話音声両方を対象に, 発話「え」に着目して知覚実験と音響分析を行った。その結果, 演技対話音声において, 韻律を表現した音響パラメータ (F0move, 発話持続時間) は聞き返し, 肯定的, 否定的な表現のような機能的なバラ言語情報を検出するのに貢献している一方で, non-modal な声質を表現した音響パラメータ (気息性, 非周期性など) は「驚き」「嫌悪」など比較的強い感情や態度を表すバラ言語情報を識別するのに貢献していることを示した。この結果から, 石井他 (2006) は modal (地声), whispery (気息性のある声), aspirated (発話末に現れる強い息漏れ), creaky (基本周波数が通常発声よりも低くパルスのな声), harsh (雑音的で耳障りのある声), pressed (喉頭を力んだ声) およびこれらの組み合わせを用いた声質の表現を提案した。

声質に影響を与える声帯振動の状態を調べる方法の一つに Electroglossography (EGG) がある (図 1)。EGG とは左右の甲状軟骨外側の皮膚面に電極を置いて高周波電流を流し, 声帯開閉による電気的インピーダンスの変化を記録するものである。EGG は非侵襲的で, 操作が容易で, かつ装置が高価でないなど, 音声の研究のみならず臨床においても優れた特徴を持っている機材としてよく知られている (石毛他 1996)。



図 1-1 Electroglottography (EGG)³

声帯振動の 1 周期において、声門閉鎖の時間を閉鎖期、声門が開いている時間を開放期と言う。開放期のうち、声帯が外向きに運動し声門面積が増加する時間を開大期 (t_o)、声帯が内向きに運動し声門面積が減少する時間を閉小期 (t_c) と呼ぶ。基本周期 (開放期 + 閉鎖期) を T と記すと、声門開放時間率 (Open quotient) O_q は公式 (1) で、声門開閉速度率 (Speed quotient) S_q は公式 (2) で表される (榎原 2015)。

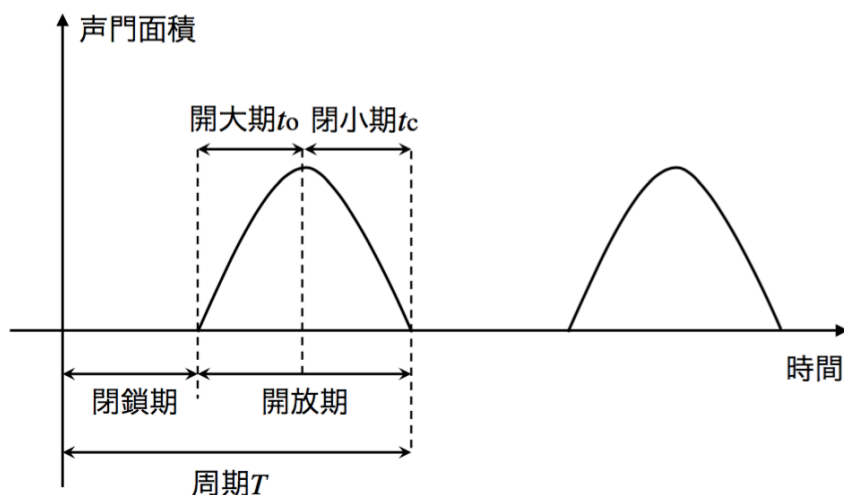


図 1-2 声門面積の変化と時間区分 (榎原 (2015) を参考に作成)

³図 1-1 では、装置の例として Glottal enterprises 社の EG2-PCX2 を示す。
<https://www.glottal.com/Electroglottographs.html> (2019 年 10 月 20 日閲覧)

$$O_q = \frac{t_o + t_c}{T} \quad (1)$$

$$S_q = \frac{t_o}{t_c} \quad (2)$$

EGG 分析の応用として、Ishi et al. (2010) では、EGG 信号の分析により、りきみ声の特徴は F0 あるいは周期性によらず、声門の閉鎖区間の時間長が開放区間より著しく長い特徴を持つことであるとし、この特徴を利用してりきみ声の区間検出が可能であると報告されている。

声質の音響関連量としては、さらに倍音を用いた関連の音響量が利用されることもある (前川・西川 2019)。第 1 倍音 (=F0 基本周波数) のエネルギーを H1 としたとき、第 2 倍音のエネルギーを H2、F1、F2、F3 のエネルギーを A1、A2、A3 とすると、H1-H2、H1-A1、H1-A2、H1-A3 など定義される音響関連量が声質に関係しているとされる。これらの特徴と発声様式との関係に関する研究には Gordon & Ladefoged (2001) 等がある。歌声メカニズムに関する研究にもこれらの特徴量が用いられており、例えば Yoshinaga & Kong (2011) では、喉詰型発声の特徴は低い声門開放時間率、高い声門開放速度率および低い H1-H2、H1-A3 であると報告している。

Scherer (2003) の研究方法にならい、日本語母語話者による一部の先行研究の音響的特徴をまとめたものを付表 1 に示す。本研究では中国語を母語とする日本語学習者についても分析するため、背景言語である中国語の一部の感情音声研究の音響特徴についてまとめたものも付表 2 に示す。

以上紹介してきた感情音声の研究をまとめると、感情が音声に与える影響は、通言語的な共通性として心理的活性度の高さと声の高さ、強さと関連するとされる。ただし、感情音声は、文化差として社会と文化的要因による影響を受け、感情の伝達には誇張や抑制が見られるため、これまで述べてきたような生成と知覚の相違が感情音声知覚の誤解を引き起こす可能性がある。

1.4. 本論文の目的と構成

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る (Erickson 2005)。母語における感情音声の仕組みについてはまだ未解明な点が多いが、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、学習者による外国語発音における感情音声の実態についても研究する必要がある。しかし、これまでの日本語教育においては、感情音声の聞き取り教材 (西端 1996) がわずかながら開発されているものの、感情は普遍的なものであるため指導する必要があるという認識が通用しており、日本語教育の現場において感情の表出、理解が扱われることがほとんどなかった (中林 2011)。そのため、日本語学習者による感情音声の生成、知覚についてはこれまでにあまり多くの研究が行われてこなかった。

日本語を非母語とする日本語話者が国内外で増える今日、世界で最も人数の多い中国語を母語とする日本語学習者⁴の日本語音声指導は重要である。しかし、学習者の一番言いたいことが伝わるような音声を生成するために何が必要なのか、学習者の日本語生成と知覚がどのように改善できるかについては、未解明の部分が多い。このことは、日本語母語話者と非母語話者によるコミュニケーションで誤解が生じさせる原因となっていると考えられる。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。

本研究では、このような背景をもとに、中国語を母語とする日本語学習者が日本語音声によって生成した感情は母語話者とどう違うか、どの程度伝わるか、伝わらない原因が何か、また音響分析および声質に関する検討を通して生成と知覚の関係について明らかにすることを目的とする。

まず、第2章では、日本語母語話者が生成した日本語感情音声は、学習者が生成した日本語感情音声より正しく聞き取られるかどうか、また異なった知覚パターンが示されるかについて検討する。このため、日本語母語話者および中国語を母語とする日本語学習者が生成した日本語感情音声を収録し、発話者とは別の日本語母語話者および中国人学習者を聴取者として、音声を示す感情を判定させる知覚実験を行なう。

⁴ Japan Foundation (2015) によれば世界でもっとも日本語学習者が多いのは、中国語を母語とする日本語学習者が 953,283 人、ついでインドネシア語を母語とする学習者で 745,125 人、韓国語を母語とする学習者で 556,237 人となっている。

感情音声の生成には、感情が音声に与える影響が音響的側面から議論されている (Bachorowski & Owren 1995, Banse & Scherer 1996 他)。これらの研究においては、感情音声には声の高さ、長さ、強さなどの複合的な音響特徴量が関与しており、各言語によってその音響的特徴量が異なることも指摘されている (Johnstone 2001, Scherer 2003)。感情音声の知覚に寄与する音響的特徴量が多数あることも認められている (Murray & Arnott 1995, Burkhardt & Sendlmeier 2000)。そこで第3章では、第2章で観察できた結果の原因を明らかにするため、知覚実験に用いた感情音声を対象に音響分析を行なう。日本語母語話者および中国人学習者による感情音声にあらわれる音響特徴量(基本周波数最小値(F0Level), 基本周波数最大値(F0Max), 基本周波数の範囲(F0Span), 基本周波数平均値(F0Mean), 持続時間長(Duration) およびインテンシティの中央値(MedianIntensity))の6つのパラメータを抽出し、発話者グループ間および感情間の違いについて検討する。

感情音声においては、音響特徴量以外の要因、特に声質が大きく関わっている (Gordon & Ladefoged 2001, Menezes & Maekawa 2006, Erickson et al. 2011, Scherer et al. 2015) ため、第4章は、EGGを用いて声帯運動を計測することで、感情音声を音声生理学的側面から観察し、中国語を母語とする日本語学習者が日本語で生成した感情音声と日本語母語話者が生成した感情音声は異なった発声様式が示されるのか、主にEGG信号によるOpen Quotientの違いは何かについて音声学的分析を通して検討する。また、感情音声の知覚と生成の関係性を明らかにするために、知覚実験で得られた同定率と各音響的特徴、声質との関係についても検討を行なう。

第3章と第4章で見られた結果は中国人学習者の感情音声の音響的特徴が背景言語である中国語の影響によるものなのかについて明らかにするため、中国語の感情音声について同様の分析を行なって検討する必要があると考えられる。また、スペクトル特性に関する分析も加え、声質と音響特徴量を複合的に検討する必要があると知られている (Gobl & Chasaide 2003)。このため第5章では、さらに、日本語母語話者が生成した日本語感情音声と中国語を母語とする日本語学習者が生成した日本語と中国語感情音声に現れる声質的特徴(スペクトル特性とEGGによるOpen Quotient)および音響特徴量の複合的な比較をすることで、日中両言語の感情音声の発声様式の違いは何か、また、中国語を母語とする日本語学習者の感情表出様式が背景言語の影響によるものなのかについて検討する。スペクトル特性のパラメータとして、基本周波数と第1フォルマント周波数帯域の振幅値の差(H1-A1)、第1フォルマントと第3フォルマント周波数周辺帯域でフィルタリングした

振幅包絡の同期性 (F1F3syn) について測定を行ない、発話時の Open Quotient の変動を視覚的に捉える Oq-valued VRP (Voice Range Profile) を用いて、同程度の F0 と power における Open Quotient が発話者グループによってどのような違いがあるかについて考察した結果についてまとめる。

最後に、第 6 章ではこれらの結果をまとめ、総合考察を行う。

第2章 感情音声の知覚

2.1. 実験の目的

本実験では以下の三点を目的とする。第一に、中国語を母語とする日本語学習者は日本語音声に含まれる感情を一体どの程度知覚できるのか、第二に、学習者が生成した日本語感情音声は日本語母語話者にどの程度伝わるか、第三に、学習者に混同されやすい感情パターンは何かについて音声学的分析を通して明らかにすることを目的とする。

2.2. 知覚実験

知覚実験では日本語母語話者と中国人学習者が生成した日本語感情音声を収録し、発話者とは別の日本語母語話者と中国人学習者を聴者として、音声を聞いて感情の判別をってもらう実験を行った。

2.2.1. 方法

感情の選定

図 2-1 はこれまでの感情音声に関する論文 32 本に取り上げられた感情の種類と頻度を示している。先行研究においては Ekman (1992) の基本 6 感情に基づき、「怒り」「悲しみ」「喜び」「中立」「驚き」「恐れ」「嫌悪」を対象とするものが多かった。多くの研究において、「中立」は他の感情と比較するためのベースラインとして用いられていた。本実験ではこれにしたがい、Scherer (1999) の感情選択と同様に、次の 7 種類の感情、「中立(Neutral)」「喜び(Happy)」「怒り(Angry)」「悲しみ(Sad)」「驚き(Surprised)」「恐れ(Afraid)」「嫌悪(Disgusted)」を選定した。

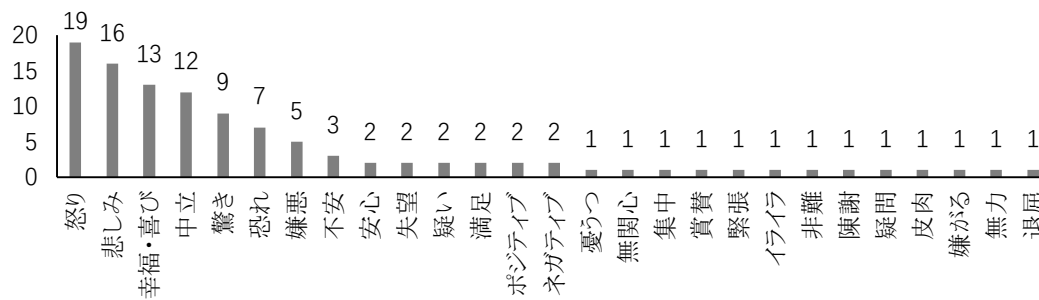


図 2-1 先行研究で取り上げられた感情の頻度

タスク語の選定

宮島他（2013）などを参考に、次の点に留意してタスク語を選定した。まず、表現の多様性が文意に影響を受けないようにするため、単語自体に感情のない有意味単語を用いた。次に、アクセント核の有無は日本語学習者の感情知覚に影響を与える可能性があるため、無核アクセント語と有核アクセント語の両方を用いた。さらに、ピッチ曲線をきれいに抽出できるように摩擦音と破裂音を排除し、無声子音を含まないようにした。これらの基準にしたがい、以下の計 11 個の単語をタスク語として用いた。

- 1 モーラ：え，目
- 2 モーラ：海，桃
- 3 モーラ：バナナ，匂い，まなみ
- 4 モーラ：今にも，生もの，目の前，青山

音声資料の収録

発話者は 16 名であった。日本語母語話者 8 名（男女各 4 名，平均年齢 23.94， $SD=2.81$ ）との中国人日本語学習者 8 名（男女各 4 名，平均年齢 25.43， $SD=3.26$ ，全員日本語能力試験 N1 合格者）であった。

前川・北川（2002）を参考に、収録に先立って、「中立」は「何の感情もこもっていない棒読み」と説明した。音声資料は、マイクを用いて唇と一定の距離を保ちつつ、静かな実験室内で収録した。16 名の発話者に、できるだけ自然かつ明確に感情を表出するように指示し、各感情とも 1 文につき 3 回発話してもらった。1 回目と 3 回目の録音を除き、2 回

目の録音を音声データとして分析の対象とした。合計 1,232 トークン (7 感情×11 タスク語×16 名) の音声データを収録した。

知覚実験の手順

聴者は、発話者とは別の日本語母語話者 12 名 (平均年齢 27 歳, $SD = 6.41$) および中国人日本語学習者 12 名 (平均年齢 29 歳, $SD = 3.29$, 全員日本語能力試験 N1 合格者) が知覚実験に参加した。

収録した音声資料を用いた。知覚実験は静かな実験室内で行なわれ、ヘッドフォンを通して刺激音を提示した。Praat の ExperimentMFC を用いて、音声資料をランダムに並べ替えたものを聞いてもらい、音声が示す感情と思われるものを 7 つの感情から 1 つだけ選ぶように教示した。テスト 1 の音声資料は母語話者による 616 トークン (7 感情×11 タスク語×8 名) である。テスト 2 の音声資料は学習者による 616 トークン (7 感情×11 タスク語×8 名) であった。実験は、練習、テスト 1、テスト 2 という順で行なった。実験時間は約 60 分間であった。

2.2.2. 結果と考察

知覚実験では、得られた回答から、発話者が意図した感情と聴者が同定した感情が一致した回答を正答とし、同定率を求めた。発話者 2 水準 (母語話者, 学習者) × 聴者 2 水準 (母語話者, 学習者) × 感情 7 水準 (中立, 喜び, 怒り, 悲しみ, 驚き, 恐れ, 嫌悪) の 3 要因の分散分析を行った。その結果、発話者と感情の主効果が有意であった ($F(1, 22) = 9.15$, $p = .006$; $F(6, 132) = 76.83$, $p < .001$)。さらに、聴者×発話者、発話者×感情、聴者×発話者×感情の交互作用が有意となった ($F(1, 22) = 16.86$, $p < .001$; $F(6, 132) = 9.01$, $p < .001$; $F(6, 132) = 2.77$, $p = .015$)。

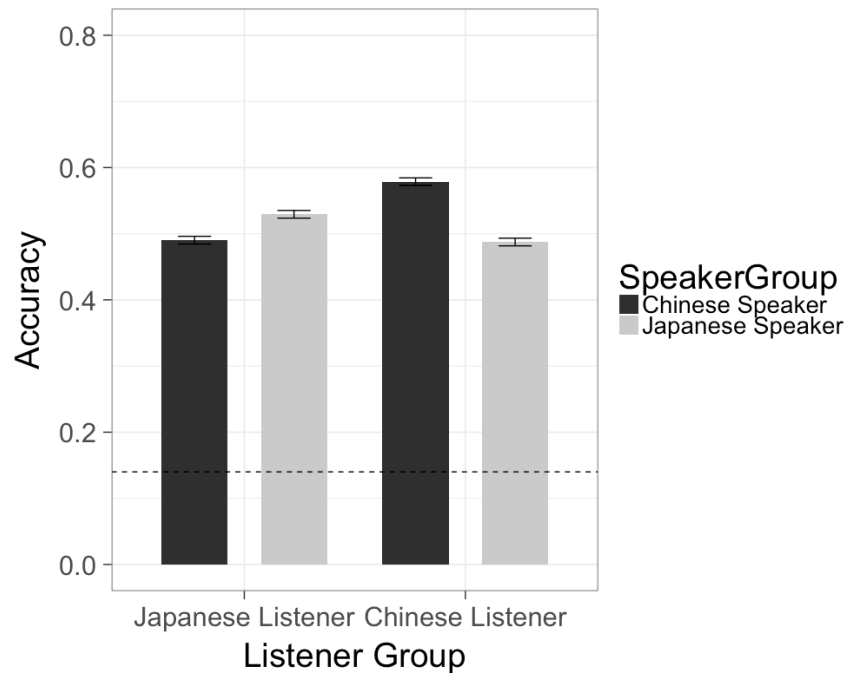


図 2-2 聴者間の同定率

聴者×発話者×感情の交互作用が有意であったため、発話者 2 水準（母語話者、学習者）×聴者 2 水準（母語話者、学習者）の 2 要因分散分析を行なった（図 2-2：図中の点線はチャンスレベルを示す， $1/7 = .14$ ）。その結果，発話者の主効果および発話者と聴者の交互作用が有意であった（ $F(1, 22) = 9.15, p = .006$; $F(1, 22) = 16.86, p < .001$ ）。また，図 2-2 の通り，日本語母語話者による発話の場合，日本語母語話者の同定率（52%）は中国人学習者（48%）より有意に高かった（ $F(1, 22) = 13.72, p = .001$ ）。一方，中国人学習者による発話の場合，中国人学習者の同定率（57%）は日本語母語話者（51%）より有意に高かった（ $F(1, 22) = 6.83, p = .015$ ）。

同定率は感情によって変わることも示された（図 2-3：Japanese Speaker：日本語母語話者発話者，Chinese Speaker：中国人学習者発話者，Japanese Listener：日本語母語話者聴者，Chinese Listener：中国人学習者聴者）。日本語母語話者による発話において，日本語母語話者の「驚き」は中国人学習者より有意に高く（ $F(1, 11) = 28.38, p < .001$ ），中国人学習者による発話においては，中国人学習者の「恐れ」「怒り」「中立」「悲しみ」「驚き」の同定率は日本語母語話者より有意に高かった（ $F(1, 11) = 15.37, p = .002$; $F(1, 11) = 4.90, p = .048$; $F(1, 11) = 82.56, p < .001$; $F(1, 11) = 36.88, p < .001$; $F(1, 11) = 5.52, p = .038$ ）。ここでは「嫌悪」

と「喜び」には聴者による差は観察されなかった ($F(1,11) = 2.85, p = .119$; $F(1,11) = 1.38, p = .264$)。

これらのことは、発話者と聴者の母語が一致している場合の感情同定率は、母語が一致していない場合の感情同定率より高いことを示しており、櫻庭他 (2001, 2004) の日米間の比較結果および Bent & Bradlow(2003)の結果と一致している。このような傾向は、Pell et al. (2009) と Chronaki et al. (2018) が提案している「内集団優位性 (in-group advantage)」を支持するものと考えられる。

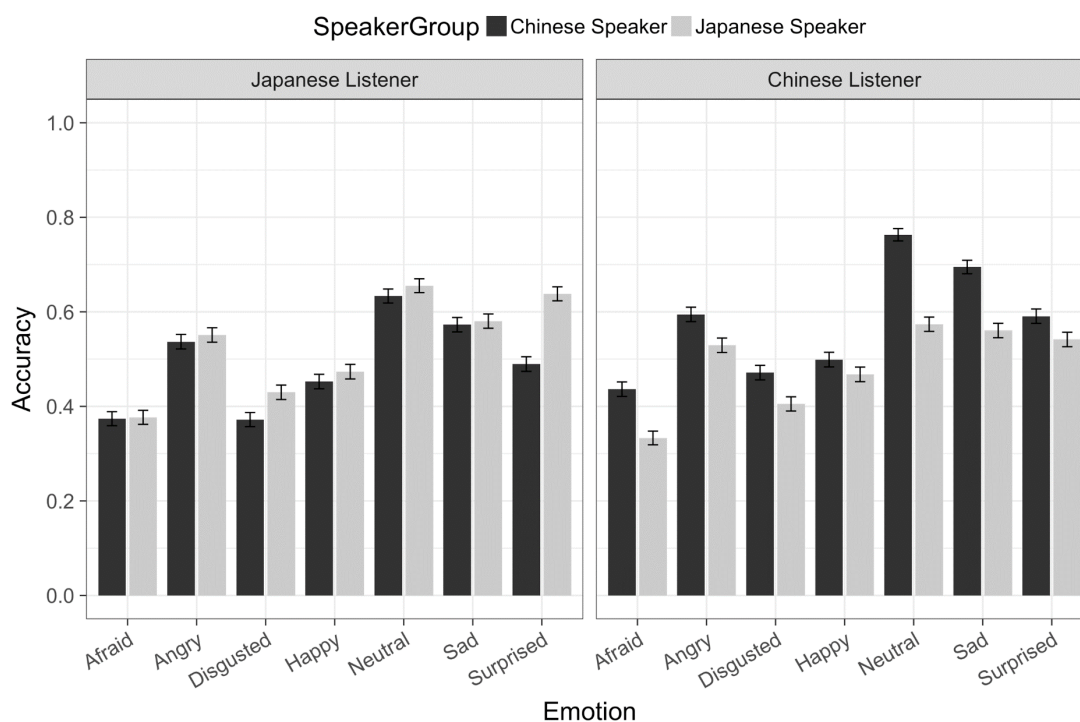


図 2-3 感情間の同定率

続いて、感情音声に対する聴者の各感情の混同傾向について検討する。まず、聴者がどの感情を選択したかを示す混同行列を表 2-1～表 2-4 に示す。縦欄は意図された感情であり、横欄は回答された感情を示す。次に、Levy (2009) にならい、公式 (3) を用いて、Overlap Score を計算することで、0 (全く重なっていない、つまり全く混同されない) から 1 (完全に重なっている、つまり完全に混同される) までの値で感情間の混同がどの程度生じやすいかについて数値化した。

$$\text{overlap}(X, Y) = \frac{|X \cap Y|}{\min(|X|, |Y|)} \quad (3)$$

さらに、昇地他（2007）にならい、聴者が各感情をどのように知覚したのかを調べるために、MDS を用いて 2 次元の知覚空間を描いた（図 2-4～図 2-7 参照）。表 2-1～表 2-2 および図 2-4～図 2-5 は日本語母語話者による感情音声を対象とした実験の結果と対応している。表 2-3～表 2-4 および図 2-6～図 2-7 は中国人学習者による感情音声を対象とした実験の結果と対応している。

表 2-1 日本語母語話者による発話に対する母語話者の感情判定 Confusion Matrix

同定率	Response						
	Afraid	Angry	Disgusted	Happy	Neutral	Sad	Surprised
Stimuli							
Afraid	0.38	0.05	0.05	0.03	0.10	0.27	0.13
Angry	0.00	0.55	0.13	0.05	0.16	0.03	0.08
Disgusted	0.01	0.12	0.43	0.01	0.11	0.16	0.16
Happy	0.02	0.12	0.01	0.47	0.14	0.02	0.23
Neutral	0.02	0.06	0.05	0.01	0.66	0.20	0.01
Sad	0.11	0.03	0.08	0.02	0.13	0.58	0.05
Surprised	0.02	0.10	0.04	0.17	0.04	0.00	0.64

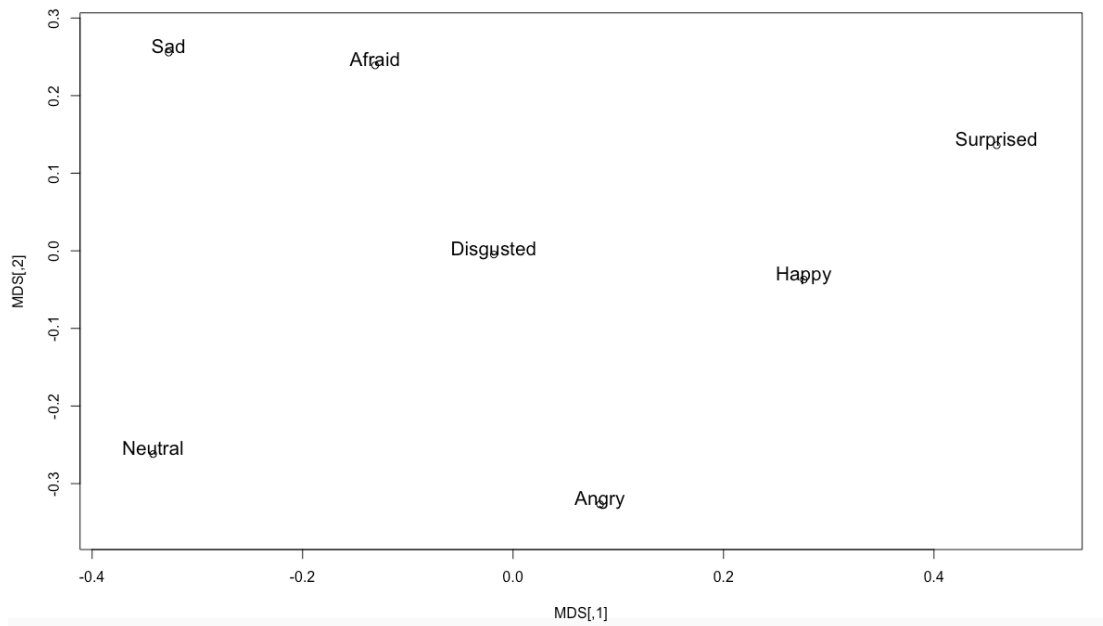


図 2-4 日本語母語話者による発話に対する母語話者の知覚行動

まず日本語母語話者が生成した感情音声について検討する。表 2-1 と図 2-4 に示す日本語母語話者の反応には、「中立」の同定率（66%）が最も高く、これは櫻庭他（2001, 2004）の結果と一致しており、「恐れ」の同定率（38%）が最も低かったことが見られた。また、MDS の結果について、図 2-4 の第 2 次元は「悲しみ」と「恐れ」、「喜び」「驚き」および「嫌悪」、「怒り」と「中立」を 3 つのグループとして分離する。さらに第 1 次元によって「喜び」「驚き」と「嫌悪」の配置が大きく変わることが観察された。

続いて表 2-2 と図 2-5 を見てみると、母語話者が生成した音声に対して、中国人学習者の同定率は「中立」の同定率（57%）が最も高く、「恐れ」の同定率（33%）が最も低いことが観察された。この結果は日本語母語話者聴者と一致しているが、それぞれの同定率は日本語母語話者より低いことがわかった。さらに、「喜び」以外の 6 種類の感情では、特に「中立」と「驚き」の同定率は、中国人学習者聴者の同定率は日本語母語話者聴者より低かった。これらのことは、日本語母語話者が生成した感情音声は中国人学習者に伝わりにくい可能性を示唆している。

また、図 2-5 の通り、第 2 次元は「怒り」「中立」とそれ以外の感情を分離している。さらに、「喜び」「驚き」および「嫌悪」の相対的距離が短いため、学習者が示した混同

傾向が母語話者より顕著であると考えられる。これは、中国人学習者は母語話者と異なる音声的手がかりを用いて感情を知覚している可能性を示唆している。

表 2-2 母語話者による発話に対する学習者の感情判定 confusion matrix

同定率	Response						
	Afraid	Angry	Disgusted	Happy	Neutral	Sad	Surprised
Stimuli							
Afraid	0.33	0.06	0.05	0.06	0.08	0.33	0.09
Angry	0.00	0.53	0.15	0.05	0.16	0.03	0.08
Disgusted	0.01	0.14	0.41	0.03	0.11	0.16	0.14
Happy	0.02	0.13	0.03	0.47	0.11	0.01	0.23
Neutral	0.01	0.07	0.04	0.02	0.57	0.25	0.02
Sad	0.12	0.04	0.07	0.04	0.15	0.56	0.02
Surprised	0.02	0.12	0.07	0.23	0.02	0.01	0.54

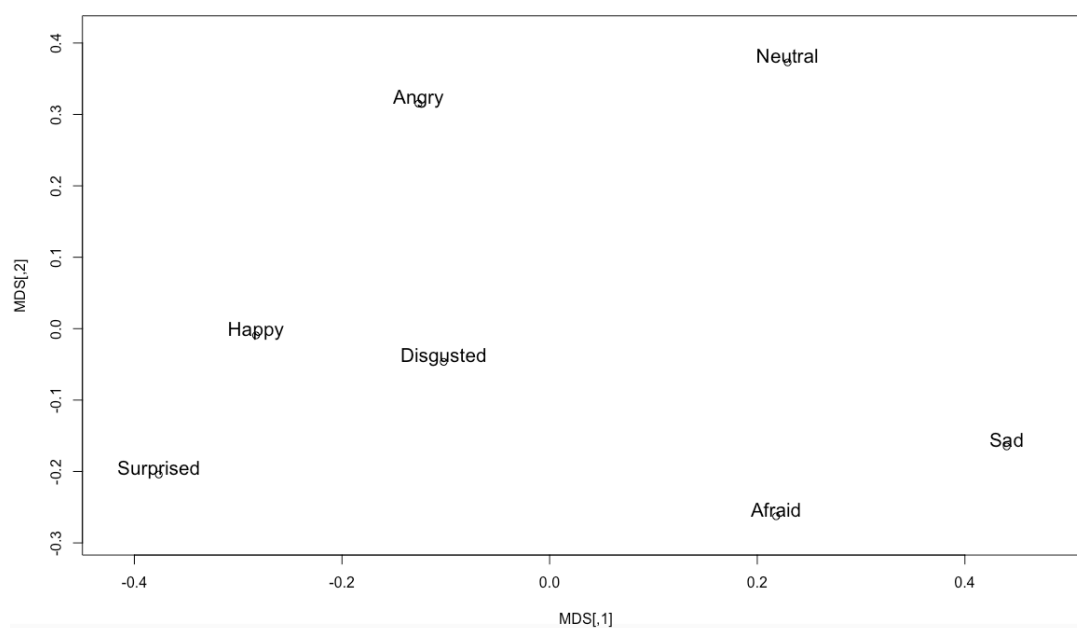


図 2-5 母語話者による発話に対する学習者の知覚行動

続いて中国人学習者による発話について検討する（表 2-3, 図 2-6）。まず、日本語母語話者の反応については、日本語母語話者聴者の知覚傾向は、日本語母語話者による発話の傾向と違い、「嫌悪」の同定率が最も低かった（37%）。さらに、日本語母語話者による発話における「驚き」（64%）と比較すると、中国人学習者による発話における「驚き」の同定率が顕著に低かった（49%）。これらのことは、中国人学習者による「嫌悪」と「驚き」は日本語母語話者に伝わりにくい可能性が示唆される。

また図 2-6 から、日本語母語話者による発話と比べると、中国人学習者による発話の「嫌悪」の知覚は「喜び」と「驚き」から大きく離れたことが観察された。なお、「喜び」と「驚き」の相対的距離がさらに小さくなり、より強い混同傾向を示した。この結果は、中国人学習者が感情音声を生成する際、日本語母語話者と違う音響的手がかりを用いて「嫌悪」を表現している可能性が考えられる。この点に関しては音響分析による検討が必要となる。

表 2-3 学習者による発話に対する母語話者の感情判定 confusion matrix

同定率	Response						
	Afraid	Angry	Disgusted	Happy	Neutral	Sad	Surprised
Stimuli							
Afraid	0.41	0.01	0.08	0.04	0.09	0.30	0.07
Angry	0.01	0.54	0.04	0.06	0.28	0.03	0.06
Disgusted	0.03	0.17	0.37	0.05	0.10	0.08	0.19
Happy	0.02	0.12	0.01	0.48	0.19	0.02	0.16
Neutral	0.01	0.05	0.05	0.03	0.68	0.17	0.00
Sad	0.10	0.01	0.05	0.01	0.19	0.63	0.01
Surprised	0.04	0.08	0.03	0.30	0.05	0.01	0.49

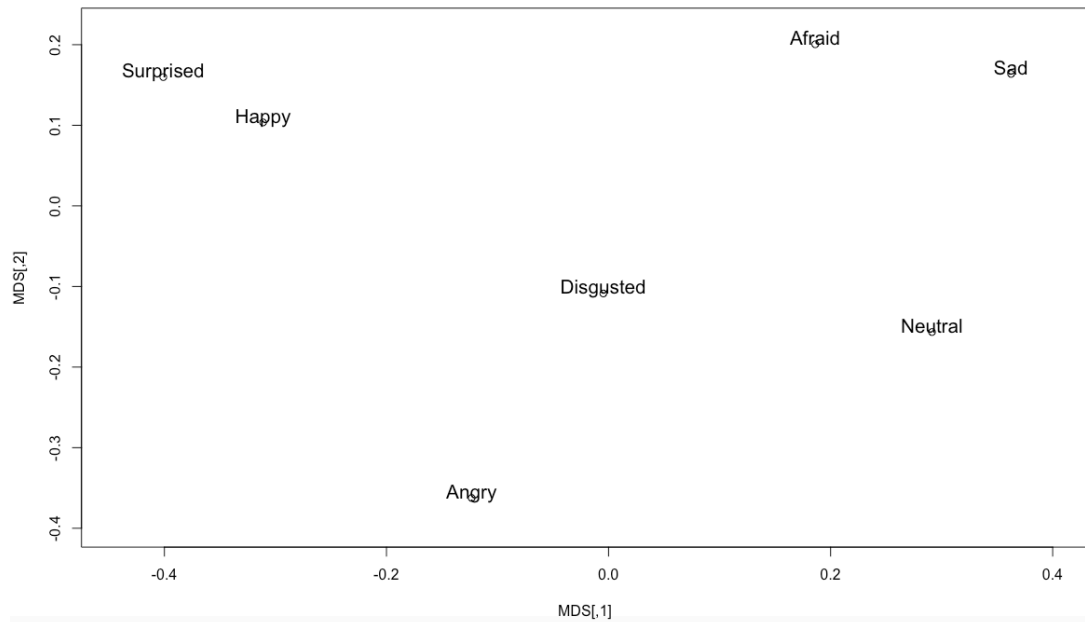


図 2-6 学習者による発話に対する母語話者の知覚行動

次に中国人学習者による発話に対して中国人学習者の知覚傾向を検討する。表 2-4 と図 2-7 に示す中国人学習者の反応を見るとわかるように、全ての感情において、日本語母語話者による発話と比較して、中国人学習者による発話の同定率が高かった。特に「悲しみ」において、中国人学習者が日本語母語話者による発話を 58%の確率で正しく判定できるのに対して、同じ中国人学習者による発話を判定する際、70%の確率で正確に判定していることが見られた。これらの結果は、中国人学習者聴者は日本語母語話者による発話より、同じ学習者による発話を正しく知覚できることが示していると考えられる。

また、図 2-7 を見ると、第 2 次元によって「怒り」は「中立」から離れ、「喜び」と「嫌悪」の相対的距離が小さくなり、「驚き」は「喜び」から離れたことが観察された。この結果は、日本語母語話者による発話と比較すると異なる点であり、中国人学習者聴者が日本語母語話者聴者と違う音響的手がかりを用いて感情音声を知覚していると考えられる。

表 2-4 学習者による発話に対する学習者の感情判定 confusion matrix

同定率	Response						
	Afraid	Angry	Disgusted	Happy	Neutral	Sad	Surprised
Stimuli							
Afraid	0.44	0.01	0.07	0.02	0.08	0.32	0.07
Angry	0.00	0.59	0.07	0.03	0.22	0.03	0.05
Disgusted	0.03	0.14	0.47	0.01	0.10	0.08	0.17
Happy	0.02	0.09	0.04	0.50	0.20	0.03	0.12
Neutral	0.00	0.03	0.05	0.01	0.76	0.14	0.01
Sad	0.09	0.01	0.04	0.00	0.15	0.70	0.02
Surprised	0.04	0.07	0.04	0.20	0.05	0.01	0.59

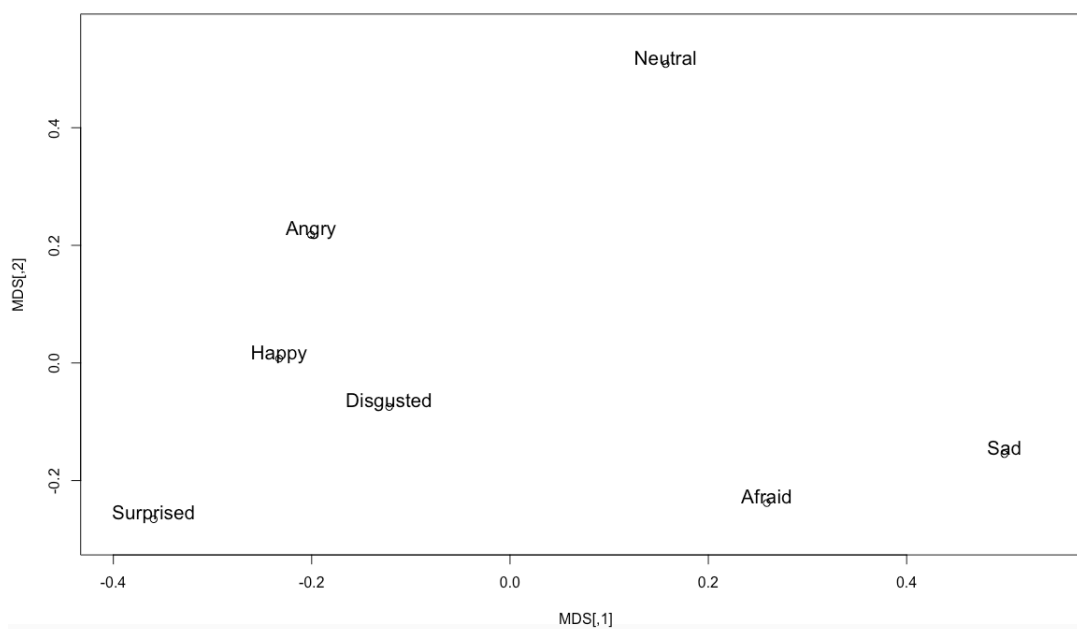


図 2-7 学習者による発話に対する学習者の知覚行動

続いて、各発話者グループおよび各聴者グループにあらわれる感情間の混同傾向を表 2-5 にまとめる。Overlap Score による結果について検討する。縦欄は 2 グループの発話者および聴者の組み合わせであり、横欄は各感情ペアの Overlap Score 結果を示す。

表 2-5 Overlap Score による感情間の混同傾向

<i>Emotion pair</i>	JS-JL	JS-CL	CS-JL	CS-CL	<i>Avg.</i>
<i>Sad + Afraid</i>	0.62	0.68	0.55	0.56	0.60
<i>Surprised + Happy</i>	0.56	0.65	0.52	0.63	0.59
<i>Disgusted + Angry</i>	0.58	0.53	0.40	0.46	0.49
<i>Sad + Neutral</i>	0.45	0.55	0.35	0.44	0.45
<i>Happy + Disgusted</i>	0.44	0.47	0.41	0.48	0.45
<i>Happy + Angry</i>	0.41	0.42	0.44	0.47	0.44
<i>Disgusted + Afraid</i>	0.40	0.49	0.35	0.40	0.41
<i>Surprised + Disgusted</i>	0.36	0.40	0.39	0.45	0.40
<i>Neutral + Afraid</i>	0.42	0.50	0.29	0.37	0.40
<i>Sad + Disgusted</i>	0.45	0.45	0.27	0.30	0.37
<i>Neutral + Disgusted</i>	0.41	0.44	0.27	0.34	0.36
<i>Neutral + Angry</i>	0.33	0.35	0.34	0.43	0.36
<i>Happy + Afraid</i>	0.35	0.35	0.27	0.26	0.31
<i>Surprised + Angry</i>	0.31	0.35	0.26	0.31	0.31
<i>Sad + Angry</i>	0.34	0.35	0.25	0.28	0.31
<i>Angry + Afraid</i>	0.33	0.36	0.26	0.27	0.30
<i>Neutral + Happy</i>	0.27	0.28	0.32	0.33	0.30
<i>Surprised + Afraid</i>	0.30	0.30	0.23	0.25	0.27
<i>Sad + Happy</i>	0.27	0.27	0.27	0.27	0.27
<i>Surprised + Neutral</i>	0.18	0.20	0.15	0.20	0.18
<i>Surprised + Sad</i>	0.19	0.21	0.16	0.16	0.18
<i>Overall mean:</i>	0.38	0.41	0.32	0.36	

(JS=Japanese speaker, CS=Chinese speaker, JL=Japanese listener, CL=Chinese listener)

表 2-5 の通り、感情間の混同傾向は、発話者と聴者に関係なく最も混同されやすい感情のペアは「悲しみ」と「恐れ」（平均 0.60）であり、最もはっきり区別されている感情ペアは「驚き」と「悲しみ」（0.18）および「驚き」と「中立」（0.18）である。また、結果の分布を考慮したうえ、日本語母語話者による発話に対して、中国人学習者による感情判

定の混同（平均 0.41）は日本語母語話者（平均 0.38）より顕著であった。さらに、日本語母語話者による発話において、日本語母語話者聴者には「悲しみ」と「恐れ」（0.62）、「驚き」と「喜び」（0.56）、「嫌悪」と「怒り」（0.58）および「悲しみ」と「中立」（0.62）の 3 ペアの混同が見られ、中国人学習者聴者には上記の 3 ペア以外、「悲しみ」と「中立」（0.55）の混同も観察された。一方で、中国人学習者による発話に対して、中国人学習者による感情間の混同（平均 0.36）は日本語母語話者（平均 0.32）より顕著であることが観察された。さらに、中国人学習者による発話において、聴者に関係なく「悲しみ」と「恐れ」および「驚き」と「喜び」の 2 ペアの混同が見られた。

Erickson et al. (2003) で、アメリカ英語母語話者、日本語母語話者と韓国人母語話者を対象に行なった知覚実験の結果によれば、日本語母語話者が発話した「怒り」の知覚において、アメリカ英語母語話者と韓国語母語話者の同定率が日本語母語話者より高くなった。その原因は日本語母語話者による発話は曖昧で混同されやすいためであると述べられている。本実験では「怒り」「嫌悪」などの感情に関して同様の傾向を観察することができた。

これらのことは、日本語母語話者による感情音声は中国人学習者による感情音声より混同されやすく、曖昧であることを示唆する。例えば「嫌悪」と「怒り」および「悲しみ」と「中立」の 2 ペアが中国人学習者による発話には見られなかった。なお、「悲しみ」と「恐れ」および「驚き」と「喜び」の 2 ペアの感情は発話者と聴者に関係なく混同が生じやすいことが示された。

第 2 章では、日本語母語話者および中国人学習者による感情音声を収録し、発話者とは異なる日本語母語話者および中国人学習者を対象とした知覚実験を行なった。その結果、発話者と聴者の母語が一致している場合の感情同定率は、母語が一致していない場合の感情同定率より高いことが観察された。さらに、日本語母語話者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」および「悲しみ」と「中立」の 4 ペアの混同が顕著であった。一方、中国人学習者による感情音声では、「悲しみ」と「恐れ」、「喜び」と「驚き」の 2 ペアが混同されやすいという結果を得た。日本語母語話者による「怒り」に関しては、Erickson et al. (2003) と同様の傾向を示した。これらの結果は、「内集団優位性 (in-group advantage)」と一致しており、日本語母語話者の発話の意図する感情は中国人学習者の発話より混同が多いことが示したものであると考えられる。

第3章 感情音声の生成

3.1. 実験の目的

知覚実験では、発話者と聴者の母語が一致している場合の感情同定率は、母語が一致していない場合の感情同定率より高いことと、日本語母語話者の発話の意図する感情は中国人学習者の発話より混同が多いことが明らかになった。

そこで本実験では以下を目的とする。なぜ知覚実験に上記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、中国語を母語とする日本語学習者が日本語で生成した感情音声と日本語母語話者が生成した感情音声に現れる音響的特徴(F0に関するパラメータ、Duration および Intensity)の違いは何かについて音声学的分析を通して中国人日本語学習者に難しいといわれる感情表現の使用実態を明らかにすることを目的とする。

3.2. 音響分析

3.2.1. 方法

音響分析に用いた音声資料は、知覚実験で収録したものをを用いた。音声に込める感情は Ekman (1992) の基本6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。日本語母語話者8名、中国人学習者8名によるそれぞれの感情を込めた日本語音声計1,232トークン(16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、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), 基本周波数F0の平均値(F0Mean), 持続時間長(Duration), インテンシティの中央値(MedianIntensity)の6つのパラメータについて測定を行なった。F0Level, F0Max, F0Span F0Mean, および MedianIntensity については、Hz から Z 得点へ変換した。Duration については、タスク語の母音と子音環境が統一されていないため、タスク語別で検討する。

3.2.2. 結果と考察

日本語母語話者および中国人学習者による発話の各音響的特徴の分析結果について、F0に関する音響的特徴、Duration、Intensityという順で検討していく。F0Levelを図3-1に示し、F0Maxを図3-2に示し、F0Spanを図3-3に示し、F0Meanを図3-4に示し、タスク語別のDurationを図3-5~図3-15に示し、MedianIntensityを図3-16に示す。また、各パラメータの感情間の違いを表3-1に示す。

3.2.2.1. F0に関する音響的特徴量の検討

図3-1の通り、F0Levelについて、発話者2水準（日本語母語話者、中国人学習者）×感情7水準（中立、喜び、怒り、悲しみ、驚き、恐れ、嫌悪）の2要因分散分析を行なったところ、発話者および感情の主効果は有意であった（ $F(1, 14) = 18.44, p < .001$; $F(6, 84) = 15.85, p < .001$ ）。発話者と感情の交互作用（ $F(6, 84) = 2.71, p = .018$ ）が有意であったため、多重比較を行なった。その結果、中国人学習者の「怒り」のF0Levelは日本語母語話者より有意に高く（ $F(1,14)=9.03, p=.009$ ）, 「嫌悪」は日本語母語話者より有意に低かった（ $F(1,14)=19.12, p<.001$ ）。また、F0Levelに関する感情間の違いを表3-1にまとめる。

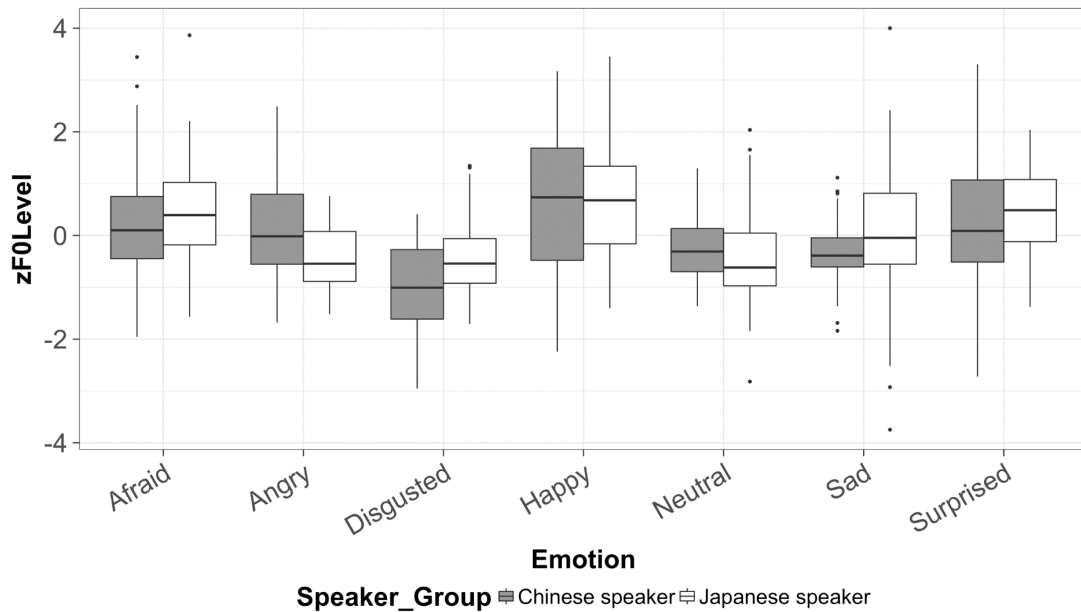


図 3-1 発話者グループ間および感情間の F0Level

図 3-2 の通り、F0Max について、発話者 2 水準（日本語母語話者、中国人学習者）×感情 7 水準（中立、喜び、怒り、悲しみ、驚き、恐れ、嫌悪）の 2 要因分散分析を行なったところ、感情の主効果は有意であり ($F(1, 14) = 42.25, p < .001$)，発話者と感情の交互作用 ($F(6, 84) = 5.13, p = .038$) は有意であった。発話者の主効果は有意ではなかった ($F(1, 14) = .11, p = .074$)。多重比較を行なった結果、中国人学習者の「悲しみ」の F0Max は日本語母語話者より有意に低く ($F(1, 14) = 11.08, p = .005$)，「中立」は日本語母語話者より有意に高かった ($F(1, 14) = 6.66, p < .021$)。また、F0Max に関する感情間の違いを表 3-1 にまとめる。

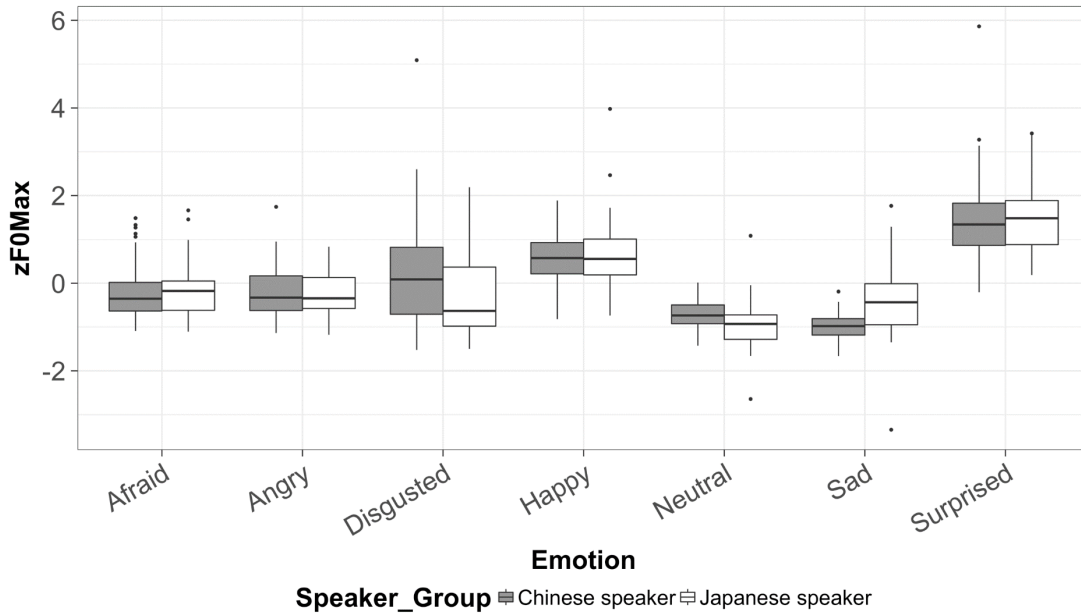


図 3-2 発話者グループ間および感情間の F0Max

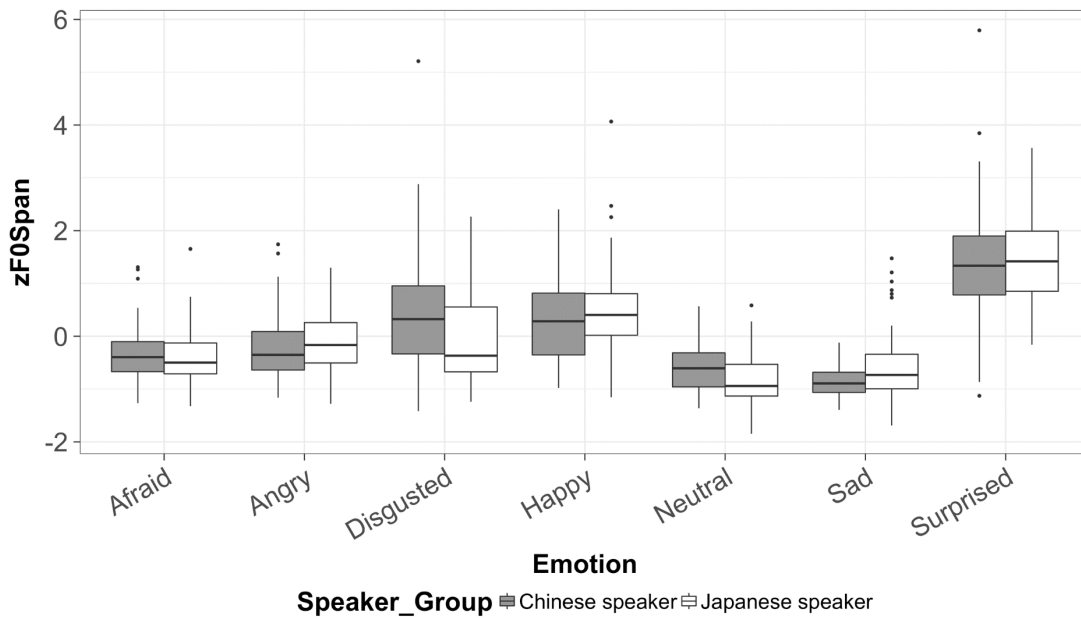


図 3-3 発話者グループ間および感情間の F0Span

図 3-3 の通り，F0Span について，発話者 2 水準（日本語母語話者，中国人学習者）×感情 7 水準（中立，喜び，怒り，悲しみ，驚き，恐れ，嫌悪）の 2 要因分散分析を行なったところ，感情の主効果は有意であり ($F(6, 84) = 73.35, p < .001$)，発話者と感情の交互作用 ($F(6, 84) = 6.56, p < .001$) は有意であった。発話者の主効果は有意ではなかった ($F(1,$

14) = .63, $p = .43$)。多重比較を行なった結果、中国人学習者の「嫌悪」の F0Span は日本語母語話者より有意に広く ($F(1,14) = 11.39, p = .004$) , 「悲しみ」は日本語母語話者より有意に狭かった ($F(1,14) = 28.94, p < .001$)。また、F0Span に関する感情間の違いを表 3-1 にまとめる。

図 3-4 の通り、F0Mean について、発話者 2 水準 (日本語母語話者, 中国人学習者) × 感情 7 水準 (中立, 喜び, 怒り, 悲しみ, 驚き, 恐れ, 嫌悪) の 2 要因分散分析を行なったところ、発話者および感情の主効果が有意であり ($F(1, 14) = 45.99, p < .001; F(6, 84) = 39.21, p < .001$) , 発話者と感情の交互作用 ($F(6, 84) = 4.90, p = .041$) が有意であった。多重比較を行なった結果、中国人学習者の「悲しみ」の F0Mean は日本語母語話者より有意に低かった ($F(1,14) = 8.53, p = .011$)。また、F0Mean に関する感情間の違いを表 3-1 にまとめる。

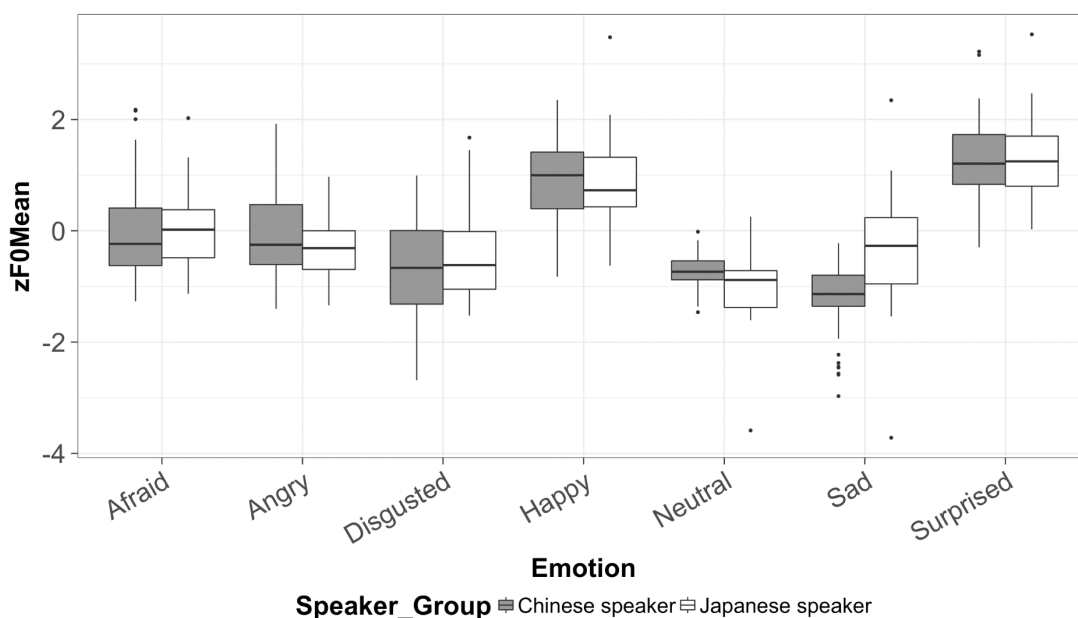


図 3-4 発話者グループ間および感情間の F0Mean

F0 に関する検討から見ると、日本語母語話者と中国人学習者の発話の間では、F0Level, F0Max, F0Span および F0Mean の 4 つのパラメータに有意差が観察された。まとめると、

中国人学習者による発話の「怒り」は日本語母語話者による発話より F0Level が高く、「悲しみ」は日本語母語話者より F0Max が低く、F0Span が狭く、F0Mean が低く、「嫌悪」は日本語母語話者より F0Level が低く、F0Span が広く、「中立」は日本語母語話者より高かった。

これらのことは、日本語母語話者と中国人学習者が F0 に関するパラメータを用いて感情音声を生成する際、「怒り」「悲しみ」「嫌悪」および「中立」に関して異なった F0 を用いる可能性が示唆された。

さらに知覚実験においては、日本語母語話者では「悲しみ」と「恐れ」および「喜び」と「驚き」の他に、「怒り」と「嫌悪」および「悲しみ」と「中立」の 2 組で混同が顕著に認められたが、中国人学習者ではそれはあまり見られなかった。表 1 の通り、中国人学習者による発話では、「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、中国人学習者が感情音声を生成する際、「喜び」と「驚き」を区別する音響学的手がかりは 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、日本語母語話者が生成した感情音声では、これらの 2 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、聴者が日本語母語話者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。さらに、感情間の F0 の違いについて、日本語母語話者による発話では 9 組の違いが見られ、中国人学習者による発話では 17 組の違いが観察された。以上の結果から、日本語母語話者の音声混同されやすかった一つの原因は、日本語母語話者の感情音声間の F0 に関するパラメータの違いが中国人学習者に比べて顕著ではなかったためだと考えられる。

表 3-1 F0 に関する音響的特徴の感情間の違い

日本語母語話者	zF0Mean	zF0Level	zF0Max	zF0Span
怒り vs. 驚き	<*	<*	<i>n.s.</i>	<*
驚き vs. 嫌悪	>*	>*	<i>n.s.</i>	<i>n.s.</i>
喜び vs. 嫌悪	<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>
中立 vs. 驚き	<*	<i>n.s.</i>	<*	<*
悲しみ vs. 驚き	<*	<i>n.s.</i>	<i>n.s.</i>	<*
恐れ vs. 驚き	<i>n.s.</i>	<i>n.s.</i>	<*	<*
中立 vs. 喜び	<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>	>*
中国人学習者	zF0Mean	zF0Level	zF0Max	zF0Span
中立 vs. 嫌悪	<i>n.s.</i>	>*	<i>n.s.</i>	<i>n.s.</i>
驚き vs. 嫌悪	>*	>*	<i>n.s.</i>	<i>n.s.</i>
悲しみ vs. 嫌悪	<i>n.s.</i>	>*	<*	<*
喜び vs. 嫌悪	>*	>*	<i>n.s.</i>	<i>n.s.</i>
喜び vs. 驚き	<i>n.s.</i>	>*	<i>n.s.</i>	<*
恐れ vs. 嫌悪	<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>
中立 vs. 驚き	<*	<i>n.s.</i>	<*	<*
悲しみ vs. 驚き	<*	<i>n.s.</i>	<*	<*
中立 vs. 喜び	<*	<i>n.s.</i>	<i>n.s.</i>	<i>n.s.</i>
怒り vs. 驚き	<*	<i>n.s.</i>	<*	<*
怒り vs. 悲しみ	>*	<i>n.s.</i>	>*	>*
喜び vs. 悲しみ	<i>n.s.</i>	<i>n.s.</i>	>*	>*
恐れ vs. 驚き	<i>n.s.</i>	<i>n.s.</i>	<i>n.s.</i>	<*
悲しみ vs. 恐れ	<i>n.s.</i>	<i>n.s.</i>	<*	<*
中立 vs. 怒り	<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>

* $p < .05$

3.2.2.2. Duration に関する検討

続いて、Duration について、各タスク語の母音と子音環境が統一されていないため、タスク語別で検討する。「e」を図 3-5 に示し、「me」を図 3-6、「umi」を図 3-7、「momo」を図 3-8、「banana」を図 3-9、「nioi」を図 3-10、「manami」を図 3-11、「imanimo」を図 3-12、「namamono」を図 3-13、「menomae」を図 3-14、「aoyama」を図 3-15 に示す。

図 3-5 の通り、タスク語「e」の Duration について、発話者 2 水準（日本語母語話者、中国人学習者）×感情 7 水準（中立、喜び、怒り、悲しみ、驚き、恐れ、嫌悪）の 2 要因分散分析を行なった。その結果、感情の主効果が有意であり ($F(6, 84) = 14.35, p < .001$)、発話者と感情の交互作用が有意であった ($F(6, 84) = 6.16, p < .001$)。発話者の主効果は有意ではなかった ($F(1, 14) = 1.74, p = .207$)。多重比較を行なった結果、中国人学習者による「e」の「喜び」「怒り」発話の Duration は日本語母語話者より有意に短く ($F(1, 14) = 14.54, p = .002; F(1, 14) = 50.88, p < .001$)、 「恐れ」発話の Duration は日本語母語話者より有意に長かった ($F(1, 14) = 29.67, p < .001$)。また、「e」の Duration に関する感情間の違いは、日本語母語話者による発話において、「嫌悪」は「驚き」より有意に長く ($t(7) = 6.58, p < .001$)。一方、中国人学習者による発話において、「恐れ」は「喜び」「中立」「怒り」「驚き」より有意に高く ($t(7) = 12.21, p < .001; t(7) = 10.60, p < .001; t(7) = 10.15, p < .001; t(7) = 6.53, p < .001$)、 「嫌悪」は「怒り」「中立」「驚き」より有意に長く ($t(7) = 6.97, p < .001; t(7) = 5.50, p < .001; t(7) = 5.38, p = .001$)、 「怒り」は「中立」より有意に短かった ($t(7) = 4.28, p = .003$)。

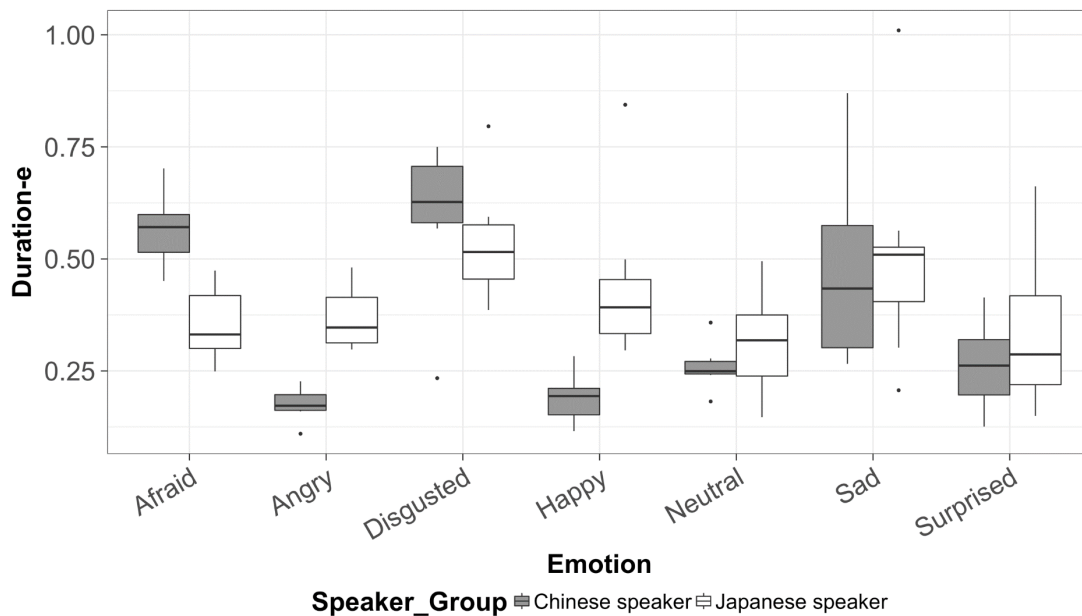


図 3-5 発話者グループ間および感情間の Duration-e

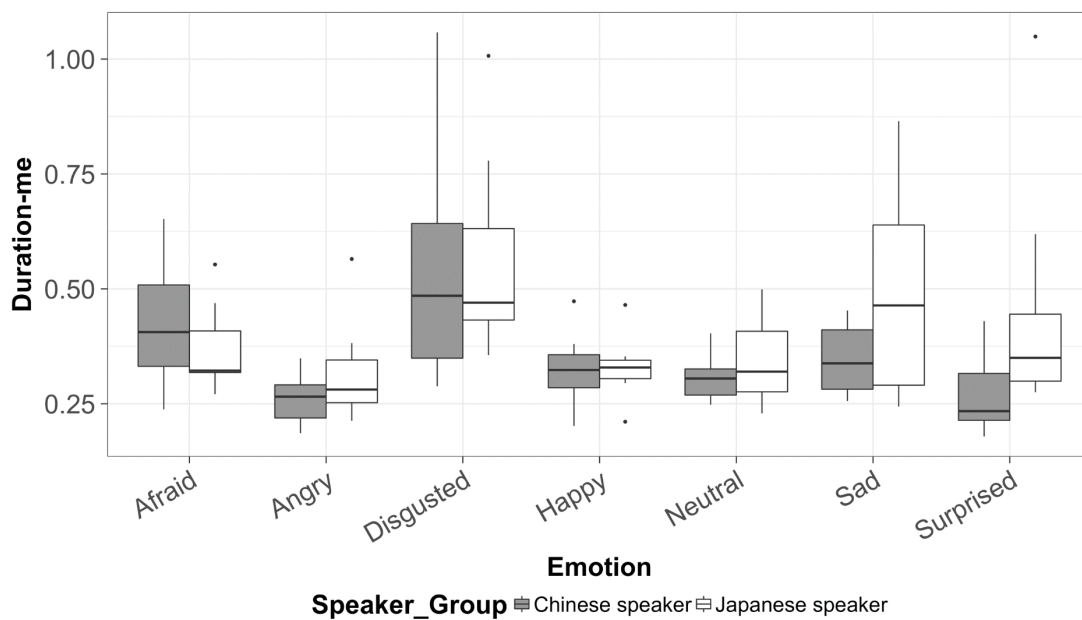


図 3-6 発話者グループ間および感情間の Duration-me

図 3-6 に示したタスク語「me」の Duration について、発話者 2 水準（日本語母語話者、中国人学習者）×感情 7 水準（中立、喜び、怒り、悲しみ、驚き、恐れ、嫌悪）の 2 要因分散分析を行なった結果、感情の主効果が有意であった ($F(6, 84)=7.96.01, p<.001$)、発話者の主効果が有意ではなかった ($F(1, 14)=4.47, p=.279$)、感情と発話者の交互作用が有意ではなかった ($F(6, 84)=4.47, p=.127$)。

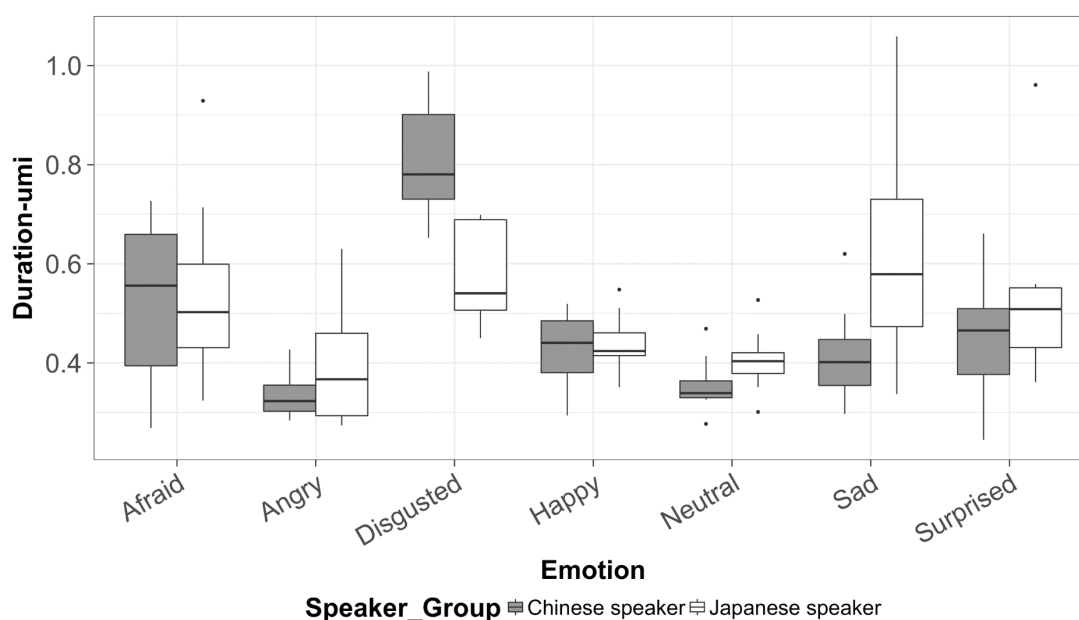


図 3-7 発話者グループ間および感情間の Duration-umi

図 3-7 の通り、タスク語「umi」の Duration について、発話者 2 水準（日本語母語話者、中国人学習者）×感情 7 水準（中立、喜び、怒り、悲しみ、驚き、恐れ、嫌悪）の 2 要因分散分析を行なった。その結果、感情の主効果が有意であり ($F(6, 84)=10.01, p<.001$)、発話者と感情の交互作用が有意であった ($F(6, 84)=4.47, p<.001$)。発話者の主効果が有意ではなかった ($F(1, 14)=.782, p=.391$)。多重比較を行なった結果、中国人学習者による「umi」の「嫌悪」発話の Duration は日本語母語話者より有意に長かった ($F(1,14)=18.44, p<.001$)。また、「umi」の Duration に関する感情間の違いは、日本語母語話者による発話において、「嫌悪」は「中立」より有意に長かった ($t(7)=4.99, p=.002$)。一方、中国人学習者による発話において、中国人学習者の「嫌悪」は「怒り」「中立」「悲しみ」「驚

き」「恐れ」より有意に長かった ($t(7) = 13.36, p < .001$; $t(7) = 10.13, p < .001$; $t(7) = 9.95, p < .001$; $t(7) = 7.72, p < .001$; $t(7) = 5.55, p = .001$)。

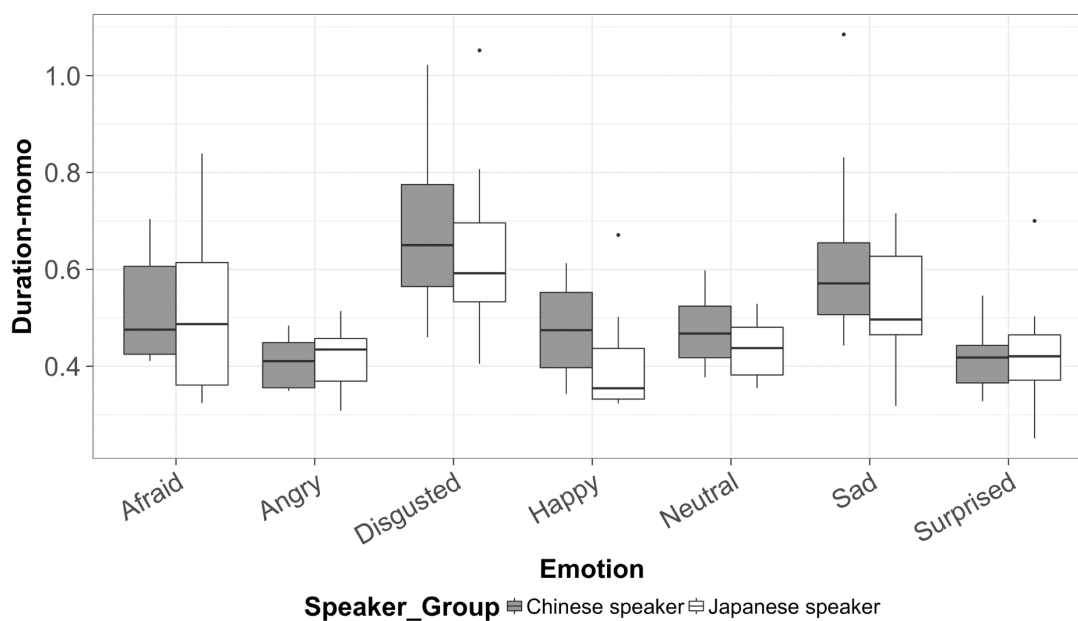


図 3-8 発話者グループ間および感情間の Duration-momo

図 3-8 に示すタスク語「momo」の Duration について、発話者 2 水準（日本語母語話者、中国人学習者）×感情 7 水準（中立、喜び、怒り、悲しみ、驚き、恐れ、嫌悪）の 2 要因分散分析を行なった結果、感情の主効果が有意であった ($F(6, 84) = 6.56, p < .001$)，発話者の主効果が有意ではなく ($F(1, 14) = 1.13, p = .307$)，感情と発話者の交互作用が有意ではなかった ($F(6, 84) = .796, p = .575$)。

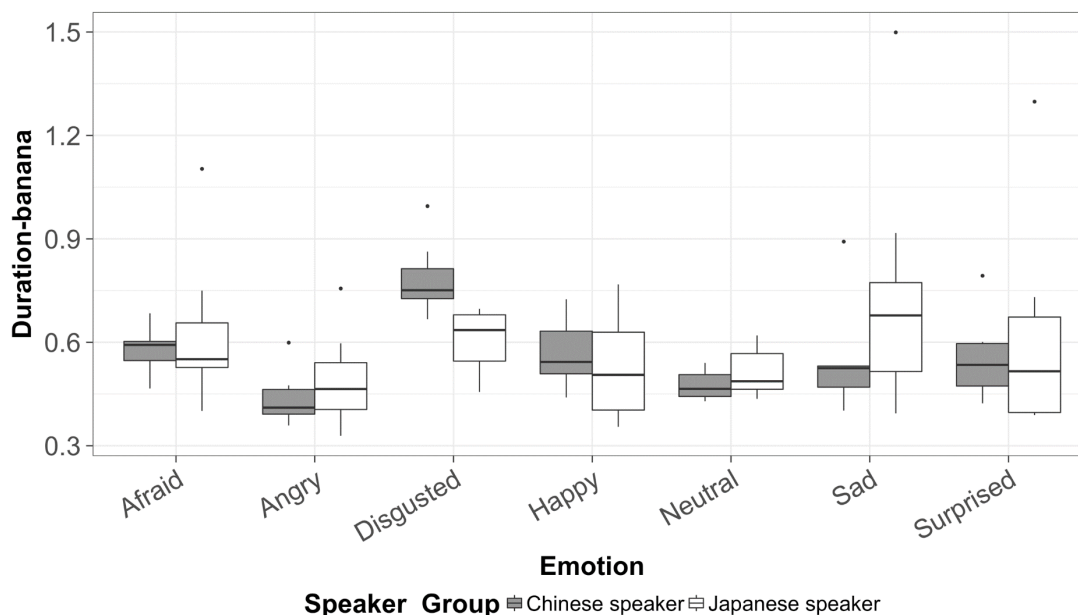


図 3-9 発話者グループ間および感情間の Duration-banana

図 3-9 の通り，タスク語「banana」の Duration について，発話者 2 水準（日本語母語話者，中国人学習者）×感情 7 水準（中立，喜び，怒り，悲しみ，驚き，恐れ，嫌悪）の 2 要因分散分析を行なった。その結果，感情の主効果が有意であり ($F(6, 84) = 7.05, p < .001$)，発話者と感情の交互作用が有意であった ($F(6, 84) = 3.44, p = .004$)。発話者の主効果は有意ではなかった ($F(1, 14) = .175, p = .681$)。多重比較では，中国人学習者による「banana」の「嫌悪」発話の Duration は日本語母語話者より有意に長かった ($F(1, 14) = 12.70, p = .003$)。また，「banana」の Duration に関する感情間の違いは，日本語母語話者による発話において，感情間の違いが観察されなかった。一方，中国人学習者による発話において，中国人学習者の「嫌悪」は「怒り」「中立」「恐れ」「喜び」より有意に長く ($t(7) = 7.39, p < .001$; $t(7) = 7.01, p < .001$; $t(7) = 5.82, p = .001$; $t(7) = 4.74, p = .002$)，「恐れ」は「怒り」より有意に長かった ($t(7) = 4.82, p = .002$)。

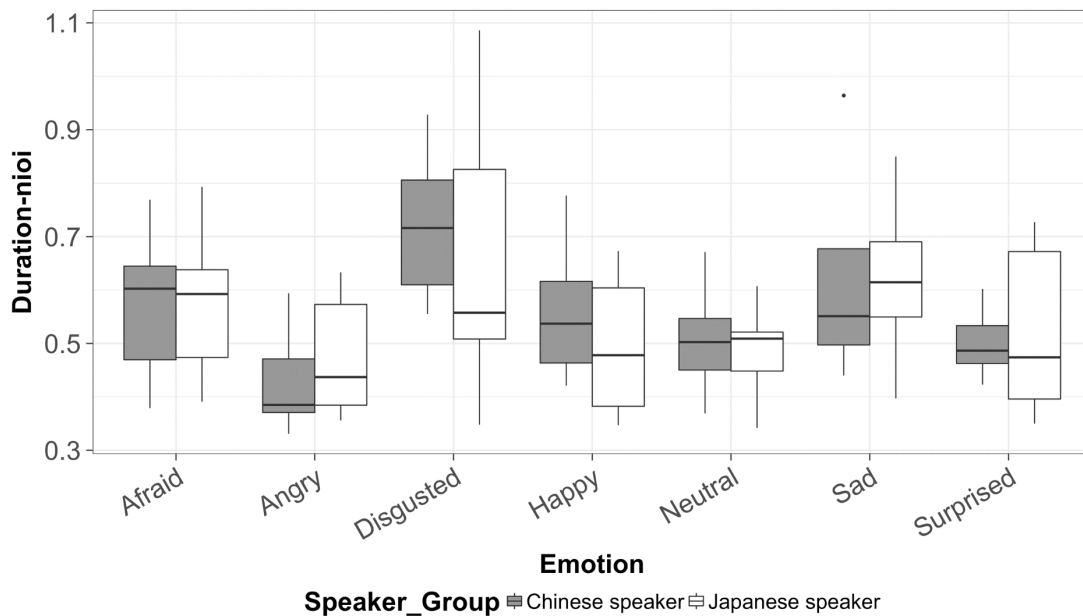


図 3-10 発話者グループ間および感情間の Duration-nioi

図 3-10 に示したタスク語「nioi」の Duration について、発話者 2 水準（日本語母語話者，中国人学習者）×感情 7 水準（中立，喜び，怒り，悲しみ，驚き，恐れ，嫌悪）の 2 要因分散分析を行なった結果，感情の主効果が有意であった ($F(6, 84) = 7.30, p < .001$)，発話者の主効果が有意ではなく ($F(1, 14) = .07, p = .797$)，感情と発話者の交互作用が有意ではなかった ($F(6, 84) = .45, p = .846$)。

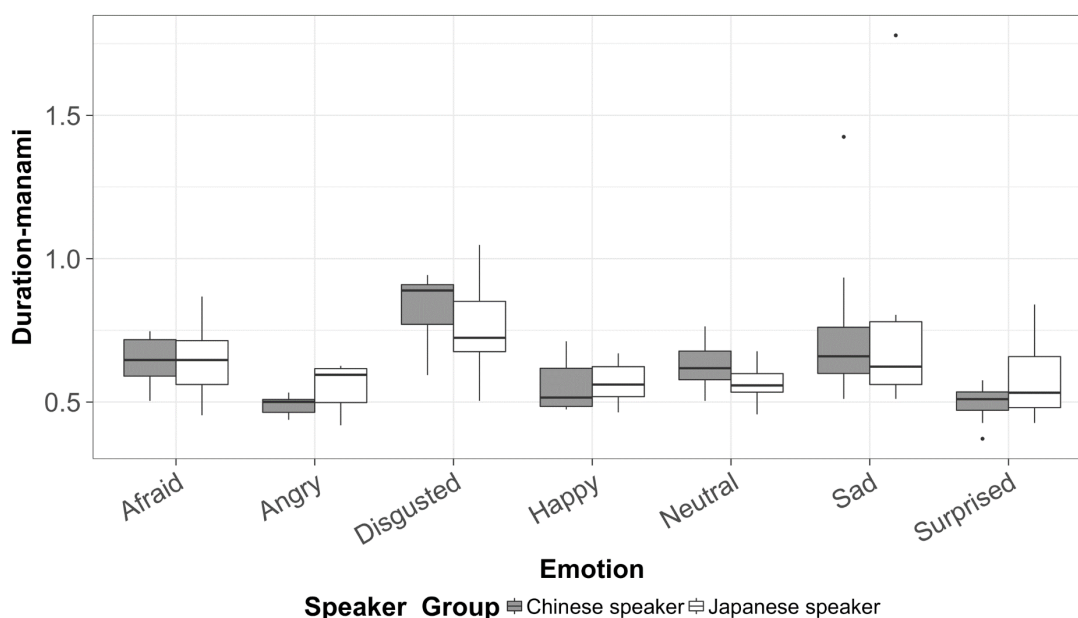


図 3-11 発話者グループ間および感情間の Duration-manami

図 3-11 にはタスク語「manami」の Duration について示す。発話者 2 水準（日本語母語話者，中国人学習者）×感情 7 水準（中立，喜び，怒り，悲しみ，驚き，恐れ，嫌悪）の 2 要因分散分析を行なったところ，感情の主効果が有意であり ($F(6, 84) = 10.32, p < .001$)，発話者と感情の交互作用が有意であった ($F(6, 84) = 2.60, p = .023$)。発話者の主効果が有意ではなかった ($F(1, 14) = .314, p = .584$)。多重比較を行なった結果，中国人学習者による「banana」の「怒り」発話の Duration は日本語母語話者より有意に短く ($F(1, 14) = 37.36, p < .001$)，「嫌悪」は日本語母語話者より有意に長く ($F(1, 14) = 6.20, p = .026$)，「驚き」は日本語母語話者より有意に短かった ($F(1, 14) = 20.54, p < .001$)。また，「manami」の Duration に関する感情間の違いは，日本語母語話者による発話において，「嫌悪」は「喜び」より有意に長かった ($t(7) = 4.63, p = .002$)。一方，中国人学習者による発話において，中国人学習者の「嫌悪」は「驚き」「喜び」「怒り」「中立」「恐れ」より有意に長く ($t(7) = 15.21, p < .001; t(7) = 12.83, p < .001; t(7) = 11.44, p < .001; t(7) = 8.25, p = .001; t(7) = 4.52, p = .002$)，「恐れ」は「驚き」「怒り」より有意に長く ($t(7) = 7.20, p < .001; t(7) = 6.64, p < .001$)，「怒り」は「中立」より有意に短かった ($t(7) = 5.34, p = .001$)。

図 3-12 の通り，タスク語「imanimo」の Duration について，発話者 2 水準（日本語母語話者，中国人学習者）×感情 7 水準（中立，喜び，怒り，悲しみ，驚き，恐れ，嫌悪）の 2 要因分散分析を行なった結果，感情の主効果が有意であった ($F(6, 84) = 2.93, p = .012$)，

発話者の主効果が有意ではなかった ($F(1, 14) = .21, p = .657$) , 感情と発話者の交互作用が有意ではなかった ($F(6, 84) = .54, p = .775$) 。

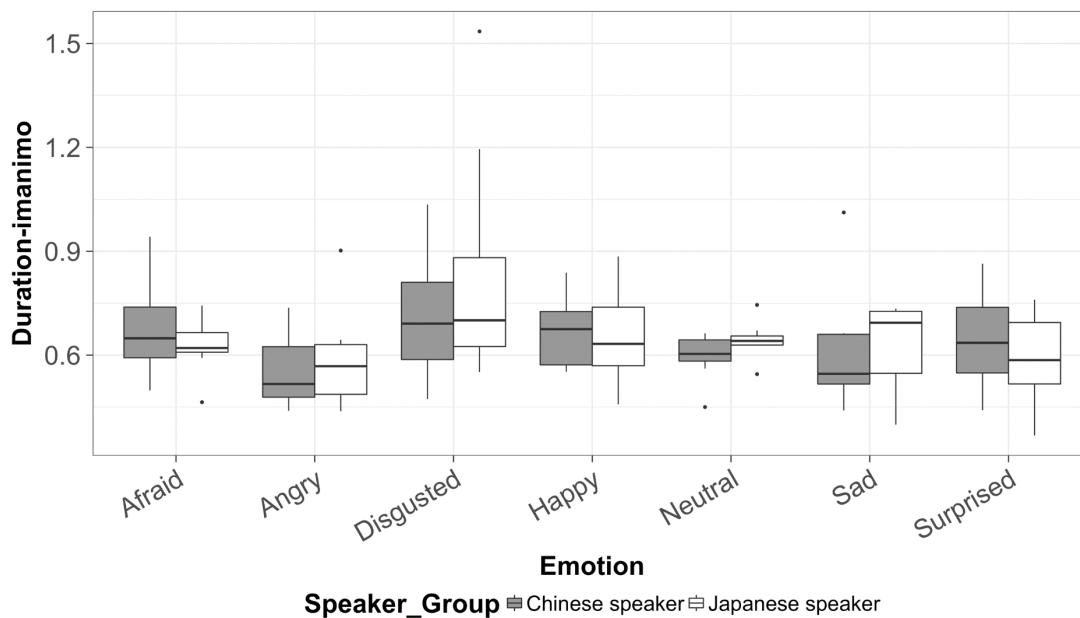


図 3-12 発話者グループ間および感情間の Duration-imanimo

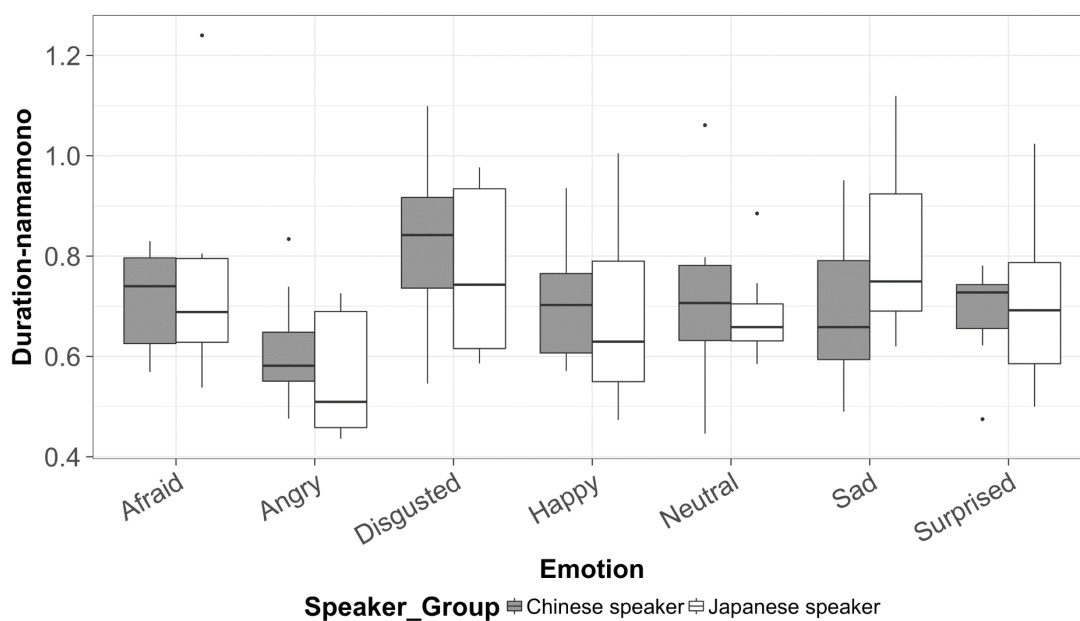


図 3-13 発話者グループ間および感情間の Duration-namamono

図 3-13 の通り，タスク語「namamono」の Duration について，発話者 2 水準（日本語母語話者，中国人学習者）×感情 7 水準（中立，喜び，怒り，悲しみ，驚き，恐れ，嫌悪）の 2 要因分散分析を行なった結果，感情の主効果が有意であった ($F(6, 84) = 3.59, p = .003$)，発話者の主効果は有意ではなかった ($F(1, 14) = .03, p = .872$)，感情と発話者の交互作用は有意ではなかった ($F(6, 84) = 1.04, p = .402$)。

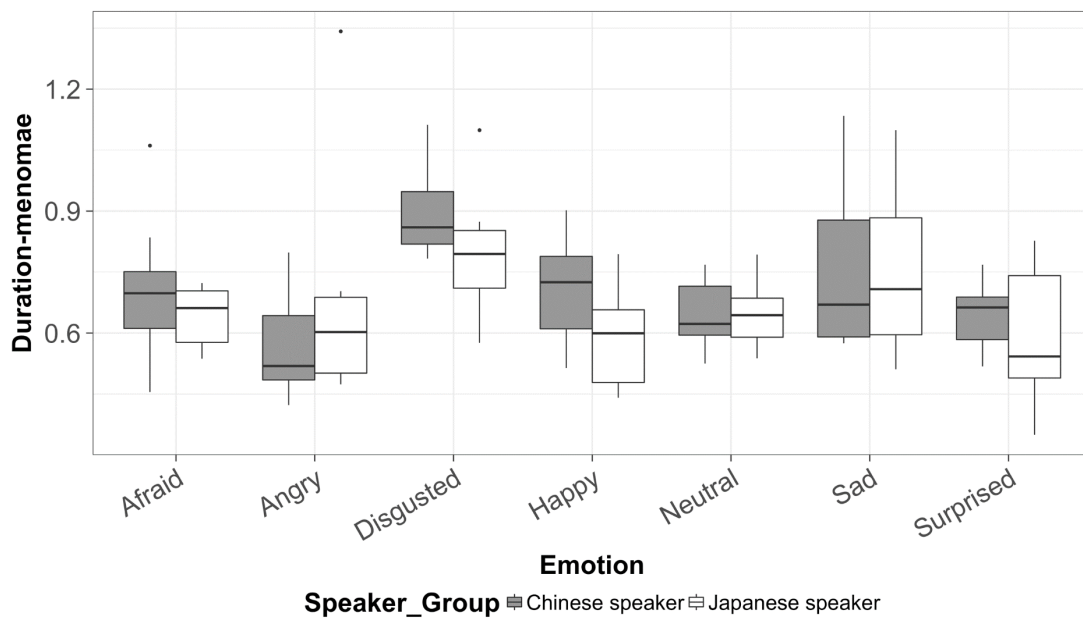


図 3-14 発話者グループ間および感情間の Duration-menomae

図 3-14 に示すように、タスク語「menomae」の Duration について、発話者 2 水準（日本語母語話者、中国人学習者）×感情 7 水準（中立、喜び、怒り、悲しみ、驚き、恐れ、嫌悪）の 2 要因分散分析を行なった結果、感情の主効果が有意であった ($F(6, 84) = 5.58, p < .001$)、発話者の主効果が有意ではなかった ($F(1, 14) = .55, p = .471$)、感情と発話者の交互作用は有意ではなかった ($F(6, 84) = 1.15, p = .343$)。

図 3-15 の通り、タスク語「aoyama」の Duration について、発話者 2 水準（日本語母語話者、中国人学習者）×感情 7 水準（中立、喜び、怒り、悲しみ、驚き、恐れ、嫌悪）の 2 要因分散分析を行なった結果、感情の主効果が有意であった ($F(6, 84) = 2.32, p = .039$)、発話者の主効果が有意ではなかった ($F(1, 14) = .64, p = .433$)、感情と発話者の交互作用が有意ではなかった ($F(6, 84) = .79, p = .579$)。

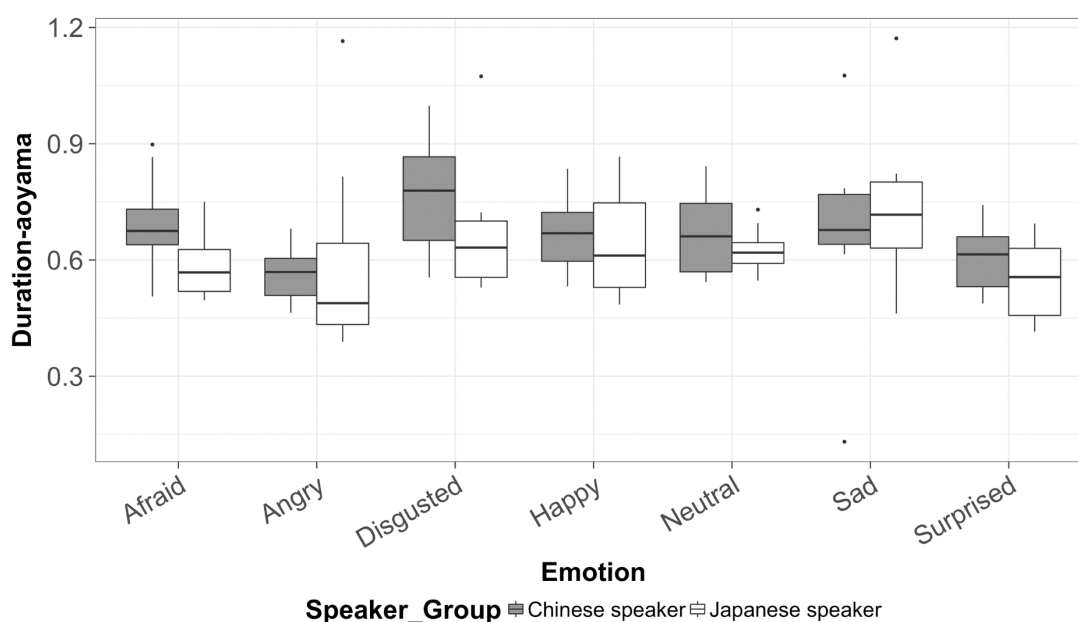


図 3-15 発話者グループ間および感情間の Duration-aoyama

Duration に関する検討から見ると、日本語母語話者による発話と中国人学習者による発話に Duration の差が見られたタスク語は「e」「me」「umi」「momo」であり、4 モーラのタスク語に Duration による差が見られなかった。また、日本語母語話者による発話では、「嫌悪」は他の感情と有意な Duration の差が観察され、中国人学習者による発話では、「怒り」「嫌悪」「恐れ」は他の感情と有意な Duration の差が見られた。

これらのことは、タスク語が 4 モーラ以上になると、感情を込めにくい可能性があることを示唆する。その原因の一つとして、有意味語単位の感情音声を表出するときには、言語学的情報が優先され音声での表出の必要性が薄れるという可能性が考えられる。さらに、日本語母語話者が感情音声を生成する際、Duration は「嫌悪」を他の感情と区別する音響的手がかりとなる可能性があり、中国人学習者が感情音声を生成する際、Duration は「怒り」「嫌悪」「恐れ」と他の感情と区別する音響的手がかりとなる可能性がある。特に、「怒り」と「嫌悪」は混同されやすい感情ペアとして、中国人学習者は Duration を用いてそれらを区別して発話する可能性があることが示唆される。

3.2.2.3. Intensity に関する検討

Intensity について、図 3-16 の通り、発話者 2 水準（日本語母語話者、中国人学習者）×感情 7 水準（中立、喜び、怒り、悲しみ、驚き、恐れ、嫌悪）の 2 要因分散分析を行なった。その結果、発話者および感情の主効果が有意であり（ $F(1, 14) = 5.30, p = .001$; $F(6, 84) = 6.41, p < .001$ ），発話者と感情の交互作用が有意であった（ $F(6, 84) = 2.34, p = .003$ ）。多重比較を行なった結果、中国人学習者による「喜び」発話の MedianIntensity は日本語母語話者より有意に大きく（ $F(1, 14) = 12.68, p = .003$ ），「悲しみ」発話の MedianIntensity は日本語母語話者より有意に小さかった（ $F(1, 14) = 5.98, p = .042$ ）。

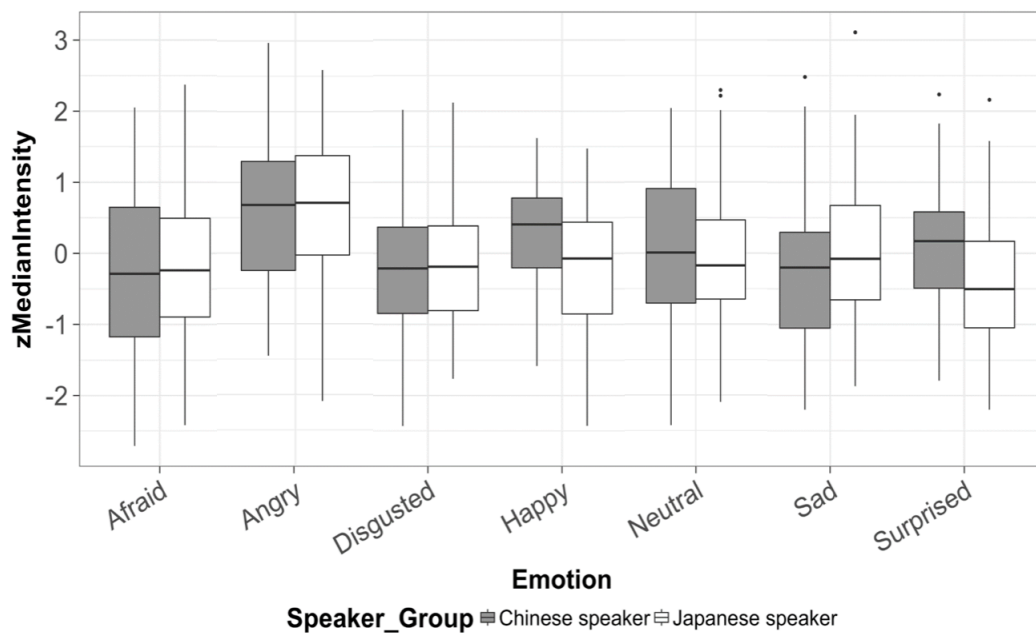


図 3-16 発話者グループ間および感情間の zMedianIntensity

MedianIntensity に関する感情間の違いは、日本語母語話者による発話において、「怒り」は「中立」より有意に大きかった ($t(7) = 4.43, p = .003$)。一方、中国人学習者による発話において、「中立」は「悲しみ」より有意に大きく ($t(7) = 5.15, p = .001$)、「喜び」は「嫌悪」より有意に大きかった ($t(7) = 6.31, p < .001$)。Intensity に関する検討から見ると、日本語母語話者による発話と中国人学習者による発話に差が見られた感情は「喜び」と「悲しみ」であり、日本語母語話者による発話では、「怒り」は他の感情と有意な Duration の差が観察され、中国人学習者による発話では、「中立」と「喜び」は他の感情と有意な Duration の差が見られた。これらのことは、日本語母語話者は Intensity を用いて「怒り」を他の感情と区別して発話する可能性が見られ、中国人学習者は Intensity を用いて「喜び」と「悲しみ」を区別する可能性があることが考えられる。

以上の音響分析の結果を踏まえ、さらに音声生成のメカニズムを探るために、次章では発声に関しての生理学的観察を行う。

第4章 EGGによる音声生理学的分析

4.1. 実験の目的

近年では声質と感情、態度の結びつきも積極的に考察され始めている（定延 2007）。感情音声においては、声の高低・長短・強弱以外の要因、特に声質が大きく関わっていることがよく知られている（Gordon & Ladefoged 2001, Menezes & Maekawa 2006, Wilson & Wharton 2006, Erickson et al. 2011, Scherer et al. 2015）。例えば、気息性と緊張性は感情音声の知覚と関係がある（Lugger & Yang 2007, Tahon et al. 2012）。また、気息性や非周期性を含んだ non-modal な声質が起きた場合、これらの強い感情を表したパラ言語情報が表現されている可能性が高い（石井他 2006）という報告もある。EGG 信号には音調に関連する F0 情報のみならず、声帯の振動様式の違いによる声質の情報が多く含まれ、パラ言語情報の記述に役に立つ（石井他 2013）。

さらに、Open Quotient は EGG 信号指標の一つであり、声帯の緊張、弛緩を示すパラメータとしてよく用いられる（Hanson et al. 2001, Henrich et al. 2001）。一般的に、喉が緊張するほど、Open Quotient が小さく、その中でりきみ声（pressed voice）がよく挙げられ、喉が弛緩するほど、Open Quotient が大きく、気息発声（breathy voice）がよく挙げられる。

本実験、中国語を母語とする日本語学習者が日本語で生成した感情音声と日本語母語話者が生成した感情音声は異なった発声様式が示されるのかについて、主に EGG 信号による Open Quotient を指標として中国人日本語学習者による感情表現の生理学的観察をすることを目的とする。

4.2. EGG 分析

4.2.1. 方法

本実験で用いた EGG 装置は「Glottal Enterprises 社 EG2-PCX2」（p14 図 1-1 参照）であった。全てのトークンについて、Open Quotient について測定を行なった。

EGG による声門波形を微分した dEGG（differentiated EGG）信号へと変換し、同一話者（日本語母語話者男性 20 代）による信号の測定実例を図 4-1 と図 4-2 に示す。得られた

dEGG 信号において、正のピークを声門閉鎖時点とし、負のピークを声門開放時点とする。O は声門開放期 (Open phase) を示し、C は声門閉鎖期 (Close phase) を示す。



図 4-1 「え」—中立

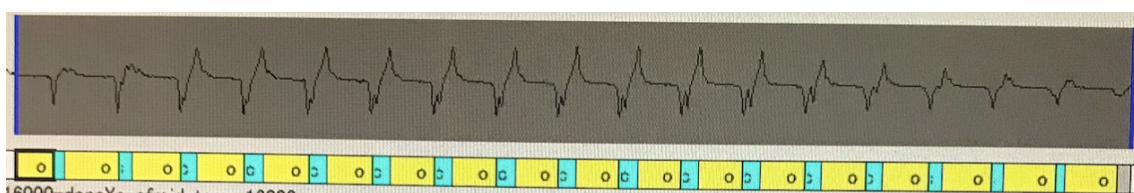


図 4-2 「え」—恐れ

EKG 信号がより正確に測定可能な男性発話者 (Fant et al. 1985, Klatt & Klatt 1990, Menezes & Maekawa 2006) による発話に注目し、知覚実験で収録した日本語母語話者男性 4 名、中国人学習者男性 4 名によるそれぞれの感情を込めた日本語音声計 616 トークン (8 名×7 感情×11 タスク語) を分析の対象とした。

4.2.2. 結果と考察

Open Quotient について、図 4-1 の通り、発話者 2 水準 (日本語母語話者、中国人学習者) ×感情 7 水準 (中立、喜び、怒り、悲しみ、驚き、恐れ、嫌悪) の 2 要因分散分析を行なった。その結果、発話者および感情の主効果が有意であり ($F(1, 14) = 62.98, p < .001$; $F(6, 84) = 16.08, p < .001$)、発話者と感情の交互作用が有意であった ($F(6, 84) = 5.42, p = .041$)。多重比較を行なった結果、中国人学習者による「怒り」発話の Open Quotient は日本語母語話者より有意に低く ($F(1, 14) = 19.12, p < .001$)、「悲しみ」発話の Open Quotient は日本語母語話者より有意に低かった ($F(1, 14) = 6.08, p = .036$)。また、Open Quotient に関する感

情間の違いは、日本語母語話者による発話観察されなかった。一方、中国人学習者による発話において、「恐れ」は「怒り」「嫌悪」「驚き」より有意に高く ($t(7) = 10.59, p < .001$; $t(7) = 8.73, p < .001$; $t(7) = 5.53, p = .001$) , 「嫌悪」は「怒り」より有意に高かった ($t(7) = 6.09, p = .002$) 。

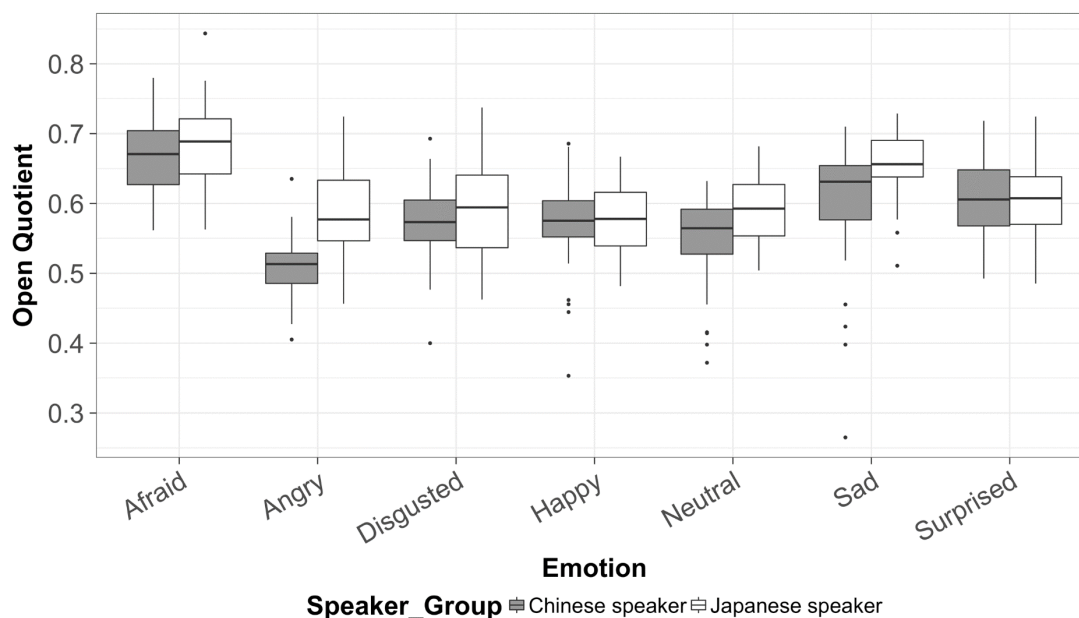


図 4-1 発話者グループ間および感情間の Open Quotient

Open Quotient に関する結果から見ると、日本語母語話者による発話と中国人学習者による発話に顕著な差が見られた感情は「怒り」と「悲しみ」であり、日本語母語話者による発話では、感情間の Open Quotient の差が見られなかった一方、中国人学習者による発話では、「恐れ」と「嫌悪」は他の感情と有意な差が観察された。特に中国人学習者による「怒り」と「悲しみ」は日本語母語話者より tense voice に近いことが考えられる。これらのことは、中国人学習者は Open Quotient を用いて「恐れ」と「嫌悪」を他の感情と区別して生成している可能性があることが示唆される。また、Scherer (1986) は「悲しみ」は弛緩した声 (lax voice) の特徴を持つと述べているが、本実験における日本語母語話者による「悲しみ」の傾向に近いと考えられる。

第5章 声質分析

5.1. 実験背景と目的

前章に述べたような中国人学習者の感情音声の音響的特徴が、背景言語である中国語の影響によるものなのかについて明らかにするため、中国語の感情音声について同様の分析を行う必要がある。まず、中国語の一部の感情音声研究について先行研究をまとめたものを付表2に示す。

本章では日本語母語話者が生成した日本語感情音声と、中国語を母語とする日本語学習者が生成した日本語および中国語感情音声に現れる声質的特徴(スペクトル特性とEGGによるOQ)ならびに音響特徴量を比較する。日中両言語の感情音声の違いは何か、また、中国語を母語とする日本語学習者の感情表出様式が背景言語の影響によるものなのかについて音声学的分析を通して明らかにすることを目的とする。

声質としてよく挙げられる種類として、creaky voice は glottal fry, vocal fry とも呼ばれ、声門閉鎖時間が長くなり、声門体積流がパルス的で極端に低いピッチが特徴づけられる発声様式の特徴をある程度有する声であることを意味する。自発音声においては、ピッチの低下と同様に、フィラー、発話末や句末、自信のない心理状態でよく観察される (Ishi 2006, 森他 2017)。Ishi et al. (2008) は、周期性と声帯パルスの類似性を利用した vocal fry の検出法を提案し、自然発話データに対して 73%の検出率と 13%の挿入誤り率を達成したと報告している。

また、気息性のある声質として、breathy voice は不完全な声門閉鎖により、声門体積流の波形は相対的な正弦波になり、声帯振動による周期音と同時に生じる乱流雑音によって生成される (森他 2017)。また、whisper や breathy は乱流雑音の程度や様態によって生じるとされている。感情音声との関連性として、会話の中で落胆などの心的状態を反映した発話では、呼吸筋や喉頭筋などの弛緩により、気息性のある低い発声 breathy voice が見られ、大きなスペクトル傾斜と低い HNR などにより特徴づけられる (Klatt & Klatt 1990, 森他 2014)。

5.2. スペクトル分析

5.2.1. 音声資料

分析に用いた音声資料は、「中立 (neutral)」「喜び (happy)」「激しい怒り (hot anger)」「押し殺した怒り (cold anger)」「悲しみ (sad)」「驚き (surprised)」「恐れ (afraid)」「嫌悪 (disgusted)」の 8 感情であった。「怒り」に関しては、F0 と Duration に関する結果は一貫しておらず、音響的な面から、別のものとして取り扱う必要がある (広瀬 2002, Scherer 2003) ため、本実験では、「怒り」を「激しい怒り」「押し殺した怒り」に分けて検討する。さらに、感情の度合いによって、各パラメータが異なる結果になる (Probst & Braun 2019) と報告されているため、本実験では各感情を 3 段階で発話してもらうこととした。「喜び」を例に挙げると、「やや喜んで」「喜んで」「非常に喜んで」という 3 段階で発話してもらい、感情が最も表出されていると考えられる第 3 段階のみを分析の対象とした。

また、タスク文の選定について、先行研究を踏まえ、感情を込めやすく、文自体に感情のない有意味文を用いた。「え」(石井 2006)、「そうですか」(前川・北川 2002, Maekawa 2004, 池本・鈴木 2008)、「そして、局所の強調の有無・場所およびそれと感情表出の関連性 (Liu 2011) があるために「オバマが当選した」の 3 つをタスク文とした。さらに、実験用の感情別のシチュエーションを提案した Liberman et al. (2002) を参考にしつつ、発話者の感情をより自然に引き出せるように、収録に用いる会話式の台本を作成した。全発話のシナリオを付録 1 に示す。感情「嫌悪」、タスク文「え」の例を以下に示す。発話者に B の発話台本の通りに発話してもらった。

あなたと同僚が誰かが吐いた汚物を片付けなければなりません。

A : 30 分以内でここを片付けなければなりませんね。

B : (嫌悪 : え) …見たくもないですね…

発話者計 20 名が生成した発話を収録した。日本語母語話者 10 名 (男女各 5 名, 平均年齢 30 歳, $SD = 4.92$) および中国人学習者 10 名 (男女各 5 名, 全員 N1 合格者, 平均年齢 28 歳, $SD = 3.67$) であった。中国人学習者が同意味のタスク文を中国語で生成した中国語感情音声も収録した。中国語感情音声タスクは、「啊 (a)」「是吗 (shi ma)」「奥巴马

当選了 (ao ba ma dang xuan le) 」である (括弧内はローマ字ピンインの表記を示す)。本章では聞き手によく伝わった発話者の感情音声のみに注目するため、日本語母語話者および中国人学習者各 3 名を聴取者として、自由記述と上手さについて判定してもらう予備的な判定実験を行なった。その結果により、感情が伝わっていない日中計 4 名 (男女各 2 名) の発話を分析対象から外した。よって、日本語母語話者 8 名 (男女各 4 名) および中国人学習者 8 名 (男女各 4 名)、計 16 名の発話者による感情音声を本章の分析対象とした。

5.2.2. 声質の分析方法

スペクトル特性は声質の特徴、特に緊張した発声と弛緩した発声を区別する有効な分析手法として知られている (Maddieson & Ladefoged 1985, 森他 2016, 2017, 前川他 2019)。本章では、スペクトル特性に関する音響関連量として、基本周波数と第 1 フォルマント周波数帯域の振幅値の差 (H1-A1)、基本周波数と第 3 フォルマント周波数周辺帯域でフィルタリングした信号の同期性 (F1F3syn) の 2 つのパラメータについて測定を行なった。F1 帯域は 100~1500 Hz, F3 帯域は 1800~4500 Hz に指定した。F1 と F3 帯域の波形振幅包絡の相互相関によって求める F1F3syn は、氣息性を示すパラメータであり、F1F3syn は 1 に近いほど氣息性が強く、0 に近いほど氣息性が弱いことを表す (Ishi 2004, 2010, 石井 2006)。H1-A1 は基本周波数成分のエネルギーをより周波数の高い倍音ないし共鳴周波数と比較することで、スペクトル傾斜を定量化したものである (前川・西川 2019)。スペクトルの傾斜が急になるほど、つまり高音域の周波数の振幅値が少ないほど、氣息性が増して知覚され、傾斜が緩やかならその逆の結果を示す (Klatt & Klatt 1990, Yoshinaga & Kong 2011)。

5.2.3. 結果と考察

日本語母語話者および中国人学習者による発話の各声質パラメータの分析結果について、声質による性差を避けるため、各パラメータは性別ごとに検討する。男性発話者による H1-A1, F1F3syn の結果を図 6-1, 図 6-2 に示し、女性発話者による H1-A1, F1F3syn の結果を図 6-3, 図 6-4 に示す。

図 6-1 の通り、男性発話者による発話の H1-A1 について、言語 3 水準 (日本語母語話者による日本語発話, 中国人学習者による日本語発話, および中国語発話) × 感情 8 水準 (中

立、喜び、激しい怒り、押し殺した怒り、悲しみ、驚き、恐れ、嫌悪)の2要因分散分析を行なったところ、感情の主効果は有意であり ($F(7, 63)=18.81, p<.001$) , 言語と感情の交互作用は有意であり ($F(14, 63)=3.58, p<.001$) , 言語の主効果は有意傾向であった ($F(2, 9)=3.64, p=.069$) 。言語と感情の交互作用が有意であったため、多重比較を行なった。その結果、日本語母語話者による日本語の「押し殺した怒り」の H1-A1 は中国語より有意に高く ($F(2, 9)=13.42, p=.002$) , 日本語母語話者による日本語の「嫌悪」は学習者による日本語と中国語より有意に高かった ($F(2, 9)=16.96, p<.001$) 。

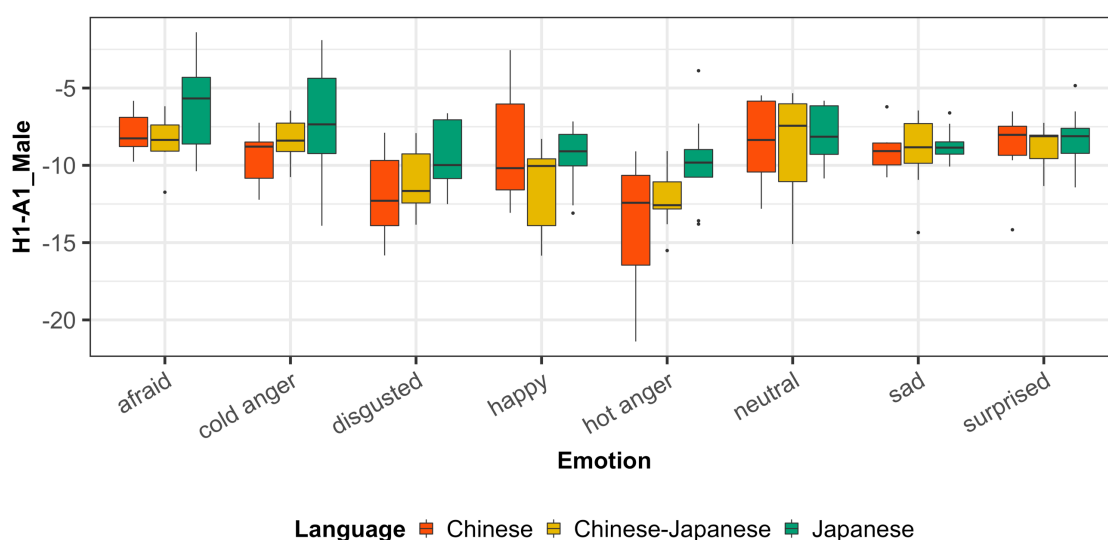


図 6-1 男性発話者による言語間および感情間の H1-A1

次に F1F3syn に関しては、図 6-2 の通り、男性発話者による発話の F1F3syn を言語 3 水準 (日本語母語話者による日本語発話、中国人学習者による日本語発話、および中国語発話) ×感情 8 水準 (中立、喜び、激しい怒り、押し殺した怒り、悲しみ、驚き、恐れ、嫌悪)の2要因分散分析を行なった結果、言語の主効果は有意であり ($F(2, 9)=17.21, p=.003$) , 言語と感情の交互作用 ($F(14, 63)=5.54, p=.004$) は有意であった。多重比較を行なった結果、日本語母語話者男性発話者による日本語の「恐れ」「押し殺した怒り」「喜び」「驚き」の F1F3syn は中国人学習者による日本語と中国語より有意に低かった ($F(2, 9)=33.15, p<.001, F(2, 9)=5.61, p=.048, F(2, 9)=15.97, p=.003, F(2, 9)=12.13, p=.007$) 。

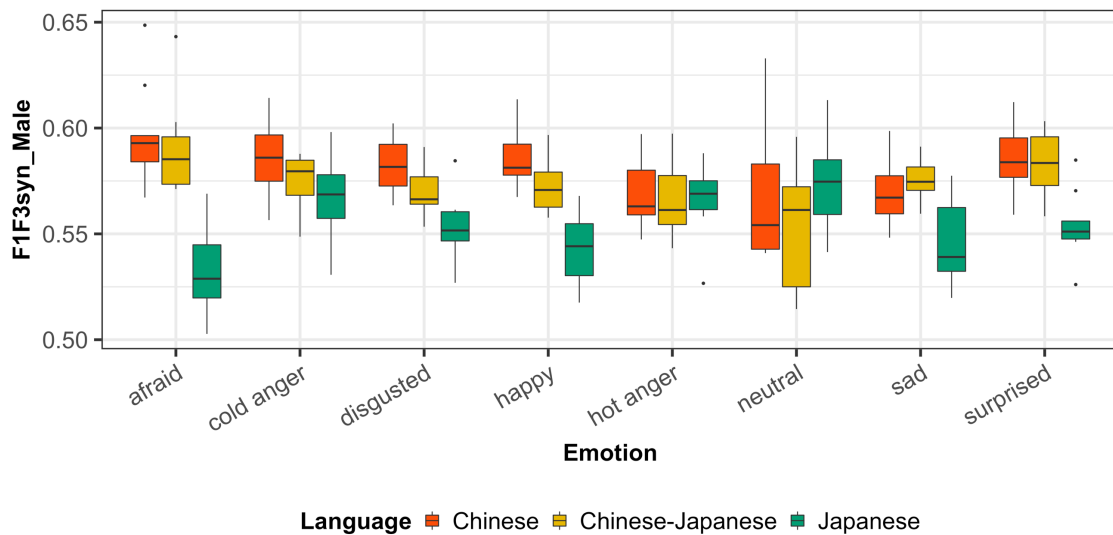


図 6-2 男性発話者による言語間および感情間の F1F3syn

図 6-3 の通り，女性発話者による発話の H1-A1 について，言語 3 水準（日本語母語話者による日本語発話，中国人学習者による日本語発話，および中国語発話）×感情 8 水準（中立，喜び，激しい怒り，押し殺した怒り，悲しみ，驚き，恐れ，嫌悪）の 2 要因分散分析を行なった結果，言語と感情の主効果は有意であった（ $F(2, 9) = 5.28, p = .034$ ， $F(7, 63) = 7.87, p < .001$ ）。言語と感情の交互作用（ $F(14, 63) = 3.84, p < .001$ ）は有意であったため，多重比較を行なった。その結果，日本語母語話者による日本語の「押し殺した怒り」の H1-A1 は中国人学習者による日本語より有意に高く（ $F(2, 9) = 5.04, p = .034$ ），日本語母語話者による日本語の「喜び」および「驚き」は中国人学習者による日本語と中国語より有意に高かった（ $F(2, 9) = 6.89, p = .015, F(2, 9) = 20.25, p < .001$ ）。

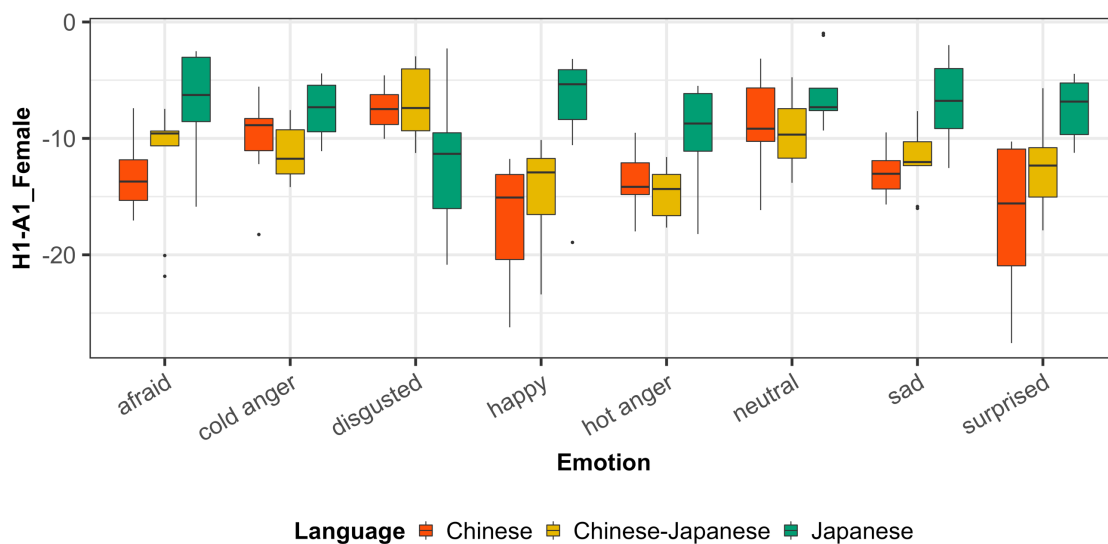


図 6-3 女性発話者による言語間および感情間の H1-A1

図 6-4 の通り，女性発話者による発話の F1F3syn について，言語 3 水準（日本語母語話者による日本語発話，中国人学習者による日本語発話，および中国語発話）×感情 8 水準（中立，喜び，激しい怒り，押し殺した怒り，悲しみ，驚き，恐れ，嫌悪）の 2 要因分散分析を行なった結果，感情の主効果は有意であった ($F(7, 63) = 14.95, p < .001$)。言語と感情の交互作用 ($F(14, 63) = 6.03, p < .001$) は有意であったため，多重比較を行なった。その結果，日本語母語話者女性発話者による日本語の「恐れ」「押し殺した怒り」「悲しみ」の F1F3syn は中国人学習者による日本語と中国語より有意に低く ($F(2, 9) = 6.48, p = .024, F(2, 9) = 7.19, p = .023, F(2, 9) = 23.83, p = .002$)，日本語母語話者による日本語の「嫌悪」は中国人学習者による日本語と中国語より有意に高い ($F(2, 9) = 13.94, p = .005$) ことが観察された。

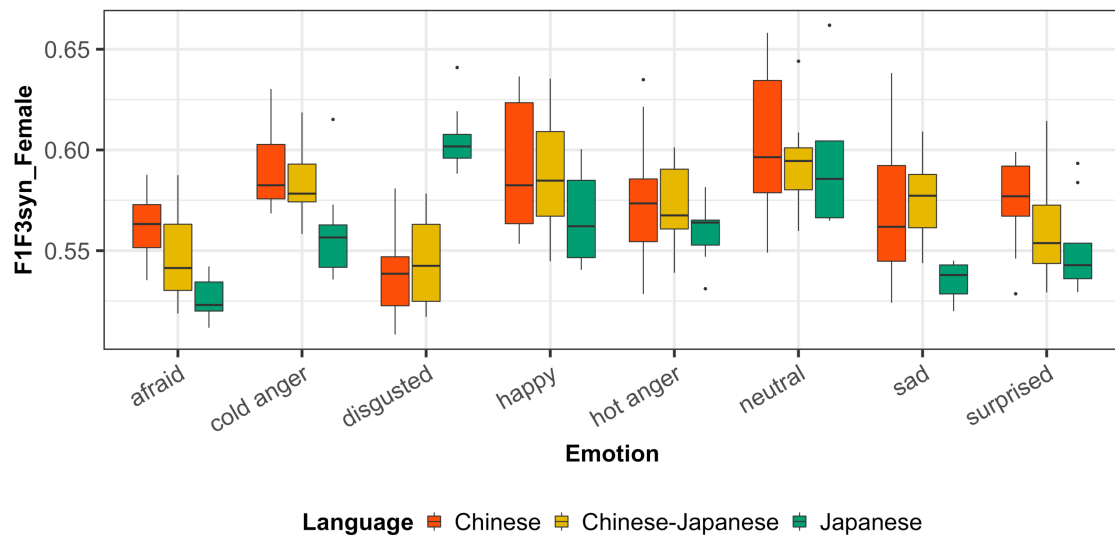


図 6-4 女性発話者による言語間および感情間の F1F3syn

スペクトル特性の分析をまとめると、男性発話者による発話の H1-A1 について、「押し殺した怒り」では中国人学習者による中国語感情音声により声帯が緊張した発声、また「嫌悪」では日本語母語話者による日本語発話により声帯が弛緩した発声が見られた。また、男性発話者による発話の F1F3syn に関しては、日本語母語話者男性発話者による日本語の「恐れ」「押し殺した怒り」「喜び」「驚き」は中国人学習者による日本語と中国語より氣息性の強い発声が観察された。次に女性発話者による発話の H1-A1 について、日本語母語話者と異なり、「喜び」「驚き」では中国人学習者による日本語感情音声と中国語感情音声ともに日本語母語話者より緊張した発声が目立ち、「押し殺した怒り」では中国人学習者による日本語発話により声帯が緊張した発声様式が見られた。また、女性発話者による発話の F1F3syn について、日本語母語話者による日本語の「恐れ」「押し殺した怒り」は中国人学習者による日本語と中国語より氣息性の強い発声が見られたが、「嫌悪」に関しては逆の傾向を示した。中国人学習者女性発話者による日本語と中国語の「嫌悪」は日本語母語話者女性発話者より氣息性が顕著に強いことが見られた。さらに、「嫌悪」に関する中国語学習者の性差も観察された。これらのことから、中国語を母語とする学習者は声帯を緊張させて感情音声を表出する可能性がある一方、日本語母語話者は氣息性の強い感情音声を生成する可能性が示唆される。

5.3. 声質と音響の複合的分析

5.3.1. 感情ごとの F0 と OQ の傾向

感情音声について、声質パラメータと音響特徴量を複合的に検討する必要があると知られている (Gobl & Chasaide 2003)。本実験で EGG を用いて収録し、EGG 信号による Open Quotient (以下 OQ) を計測した。音声資料と分析対象は前節と同様である。感情ごとおよび発話者グループごとの特徴を把握するために、F0 と OQ の平均値とを算出した。F0 の測定に関しては、LPC 逆フィルタによる残差波形の自己相関関数の最大ピークに基づいた処理を行っており、自己相関関数で F0 の sub-harmonic に対応するピークも、ある閾値を満たさなければならないという制約を追加した (Ishi 2005, 石井他 2006)。測定した結果を図 6-5 (J: 日本語母語話者による日本語発話, L: 中国人学習者による日本語発話) に示す。また発話者グループを比較しやすいように、感情別の結果を図 6-6 に示す。

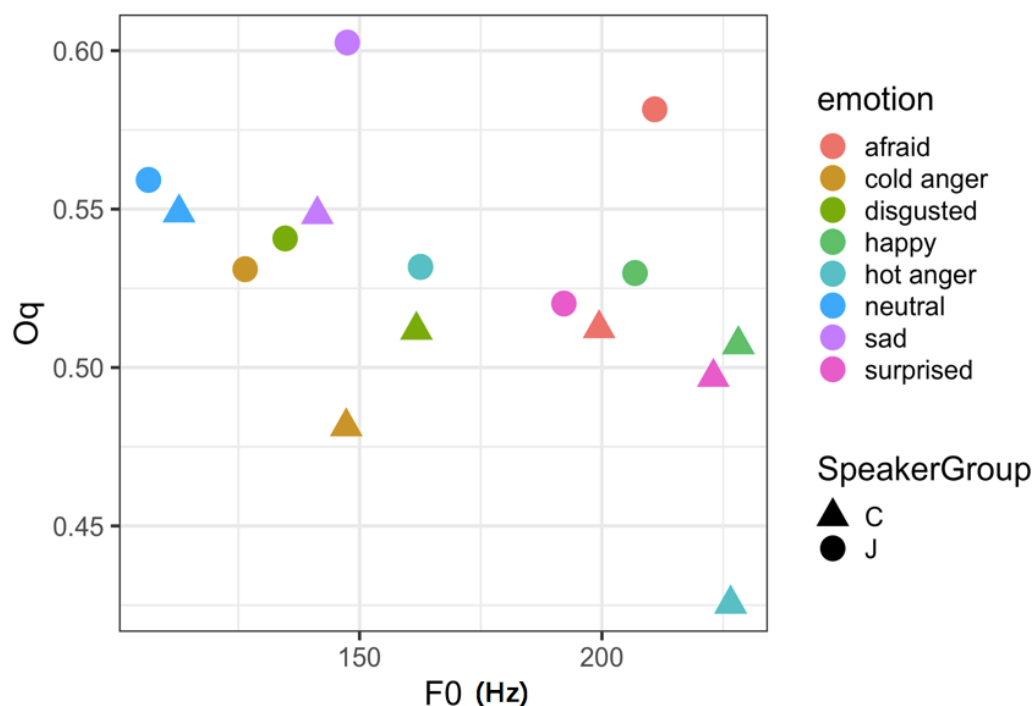


図 6-5 F0 と OQ による発話者グループ別の平均値

OQ について、発話者 2 水準×感情 8 水準の 2 要因分散分析を行なった結果、発話者および感情の主効果が有意であり ($F(1, 4) = 9.83, p = .035; F(7, 28) = 30.14, p < .001$)、発話者と感情の交互作用が有意であった ($F(7, 28) = 8.27, p < .001$)。多重比較を行なった結果、日本語母語話者による「激しい怒り」「押し殺した怒り」「恐れ」「嫌悪」発話の OQ は中国人学習者より有意に大きかった ($F(1, 4) = 19.68, p = .011; F(1, 4) = 18.74, p = .012; F(1, 4) = 39.19, p = .003; F(1, 4) = 8.28, p = .047$)。

また、F0 について、発話者 2 水準×感情 8 水準の 2 要因分散分析を行なったところ、発話者および感情の主効果が有意であり ($F(1, 4) = 29.09, p = .006; F(7, 28) = 33.47, p < .001$)、発話者と感情の交互作用が有意であった ($F(7, 28) = 5.49, p < .001$)。多重比較を行なった結果、日本語母語話者による「激しい怒り」「押し殺した怒り」「嫌悪」発話の F0 は中国人学習者より有意に低かった ($F(1, 4) = 35.78, p = .004; F(1, 4) = 8.29, p = .045; F(1, 4) = 8.88, p = .041$)。

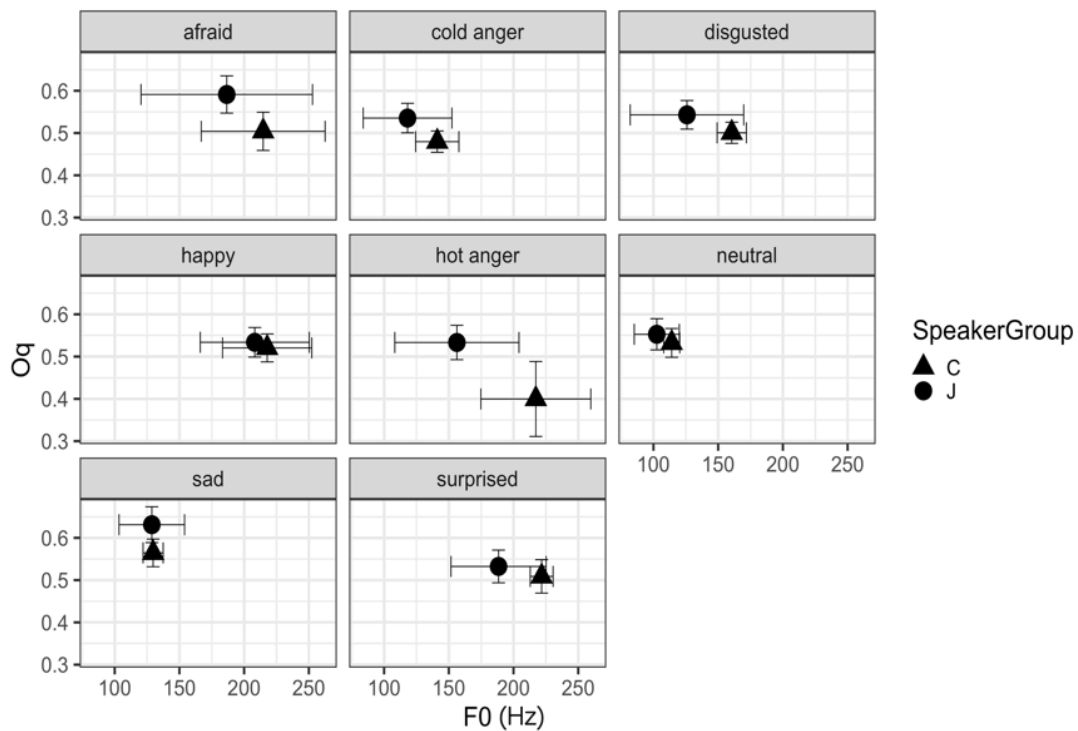


図 6-6 F0 と OQ による感情別の平均値と標準偏差

以上をまとめると、中国人学習者による「激しい怒り」「押し殺した怒り」「嫌悪」発話は日本語母語話者よりも OQ が大きく、F0 が低いことが観察された。これらの結果は、中国人学習者はより緊張した発声でこれら 3 種類の感情を表現しており、tense voice に近い発声様式を用いることを示唆している。

5.3.2. Oq-valued VRP

前節のような日本語母語話者と中国人学習者の感情表出様式の違いを考察するため、若狭他 (2019) にならい、発話時の OQ の変動を視覚的に捉える Oq-valued VRP (Voice Range Profile) を用いて、同程度の F0 と power における OQ が発話者グループによってどのような違いがあるかについて考察する。

F0 は 10Hz ごとに、power は 10dB ごとに区切り、該当する分析フレームにおける OQ の平均値を色に対応させた。F0 を横軸、Power を縦軸、OQ をカラーマップにより表示した。色と値の対応は、Oq が小さいほど、つまり声帯が緊張するほど赤くなり、OQ が大きいほど、つまり声帯が弛緩するほど緑になるように指定した。これにより、異なる power と F0 において、それぞれの OQ の平均値を視覚により直感的に理解することが可能になる。

感情が最もよく伝わった日本語母語話者および中国人学習者各 1 名による平均値を図 6-7、図 6-8 に示す。学習者の背景言語である中国語が学習者の感情表現への影響を分析するため、同じ中国人学習者が生成した中国語感情音声を図 6-9 に示す。

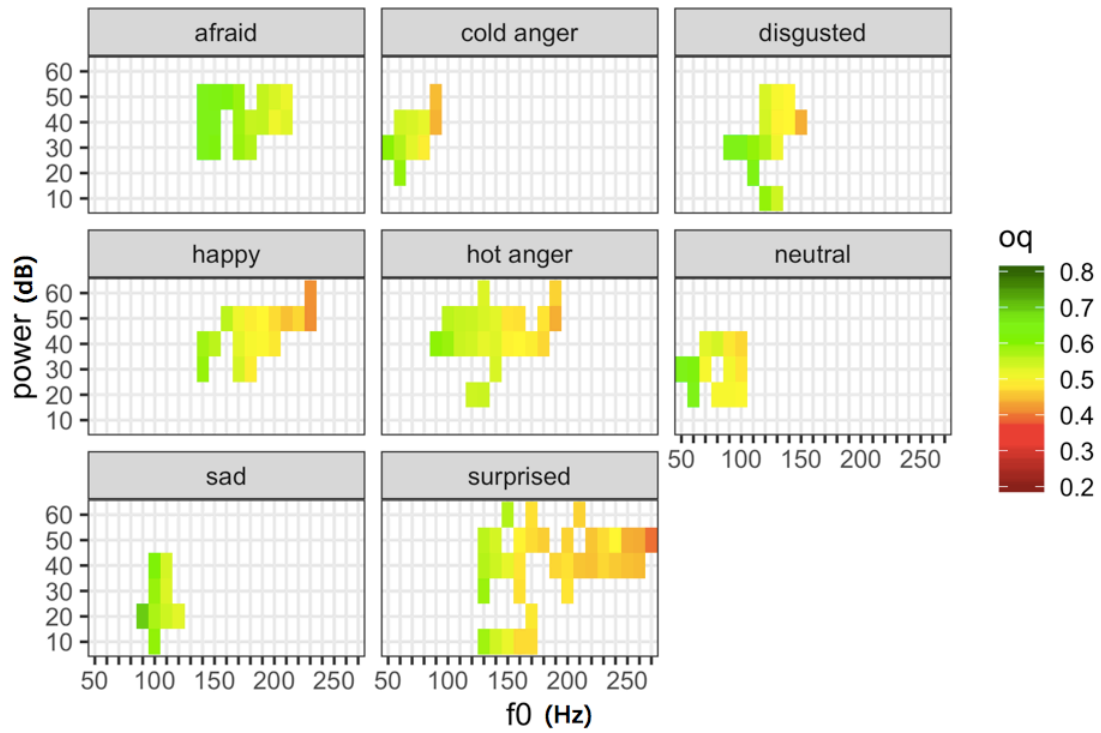


図 6-7 日本語母語話者による日本語感情音声の Oq-valued VRP

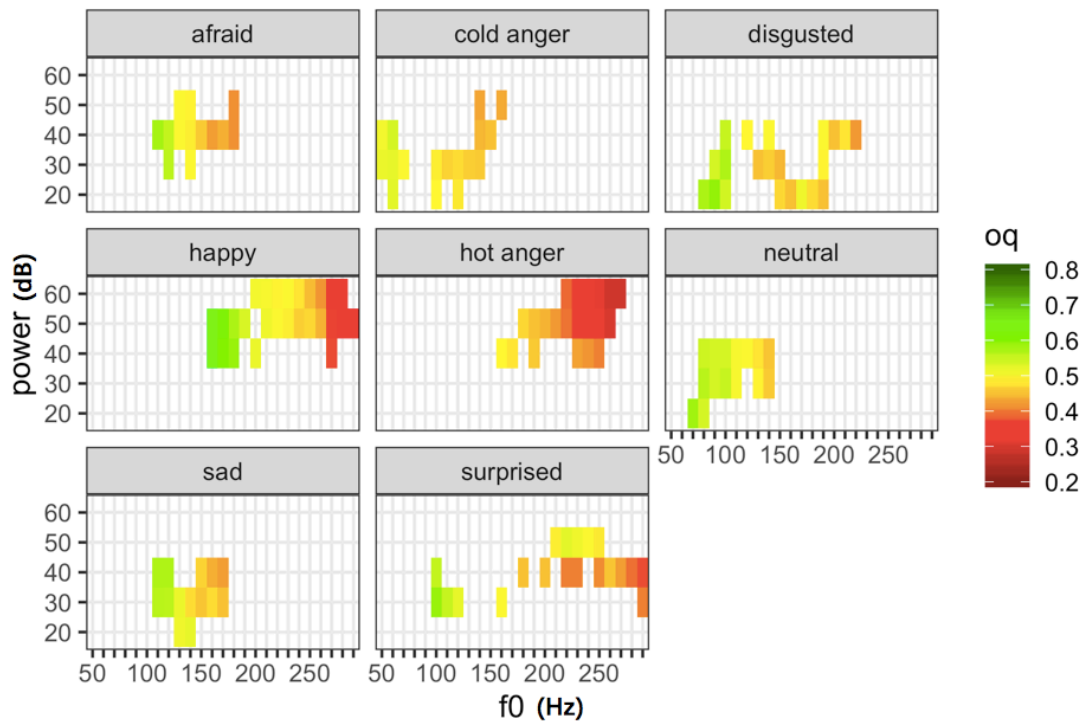


図 6-8 中国人学習者による日本語感情音声の Oq-valued VRP

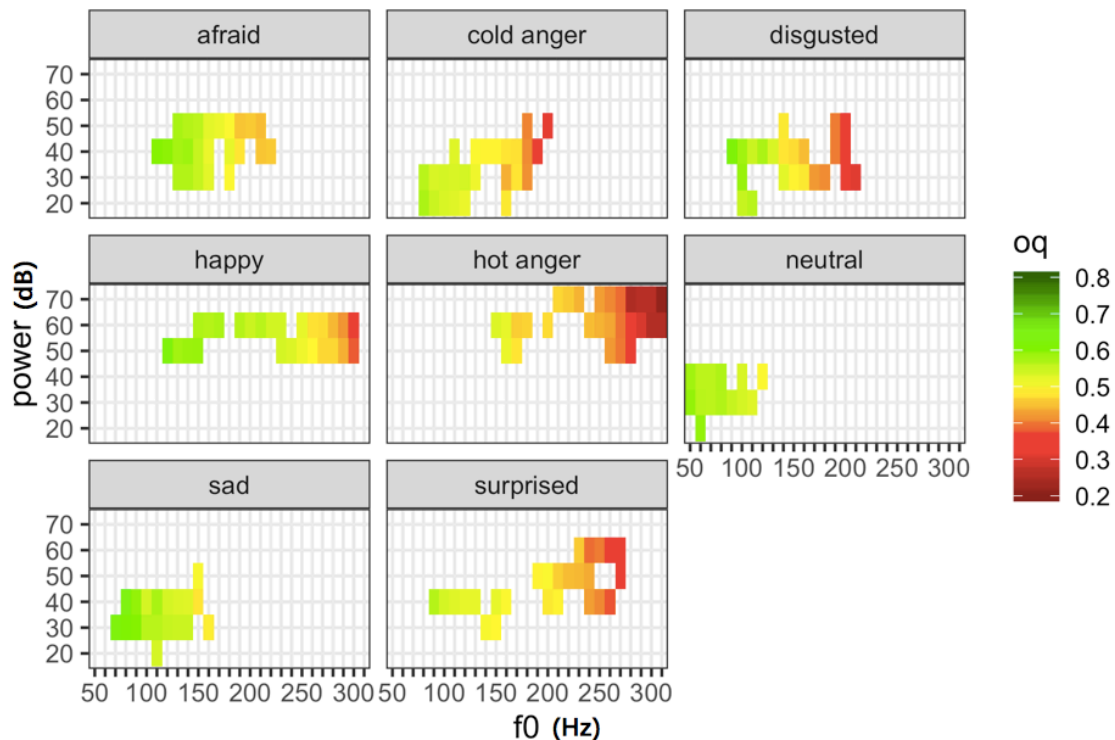


図 6-9 中国人学習者による中国語感情音声の Oq-valued VRP

結果として、発話者グループに関係なく、F0が高く、Powerが大きくなるほど、OQが小さくなっていることが観察された。この結果は、OQはF0と負の相関がある報告 Henrich et al. (2005)と一致している。

一方で、発話者によって発声様式が異なる傾向も観察された。例えば中国人学習者による「激しい怒り」（図 6-8）の特徴は、F0が230—280Hzの範囲で、OQが0.2—0.3まで顕著に下がることから「りきみ声」と考えられる。しかし、この特徴は日本語母語話者による「激しい怒り」では見られなかった。

同様の傾向は「喜び」の高音域にも見られた。さらに、図 6-9 の通り、中国語感情音声における「悲しみ」「恐れ」「中立」以外の感情に関して、OQが小さい値（赤色）を取っている部分が広く分布している。特に「激しい怒り」にも「りきみ声」を観察できた。これらの結果は、母語である中国語の感情表出様式が、学習した言語である日本語の感情表出に影響を与えた可能性を示唆するものと考えられる。

第6章 総合考察

6.1. 実験結果のまとめ

第2章では日本語母語話者および中国人学習者による感情音声の知覚について検討した。発話者と聴者の母語が一致している場合の感情同定率は、母語が一致していない場合の感情同定率より高いことが観察された。さらに、日本語母語話者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」および「悲しみ」と「中立」の4ペアの混同が顕著であり、中国人学習者による感情音声では、「悲しみ」と「恐れ」、「喜び」と「驚き」の2ペアが混同されやすいという結果を得た。これらの結果は、「内集団優位性 (in-group advantage)」の考えを支持しており、日本語母語話者の発話の意図する感情は中国人学習者の発話より混同が多いことを示したものである。

第3章では日本語母語話者および中国人学習者による感情音声にあらわれる音響的特徴の違いについて検討した。まずは F0 に関するパラメータについて、日本語母語話者と中国人学習者の発話の間では、F0Level, F0Max, F0Span および F0Mean の4つのパラメータに有意差が観察された。例えば、中国人学習者による発話の「怒り」は日本語母語話者による発話より F0Level が高く、「悲しみ」は日本語母語話者より F0Max が低く、F0Span が狭く、F0Mean が低く、「嫌悪」は日本語母語話者より F0Level が低く、F0Span が広く、「中立」は日本語母語話者より高かった。これらのことは、日本語母語話者と中国人学習者が F0 に関するパラメータを用いて感情音声を生成する際、「怒り」「悲しみ」「嫌悪」および「中立」に関して異なった F0 を用いる可能性を示唆する。Duration については、日本語母語話者による発話と中国人学習者による発話に Duration の差が見られたタスク語は「e」「me」「umi」「momo」であり、4モーラのタスク語に Duration による差が見られなかった。また、日本語母語話者による発話では、「嫌悪」は他の感情と有意な Duration の差が観察され、中国人学習者による発話では、「怒り」「嫌悪」「恐れ」は他の感情と有意な Duration の差が見られた。これらのことは、タスク語が4モーラ以上になると、感情を込めにくい可能性があることを示唆する。その原因の一つとして、有意味単語単位の感情音声を表出するとき、喉の緊張を長く続けることが困難である可能性がある。さらに、日本語母語話者が感情音声を生成する際、Duration は「嫌悪」を他の感情と区別する音響的手がかりとなる可能性があり、中国人学習者が感情音声を生成する際、Duration は「怒り」

「嫌悪」「恐れ」と他の感情と区別する音響的手がかりとなる可能性があると考えられる。特に、「怒り」と「嫌悪」は混同されやすい感情ペアとして、中国人学習者は Duration を用いてそれらを区別して発話する可能性があることが示唆される。なお、Intensity について、日本語母語話者による発話と中国人学習者による発話に差が見られた感情は「喜び」と「悲しみ」であった。また、日本語母語話者による発話では、「怒り」は他の感情と有意な Duration の差が観察され、中国人学習者による発話では、「中立」と「喜び」は他の感情と有意な Duration の差が見られた。これらのことは、日本語母語話者は Intensity を用いて「怒り」を他の感情と区別して発話する可能性が見られ、中国人学習者は Intensity を用いて「喜び」と「悲しみ」を区別する可能性があることが考えられる。

第 4 章では日本語母語話者および中国人学習者による感情音声にあらわれる Open Quotient の違いについて検討した。日本語母語話者による発話と中国人学習者による発話に Open Quotient の有意差が見られた感情は「怒り」と「悲しみ」であった。また、日本語母語話者による発話では、感情間の Open Quotient の差が見られなかった一方、中国人学習者による発話では、「恐れ」と「嫌悪」は他の感情と有意な差が観察された。特に中国人学習者による「怒り」と「悲しみ」は日本語母語話者より tense voice に近いことが考えられる。これらのことは、中国人学習者は Open Quotient を用いて「恐れ」と「嫌悪」を他の感情と区別して生成している可能性があることが考えられる。

第 5 章では中国語を母語とする日本語学習者が発話した中国語感情音声に加え、日本語母語話者と中国語を母語とする日本語学習者が発話した日本語感情音声に現れる声質的特徴 (H1-A1, F1F3syn および Open Quotient) および音響特徴量の複合的な比較を行なった。男性発話者による H1-A1 に関して、「押し殺した怒り」では中国人学習者による中国語感情音声よりも声帯が緊張した発声、「嫌悪」では日本語母語話者による日本語発話よりも声帯が弛緩した発声が見られ、男性発話者による発話の F1F3syn に関しては、日本語母語話者男性発話者による日本語の「恐れ」「押し殺した怒り」「喜び」「驚き」は中国人学習者による日本語と中国語より氣息性の強い発声が見られた。次に、女性発話者による H1-A1 に関して、「喜び」「驚き」では中国人学習者による日本語感情音声と中国語感情音声ともに日本語母語話者よりも緊張した発声が目立っており、「押し殺した怒り」では中国人学習者による日本語発話よりも声帯が緊張した発声様式が見られた。また、女性発話者による発話の F1F3syn について、日本語母語話者による日本語の「恐れ」「押し殺した怒り」は中国人学習者による日本語と中国語よりも氣息性の強い発声が見られ、中国人学

習者女性発話者による日本語と中国語の「嫌悪」は日本語母語話者女性発話者より氣息性は顯著に強い結果が得られた。これらの結果は、中国語を母語とする学習者は声帯が緊張した発声を用いて感情音声を表現する可能性がある一方、日本語母語話者は氣息性の強い感情音声を生成する可能性が示唆された。

6.2. 生成と知覚の関係

第2章～第4章の結果に基づき、感情音声の生成と知覚の関係を考察する。収録した発話に現れた音響的特徴および声質パラメータの Open Quotient (OQ) と知覚実験で得られた感情同定率の相関分析を行った結果を表 5-1 に示す。

表 5-1 音響的特徴と同定率の相関 (r)

母語話者	各感情の同定率						
	恐れ	怒り	嫌悪	喜び	中立	悲しみ	驚き
F0mean	.13	.01	.12	-.06	.07	-.02	.27
F0level	.18	.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
OQ	.40***	-.37	.13*	.06	-.17	-.36**	-.24
学習者	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
OQ	.64***	-.18	-.46*	-.21*	-.11	-.39**	.18

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

中国人学習者による発話では、知覚実験に混同が顕著だった感情ペアについて、「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 ($r(168)=-.47, p<.001$), Intensity が弱いと高くなる ($r(168)=-.38, p<.01$)。「恐れ」の同定率は F0mean が高いと高くなり ($r(168)=.57, p<.01$), Intensity が強いと高くなる ($r(168)=.23, p<.01$)。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 ($r(168)=.62, p<.001$)。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 ($r(168)=-.32, p<.05$) ことが観察された。日本語母語話者による発話では、上記のような感情の区別に手がかりとされた音響特徴と同定率の相関は見られなかった。

以上の感情音声の生成と知覚の関連についてまとめる。中国人学習者による発話では、「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、聴者は中国人学習者の「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。

知覚実験では、日本語母語話者による感情音声では、4 ペアの混同が顕著であり、中国人学習者による感情音声では、2 ペアが混同されやすいという結果を得た。音響分析および同定率との相関分析の結果から、知覚実験の結果の原因は、日本語母語話者の感情音声間の音響的特徴の違いが中国人学習者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。

中国人学習者の発話では「恐れ」と「悲しみ」の区別には F0mean と Intensity を制御して表しており、「驚き」と「喜び」では F0span や Duration を、「嫌悪」と「怒り」でも F0span を制御していることが示された。また、中国人学習者は Open Quotient を用いて「恐れ」と「嫌悪」を他の感情と区別して生成している可能性があることが考えられる。このような音響的特徴と声質の顕著な違いは日本語母語話者には見られず、中国人学習者は日

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。

6.3. 感情音声の発声様式における母語影響

第5章は日本語母語話者と中国人学習者による感情音声を対象に、スペクトル特性に関する手法を加え、音響特徴量および声質による複合的な検討を通して比較した。その結果、発話者の母語によって発声様式が異なった。スペクトル分析の結果により、中国人学習者が発話した「押し殺した怒り」と「嫌悪」は日本語母語話者より声帯が緊張した発声が見られた。また、「喜び」「悲しみ」「驚き」では、中国人学習者女性発話者による日本語感情音声と中国語感情音声ともに日本語母語話者より緊張した発声が観察された。一方、日本語母語話者が発話した「恐れ」「押し殺した怒り」は中国人学習者による日本語と中国語より氣息性の強い発声が見られた。

さらに、Oq-valued VRP の結果により、中国人学習者は「激しい怒り」「押し殺した怒り」「嫌悪」発話は日本語母語話者より OQ が大きく、F0 が低かったことが明らかになった。さらに、中国人学習者による日本語と中国語の「激しい怒り」ともに高音域で OQ が顕著に下がった。この特徴は日本語母語話者による発話では見られなかった。これらの結果は、日本語母語話者は中国人学習者より氣息性が強い発声が見られ、中国人学習者が生成した感情音声は日本語母語話者より緊張した発声で表出しており、またこのような発声様式は中国語の感情表出方法に影響されている可能性を示した。

Sadanobu (2004) は日本語におけるりきみ声について、苦しみか関心以外の場合はりきみ声を使うケースが少ないと述べている。Sadanobu (2004)の言う「りきみ声」は、本論文で扱う Vocal fry に相当するため、完全にも本論文で述べてきた口頭の緊張を伴う「りきみ」と完全に同じものではないが、感情の表出に用いられるという点で一部の特性が共通している。本実験の結果は、日本語母語話者とは異なり、中国人発話者はより広い範囲かつ高い頻度でりきみ声などの声質の変化を用いてパラ言語情報を伝達する可能性を示していると考えられる。

6.4. 今後の展望

本研究では、日本語母語話者と中国語を母語とする日本語学習者が発話した感情音声を対象に、知覚実験を通して母語話者と学習者の知覚傾向の違いを検討した。また音響分析、EGG 分析、スペクトル分析および VRP による複合的分析を通して日本語母語話者と中国人学習者の感情表出様式の違いについて検討した。本研究では限られた数の被験者および聴取者を対象とした実験を行ったため、引き続き双方の人数を増やし、より信頼性の高いデータを得るために今後努力する必要がある。

近年、感情音声の度合いによる音響特徴量の差に着目した研究もはじめられている。Probst & Braun (2019) では、「喜び」「激しい怒り」「押し殺した怒り」「悲しみ」「恐れ」「嫌悪」および「中立」を対象に、6名のドイツ語母語話者(男女各3名)に感情ごとに3段階の度合い(low, medium and extreme)で無意味語を発話してもらい、音響分析の結果、感情の度合いによってF0 平均値とF0 レンジが線形に変化することを報告している。本研究のさらなる課題として、日本語と中国語感情音声においても、このような音響特徴の感情の度合いによる線形変化が観察され得るのか、また感情の度合いがどのように音声に表出されるのかについて、収録済みのデータを用い、さらに分析していきたい。

さらに、タスク文の局所の強調の有無・場所およびそれと感情表出の関連性を調べた Liu (2011) では、中国語母語話者が「喜び」「怒り」「悲しみ」「驚き」「中立」を込めた発話を対象に音響分析を行なった。その結果、「怒り」において、強調された音節のF0 とDuration が顕著に変更することが見られた。本研究で扱った日本語の「オバマが当選した」ならびに中国語の「奥巴马当选了」のような文単位のタスク文においても今後、知覚実験を通して局所の強調の有無・場所を調べ、音響・声質分析を行ない、感情表出の関連性を明らかにしたいと考えている。また、同様の手法で、「驚き」以外の感情に関して、段落単位のタスクを作成し、より長い感情音声を対象に分析することも試みたい。

感情音声研究の限界として、被験者が実際に経験している感情を客観的に測定することが難しい点がある(森他 2014)。本研究では実験条件の統制のために、同じ発話内容で感情だけ異なる音声を収録し、感情が音声に与える影響を直接評価することを優先したが、感情音声研究の中で、真の感情経験を伴った自発音声も注目され始めている(Campbell & Erickson 2004, Ishi et al. 2010, Scherer et al. 2015)。将来的には、自然対話における感情音声の表出も観察していきたいと考えているが、さらにマン—マシン・インターフェースでの応用なども視野に入れ、例えばロボットと対面場面においてもどのように音声によって感情が表出されるのかについてもデータを収集し、分析していきたいと考えている。

パラ言語情報の一つである態度音声も、感情音声同様、コミュニケーションで重要な役割を果たしているとされている（Cheang & Pell 2008, Burgers et al. 2012, Rivière et al. 2018, Truesdale et al. 2018）。例えば Gu et al. (2011) は、中国語母語話者が生成した態度音声（音声による態度情報の伝達）を用い、知覚実験と音響分析を行なっている。このような先行研究を踏まえ、態度ペアとなる「友好」「敵対」, 「丁寧」「失礼」, 「本気」「冗談」, 「賞賛」「非難」を対象に、日本語母語話者と中国語を母語とする日本語学習者に文単位のタスク文を発話してもらうことを検討している。その際には、発話者が容易にその態度を表現できるように、発話シチュエーションとシナリオも提供する。例えば、「もう一度繰り返しましょうか？」をタスク文とし、中立発話の指示では「教科書を音読するような口調で読み上げてください（中立）」, 丁寧発話の指示では「あなたは会社のアフターサービスの電話受付係です。顧客からの電話で修理センターの場所を伝えました。しかし顧客は聞き取れなかったようです。あなたは顧客に：「（丁寧）？」」, 失礼発話の指示では「あなたは会社の部長です。部下が何度も重大ミスを起こしてしまいました。あなたは会社の注意事項について、その部下に：「（失礼）？」」といったようなシナリオを通して、状況に即して音声がどのように変化するのかといった研究も興味深く、将来的に多くの研究テーマに発展させることが可能と考えられる。

これらの感情・態度音声に関して多角的な研究を今後も続け、パラ言語情報に関わる音響・生理学的特徴をさらに検討し、より自然で円滑なコミュニケーションとは何かについて考察を重ねていきたいと思う。また、できれば中国語を母語とする日本語学習者に有益な知見を提供し、日本語学習指導に役立つツールの開発にも貢献したいと考えている。

参考文献

- Albas, D., McCluskey, K., Albas, C. (1976). Perception of the emotional content of speech: A comparison of two Canadian groups. *Journal of Cross-Cultural Psychology*, 7(4), 481–489.
- Altrov, R. Pajupuu, H. (2015). The influence of language and culture on the understanding of vocal emotions. *Journal of Estonian and Finno-Ugric Linguistics*, 6, 11–48.
- Amir, N., Almogi, B., Gal, R. (2004). Perceiving Prominence and Emotion in Speech—A Cross Lingual Study. *In Proceedings of Speech Prosody*, 23–26.
- 赤木 正人 (2010). 音声に含まれる感情情報の認識:感情空間をどのように表現するか 日本音響学会誌, 66, 393–398.
- 浅井 拓也 (2017). 発話目的の違いに注目した感情音声発話時の調音運動の比較に関する研究 北陸先端科学技術大学院大学情報科学研究科, 修士学位論文.
- Bachorowski, J., Owren, M. (1995). Vocal expression of emotion: Acoustic properties of speech are associated with emotional intensity and context. *Psychol Sci*, 6, 219–224.
- Banno, H., Hata, H., Morise, M., Irino, T., Kawahara, H. (2007). Implementation of realtime STRAIGHT speech manipulation system: Report on its first implementation. *Acoustic Science and Technology*, 28, 140–146.
- Banse, R., Scherer, K. (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology*, 70(3), 614–636.
- Bent, T., Bradlow, A. (2003). The interlanguage speech intelligibility benefit. *Journal of the Acoustical Society of America*, 114(3), 151–156.
- Bruce, G., Touati, P. (1992). On the analysis of prosody in spontaneous speech with exemplification from Swedish and French. *Speech Communication*, 11, 453–458.
- Burgers, C., Mulken, M., Schellens, P. (2012). Type of evaluation and marking of irony: The role of perceived complexity and comprehension. *Journal of Pragmatics*, 44, 231–242.
- Burkhardt, F., Sendlmeier, W. (2000). Verification of acoustical correlates of emotional speech using formant-synthesis. *In Proceedings of ISCA ITRW on speech and emotion*, 114(3), 1600–1610.
- Cabral, J. P., Oliveira, L. C. (2006). Emovoice: a system to generate emotions in speech. *In Ninth International Conference on Spoken Language Processing*.

- Campbell, N. (2000). Databases of Emotional Speech. *In Proceedings of ISCA (International Speech Communication and Association) ITRW on Speech and Emotion*, 34–38.
- Campbell, N., Erickson, D. (2004). What do people hear? A study of the Perception of Non-verbal Affective Information in Conversational Speech. *Journal of the Phonetic Society of Japan*, 8, 9–28.
- Cheang, H. S., Pell, M. D. (2008). The sound of sarcasm. *Speech communication*, 50(5), 366-381.
- Chronaki, G., Wigelsworth, M., Pell, M., Kotz, S. (2018). The development of cross-cultural recognition of vocal emotion during childhood and adolescence. *Scientific Reports*, 8(1), 1-17.
- Chuenwattanapranithi, S., Xu, Y., Thipakorn, B., Maneewongvatana, S. (2008). Encoding emotions in speech with the size code — A perceptual investigation. *Phonetica*, 65, 210-230.
- Doval, B., C. d’Alessandro. (1997). Spectral correlates of glottal waveform models: an analytic study. *In Proceedings of ICASSP*, 1295–1299.
- Drioli, C., Tisato, G., Cosi, P., Tesser, F. (2003). Emotions and voice quality: experiments with sinusoidal modeling. *In ISCA Tutorial and Research Workshop on Voice Quality: Functions, Analysis and Synthesis*.
- Ekman, P. (1992). An Argument for Basic Emotions. *Cognition And Emotion*, 6, 169–200.
- Ekman, P., Friesen, W. V. (1969). The repertoire of nonverbal behavior: Categories, origins, usage, and coding. *Semiotica*, 1, 49-98.
- Elfenbein, H. A. (2013). Nonverbal dialects and accents in facial expressions of emotion. *Emotion Review*, 5(1), 90-96.
- エリクソン ドナ・大西 真由子・栗原 はるか (2003). 感情音声の認識に関する男女差 岐阜私立女子短期大学研究紀, 53, 85–90.
- 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.
- エリクソン ドナ・昇地 崇明・メネゼス キャロリーヌ・河原 英紀・榊原 健一・リリヤール アルベール (2011). 日・米母語話者による感情音声知覚の比較——基本周波数以外の音響的特徴から得られる情報に注目して—— 杉藤美代子 (編) 音声文法 (pp. 43–60) くろしお出版.
- Erickson, D. (2005). Expressive Speech: Production, perception and application to speech synthesis.

- Acoustic Science and Technology*, 26, 317–325.
- Erickson, D., Zhu, C., Kawahara, S., Suemitsu, A. (2016). Articulation, Acoustics and Perception of Mandarin Chinese Emotional Speech. *Open Linguistics*, 2, 620–635.
- Estill, J. (1992). *Primer of compulsory figures*. Santa Rosa, CA: Estill Voice Training Systems.
- Fant, G., Liljencrants, J., Lin, Q. (1985). A four-parameter model of glottal flow. *STL-QPSR*, 1-13.
- Fant, G., Kruckenberg, A. (2004). An integrated view of Swedish prosody. *Speech Acoustic and Phonetics*, Kluwer Academic Publishers, 249-300.
- Forell, M. (2007). *Acoustic Correlates of Perceived Emotions in Speech*. Masters Thesis in Speech Communication, Royal Institute of Technology, KTH.
- 藤江 真也・江尻 康・菊池 英明・小林 哲則 (2005). 肯定的/否定的発話態度の認識とその音声対話システムへの応用 電子情報通信学会論文誌 3, 394–395.
- Fujisaki, H. (1996). Prosody, Models, and Spontaneous Speech. In Y. Sagisaka, N. Campbell & N. Higuchi (Eds.), *Computing Prosody: Computational Models for Processing Spontaneous Speech* (pp. 27–42). Berlin, Springer.
- Gangamohan, P., Sudarsana, R. K., Suryakanth, V. G., Yegnanarayana, B. (2014). Excitation source features for discrimination of anger and happy emotions. *In Proceedings of INTERSPEECH*, Singapore, 1253–1257.
- Gobl, C., Ní Chasaide, A. (2003). The role of voice quality in communicating emotion, mood and attitude. *Speech Communication*, 40, 189-212.
- Gordon, M., Ladefoged, P. (2001). Phonation types: a cross- linguistic overview. *Journal of Phonetics*, 29, 383-406.
- Gu, W., Zhang, T., Fujisaki, H. (2011). Prosodic analysis and perception of Mandarin utterances conveying attitudes. *In Twelfth Annual Conference of the International Speech Communication Association*, 1069-1072.
- Hanson, H. (1997). Glottal characteristics of female speakers: Acoustic correlates. *Journal of the Acoustical Society of America*, 101, 466-481.
- Hanson, H., Stevens, K., Kuo, H., Chen, M., Slifka, J. (2001). Towards models of phonation. *Journal of Phonetics*, 29, 451- 480.

- Henrich, N., d'Alessandro, C., Doval, B. (2001). Spectral correlates of voice open quotient and glottal flow asymmetry: Theory, limits and experimental data. *In Proceedings of EUROSPEECH 2001*, 47-50.
- Henrich, N., d'Alessandro, C., Doval, B., Castellengo, M. (2005). Glottal open quotient in singing: Measurements and correlation with laryngeal mechanisms, vocal intensity, and fundamental frequency. *The Journal of the Acoustical Society of America*, 117(3), 1417-1430.
- 広瀬 啓吉 (2002). 音声情報処理におけるパラ・非言語情報 日本音響学会講演論文集, 1(1-10-10), 243-246.
- 馮 丹妮 (2012). 中国語の発話態度の音響的特徴と認識——中国語母語話者と日本語母語話者を対象に—— 神戸大学大学院国際文化学研究科, 修士学位論文.
- Ibrakhim, I. (2005). 日本語音声の感情情報表現に見られる韻律的特徴の研究 大阪大学大学院文学研究科, 博士学位論文.
- イ ボギョン (2007). 韓国語を母語とする日本語学習者における日本語の「パラ言語情報」に関する実験音声学的研究——「問い返し」と「疑い」の実現および知覚の側面から—— 大阪大学大学院文学研究科, 博士学位論文.
- 飯田 赤美・伊賀 聡一郎・安村 通晃 (1997). 音声にみられる感情表現について：聴取実験結果の分析 音声言語情報処理, 15-20.
- 池本 真知子・鈴木 直人 (2008). 感情表出時の声質評価尺度の作成 感情心理学研究, 15(2), 80-88.
- 今泉 敏 (1994). 声の聴覚心理的評価 VI, 日本音声言語医学会 (編) 第2版 声の検査法基礎編, (pp. 151-172) 医歯薬出版, 東京.
- Ishi, C. T., Mokhtari, P., Campbell, N. (2003). Perceptually-related Acoustic-Prosodic Features of Phrase Finals in Spontaneous Speech. *In Proceedings of Eurospeech 2003*, 405-408.
- Ishi, C. T. (2004). Analysis of autocorrelation-based parameters for creaky voice detection. *In Proceedings of Speech Prosody 2004*, 643-646.
- Ishi, C. T. (2004). A new acoustic measure for aspiration noise detection. *International Conference on Spoken Language Processing 2004*, 2, 941-944.
- Ishi, C. T., Ishiguro, H., Hagita, N. (2005). Proposal of acoustic measures for automatic detection of voice vocal fry. *In Proceedings of Eurospeech 2005*, 481-484.

- 石井 カルロス寿憲・石黒 浩・萩田 紀博 (2006). 韻律および声質を表現した音響特徴と対話音声におけるパラ言語情報の知覚との関連 情報処理学会論文誌, 47, 1782–1792.
- Ishi, C. T., Ishiguro, H., Hagita, N. (2007). Acoustic and EGG analysis of pressed phonation. *International Congress of Phonetic Sciences ICPHS 2007*, 2057–2060.
- Ishi, C. T., Sakakibara, K., Ishiguro, H., Hagita, N. (2008). A Method for Automatic Detection of Vocal Fry. *IEEE Transactions on Audio, Speech, and Language Processing*, 16, 47–56.
- Ishi, C. T., Ishiguro, H., Hagita, N. (2008). Automatic extraction of paralinguistic information using prosodic features related to F0, duration and voice quality. *Speech Communication*, 50(6), 531–543.
- Ishi, C.T., Ishiguro, H., Hagita, N. (2010). Analysis of the Roles and the Dynamics of Breathy and Whispery Voice Qualities in Dialogue Speech. *EURASIP Journal on Audio, Speech, and Music Processing*, 2010, 1-12.
- 石井 カルロス寿憲 (2015). 音声対話中に出現するパラ言語情報と音響関連量 日本音響学会誌, 71(9), 476-483.
- Ishige, M., Niimi, S., Mori, K. (1996). Electroglottography (EGG). *The Japan Society of Logopedics and Phoniatrics*, 37, 345–354.
- Ito, M. (2004). Politeness and voice quality — The alternative method to measure aspiration noise. *In Proceedings of Speech Prosody 2004*, 213-216.
- Izard, C. E. (1977). *Human Emotions*, Plenum Press.
- Japan Foundation. (2015). Survey on Japanese-Language Education Abroad 2015. <https://www.jpf.go.jp/e/project/japanese/survey/result/survey15.html> (accessed Dec 12, 2018)
- Johnstone, T. (2001). The effect of emotion on voice production and speech acoustics. PhD Thesis, University of Western Australia.
- Kane, J., Scherer, S., Morency, L., Gobl, C. (2013). A comparative study of glottal open quotient estimation techniques. *In Proceedings of INTERSPEECH 2013*, 1658-1662.
- 甲斐 朋子・田渕 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.
- Kasuya, H., Yoshizawa, M., Maekawa, K. (2000). Roles of voice source dynamics as a conveyer of paralinguistic features. *in Proceedings of the 6th International Conference on Spoken Language Processing (ICSLP 2000)*, 345–348.

- 粕谷 英樹・楊 長盛 (1995). 音源から見た声質 (< 小特集> 声質: 音声言語の多様性に迫る) 日本音響学会誌, 51(11), 869-875.
- Klasmeyer, G., Sendlmeier, W. F. (2000). Voice and Emotional States. In *Voice Quality Measurement, Singular Thomson Learning*. 339-358.
- Klatt, D. H., Klatt, L. C. (1990). Analysis, synthesis, and perception of voice quality variations among female and male talkers. *Journal of the Acoustical Society of America*, 87(2), 820-857.
- 国際音声学会(編) 竹林 滋・神山 孝夫(訳) (2003). 国際音声記号ガイドブック 大修館書店.
- 郡 史郎 (1989). 発話の音調を規定する要因——日本語イントネーション論—— 吉沢典男教授追悼論文集, 116-127.
- 郡 史郎 (1998). 「日本語のイントネーション: 型と機能」『アクセント・イントネーション・リズムとポーズ』三省堂
- Koolagudi, S. G., Rao, K. S. (2012). Emotion recognition from speech: a review. *International journal of speech technology*, 15(2), 99-117.
- Ladefoged, P. (1983). The linguistic use of different phonation types. In D. Bless & J. Abbs (Eds.), *Vocal fold physiology: Contemporary research and clinical issues*, 351-360. San Diego: College Hill Press.
- Laver, J. (1980). Phonatory settings. *The Phonetic Description of Voice Quality*, Ch.3, 93-135, Cambridge University Press, Cambridge.
- Levy, E. S. (2009). On the assimilation-discrimination relationship in American English adults' French vowel learning. *The Journal of the Acoustical Society of America*, 126(5), 2670-2682.
- Li, A., Fang, Q., Hu, F., Zheng, L., Wang, H., Dang, J. (2010). Acoustic and articulatory analysis on Mandarin Chinese vowels in emotional speech. *7th International Symposium on Chinese Spoken Language Processing*.
- Li, X., Luo, M. (2015). On the difference between perceived emotional speech in Japanese learners between native speakers —— “joy” “anger” as the center. *The Science Education Article Collects*, 3, 176-178.
- Li, X., Luo, M. (2015). Japanese learner of Chinese perception of the extent of the Japanese emotional speech and prosodic features. *The Science Education Article Collects*, 12, 162-163.

- Li, X. (2016). Perception of Japanese Emotional Speech by Chinese Learners. MD Thesis, Dalian University of Foreign Languages.
- Li, X., Albin A., Hayashi, R. (2018). Perception of Japanese emotional speech by Chinese learners of L2 Japanese: Comparing the perception of learner vs. native speech. *The 13th Phonetic Conference of China PCC 2018*.
- Li, X., Albin, A., Ishi, C., Hayashi, R. (2019). Japanese Emotional Speech Produced by Chinese Learners and Japanese Native Speakers: Differences in Perception and Voice Quality. *International Congress of Phonetic Sciences ICPHS 2019*.
- 李 歆玥・石井 カルロス寿憲・林 良子 (2019). EGG を用いた日本語感情音声の分析 —— 日本語母語話者および中国人学習者による発話を対象に—— 日本音響学会秋季研究発表会講演論文集.
- 李 歆玥・エレン アルビン・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.
- 李 歆玥・エレン アルビン・林 良子 (2018). 中国語を母語とする日本語学習者による感情音声の生成と知覚の関係 第 32 回日本音声学会全国大会予稿集, 261-266.
- 李 歆玥・羅 米良・林 良子 (2018). 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 音声研究, 22(2), 31-38.
- Liberman, M., Davis, K., Grossman, M., Martey, N., Bell, J. (2002). Emotional Prosody Speech and Transcripts. Linguistic Data Consortium. <https://catalog ldc.upenn.edu/LDC2002S28> (accessed June 10, 2019)
- Liu, Y. (2011). A study of emotional speech in Mandarin. MD Thesis, Nanjing Normal University.
- Liu, P., Pell, M. D. (2012). Recognizing vocal emotions in Mandarin Chinese: A validated database of Chinese vocal emotional stimuli. *Behavior research methods*, 44(4), 1042-1051.
- Lugger, M., Yang, B. (2007). The relevance of voice quality features in speaker independent emotion recognition. *In Proceedings of ICASSP 2007*, 4, 17-20.
- Maddieson, I., Ladefoged, P. (1985). “Tense” and “lax” in four minority languages of China. *J. Phonetics*, 13, 433-454.
- Maekawa, K. (2004). Production and perception of ‘Paralinguistic’ information. *Proceedings of Speech Prosody 2004*, 367-374.

- Maekawa, K. Kikuchi, H. (2005). Corpus-based Analysis of Vowel Devoicing in Spontaneous Japanese—An Interim Report— *Voicing in Japanese*, 84, 1–12.
- 前川 喜久雄・北川 智利 (2002). 音声はパラ言語情報をいかに使えるか 認知科学, 9, 46–66.
- 前川 喜久雄・西川 賢哉 (2019). 『日本語話し言葉コーパス』への声質情報付与と予備的分析 言語資源活用ワークショップ 2019 発表論文集, 205–221.
- 丸本 徹・キャンベル ニック. (2000). 波形接続型音声合成方式における発話様式の制御 日本音響学会研究発表会講演論文集, 2000(1), 213-214.
- Mauss, I. B., Robinson, M. D. (2009). Measures of emotion: A review. *Cognition & Emotion*, 23, 209-237.
- 目加田 慶人・向笠 美幸・長谷川 光司・春日 正男・松本 修一・小池 淳 (1999). 音声信号に含まれる感情表現の分析 映像情報メディア学会誌, 53, 769–772.
- Menezes, C., Maekawa, K. (2006). Paralinguistic effects on voice quality: A study in Japanese. *Proceeding of Speech Prosody 2006*, 656-659.
- Meyer, J., Dentel, L., Meunier, F. (2017). Categorization of Natural Whistled Vowels by Naïve Listeners of Different Language Background. *Frontiers in Psychology*, 8, 1–12.
- Murray, I., 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, 1097-1108.
- Murray, I., Arnott, J. (1995). Implementation and testing of a system for producing emotion-by-rule in synthetic speech. *Speech Community*, 16, 369–390.
- 南 博 (1958). 初歩心理学：人間関係をよくする道具 東京：光文社.
- 宮島 崇浩・菊池 英明・白井 克彦・大川 茂樹 (2013). 演技指示の工夫が与える音声表現への影響——表現豊かな演技音声表現の獲得を目指して—— 音声研究, 17, 14–25.
- Mokhtari, P., Campbell, N. (2003). Automatic measurement of pressed/breathy phonation at acoustic centres of reliability in continuous speech. *IEICE Trans. on Inform. and Systems*, E-86-D, 3, 574-582.
- 森 大毅 (2014). 話し言葉が伝えるもの 国語研プロジェクトレビュー, 4, 183–190.
- 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—— コロナ社.

- 森 大毅 (2016). 音声研究と感情 「きれいな」理想と「きたない」現実. *エモーション・スタディーズ*, 2(1), 58-60.
- 森 大毅, 藤本 雅子, 浅井 拓也, 前川 喜久雄 (2017). 日本語話し言葉コーパスにおける発声様式の自動分類 国立国語研究所 言語資源活用ワークショップ発表論文集 1, 347-354.
- 森山 剛 (1999). 音声に含まれる感情情報と物理的特徴量に関する研究 慶應義塾大学大学院理工学研究科, 博士論文.
- 中林 律子 (2009). 問い返し疑問文に表れる『嫌』『驚き』の感情の知覚——ロシア語を母語とする学習者を対象として—— *言葉と文化*, 10, 165-180.
- 中林 律子 (2011). 日本語音声に表れる感情の知覚——ロシア人日本語学習者の知覚の個人差に着目して—— *音声研究*, 15, 14-25.
- 西端 千香子 (1996). 感情・態度を持つ音声の知覚・表出訓練のための CAI 用テキストの作成——促音が含まれる発話を中心に——感情・態度を表す日本語音声の表出診断・訓練プログラムの構築に関する研究, 平成7年度科学研究費補助金, 一般研究(B)研究成果報告書, 21-40.
- 西村 良太・北岡 教英・中川 聖一 (2009). 音声対話における韻律変化をもたらす要因分析 *音声研究*, 13, 66-84.
- 野田 哲矢・矢野 良和・道木 慎二・大熊 繁 (2005). KL 情報量に基づく音声感情認識に有効な韻律特徴の評価法 *音響学会 2005 年秋季研究発表会講演論文集 1*, 211-212.
- Nunes, A., Coimbra, R. L., Teixeira, A. (2010). Voice Quality of European Portuguese Emotional Speech. *International Conference on Computational Processing of the Portuguese Language*. 142-151.
- Ohala, J. (1996). Ethological theory and the expression of emotion in the voice, *International Conference on Spoken Language Processing 1996*, 3, 1812-1815.
- Pandit, P. (1957). Nasalization, aspiration and murmur in Gujarati. *Indian Linguistics*, 17, 165-72.
- Probst, L., Braun, A. (2019). The effects of emotional state on fundamental frequency. *International Congress of Phonetic Sciences ICPHS 2019*.
- Pell, M. D., Monetta, L., Paulmann, S., Kotz, S. A. (2009). Recognizing emotions in a foreign language. *Journal of Nonverbal Behavior*, 33, 107-120.

- Pell, M. D., Paulmann, S., Dara, C., Alaseri, A., Kotz, S. A. (2009). Factors in the recognition of vocally expressed emotions: A comparison of four languages. *Journal of Phonetics*, 37(4), 417-435.
- Rivière, E., Klein, M., Champagne-Lavau, M. (2018). Using context and prosody in irony understanding: Variability amongst individuals. *Journal of Pragmatics*, 138, 165-172.
- Sadanobu, T. (2004). A Natural History Japanese Voice of Pressed. *Journal of the Phonetic Society of Japan*, 8(1), 29-44.
- 定延 利之 (2007). 話し手は言語で感情・評価・態度を表して目的を達するか? : 日常の音声コミュニケーションから見えてくること 自然言語処理, 14(3), 3-15.
- 榊原 健一 (2014). 世界の歌唱法——様々な歌唱様式における suprenormal な声—— 音響学会誌, 70, 499-505.
- 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認知がしやすいか—— 信学技報, 20, 33-38.
- 櫻庭 京子・今泉 敏・笥 一彦 (2002). 音声による感情表現の発達の検討 音声言語医学 43, 1-8.
- 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.
- 佐野 智子 (2001). 音声を用いた感情表出と感情認知の研究 青山心理学研究, 1, 29-35.
- Scherer, K. R. (1986). Vocal affect expression: A review and a model for future research. *Psychological Bulletin*, 99, 143-165.
- Scherer, K. R. (1989). Vocal measurement of emotion. In: *Plutchik, R., Kellerman, H. (Eds.), Emotion: Theory, Research, and Experience*, 4, Academic Press, San Diego, 233-259.
- Scherer, K. R. (1999). Universality of emotional expression. In D. Levinson, J. J. Ponzetti, Jr., & Jorgenson (Eds.), *Encyclopedia of human emotions*. New York: Macmillan.
- Scherer, K. R. (2003). Vocal communication of emotion: A review of research paradigms. *Speech Communication*, 40, 227-256.
- 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.
- Scherer, K. R., Banse, R., Wallbott, H., Goldbeck, T. (1991). Vocal cues in emotion encoding and decoding. *Motivation and Emotion*, 15, 123-148.

- Scherer, K. R., Ladd, R. D., Silverman, K. E. A. (1984). Vocal cues to speaker affect: testing two models. *Journal of the Acoustical Society of America*, 76, 1346–1356.
- Scherer, K. R. (2013). Vocal markers of emotion: Comparing induction and acting elicitation. *Computer Speech & Language*, 27(1), 40-58.
- Scherer, K. R., Sundberg, J., Tamarit, L., Salomao, G. L. (2015). Comparing the acoustic expression of emotion in the speaking and the sing voice. *Computer Speech and Language*, 29, 218-235.
- 昇地 崇明・Véronique Aubergé・Albert Rilliard (2007). 発話態度の文化的特性と『偽の友達』——日仏語の対照研究を通して—— 定延 利之・中川 正之 (編) 音声文法の対照 (pp. 55–78) くろしお出版.
- 重野 純 (2001). 音声に含まれた感情の認知 青山心理学研究, 1, 37–43.
- 重野 純 (2004a). 音声と表情の間で矛盾している感情の認知——日米間の比較—— 青山心理学研究, 4, 17–26.
- 重野 純 (2004b). 感情を表現した音声の認知と音響的性質 心理学研究, 74, 540–546.
- 重野 純 (2005). 日本語及び英語音声における声質の認知 青山心理学研究, 5, 37–42.
- 重野 純 (2014). 感情音声の表出に及ぼす言語の影響 音声言語医学, 55, 233–238.
- Sundberg, J., Patel, S., Bjorkner, E., Scherer, K. R. (2011). Interdependencies among voice source parameters in emotional speech. *IEEE Transactions on Affective Computing*, 2(3), 162-174.
- 鈴木 朋子・田村 直良 (2006). 表現と認知の相違から検討した感情音声の特徴 心理学研究, 77, 149–156.
- 田川 恭識 (2001). 『平静の問いかけ』と『非難の問いかけ』の弁別に対する F0 パタンの影響 聴覚研究会資料, 31, 39–55.
- Tahon, M., Degottex, G., Devillers, L. (2012). Usual voice quality features and glottal features for emotional valence detection. *In Proceedings of Speech Prosody 2012*, 693–696.
- Teixeira, J. P., Oliveira, C., Lopes, C. (2013). Vocal acoustic analysis–jitter, shimmer and hnr parameters. *Procedia Technology*, 9, 1112-1122.
- 勅使河原 三保子・伊藤 克亘・武田 一哉 (2005). 日本のアニメの音声に表された感情と性格——声のステレオタイプの音声学的研究—— 信学技報, 14, 39–44.
- Thompson, W., Balkwill, L.L. (2006). Decoding speech prosody in five languages. *Semiotica*, 158, 407–424.

- Thongkum, T. (1988). Phonation types in Mon-Khmer languages. In O. Fujimura (Ed.), *Vocal fold physiology: voice production, mechanisms and functions*, 319-334. New York: Raven Press.
- Tickle, A. (2000). English and Japanese Speakers' Emotion Vocalization and Recognition: A Comparison Highlighting Vowel Quality. *ISCA Workshop on Speech and Emotion*. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.409.4169> (July 30, 2017.)
- Timcke, R., Von Leden, H., Moore, P. (1958). Laryngeal vibrations: Measurements of the glottis wave. *AMA Arch Otolaryngol*, 68(1), 1-19.
- Trager, G. L. (1958). Paralanguage: A first approximation. *Stud. Linguist*, 13, 1-12.
- Truesdale, D. M., Pell, M. D. (2018). The sound of Passion and Indifference. *Speech Communication*, 99, 124-134.
- 轟 真紀子・武田 昌一 (2007). アナウンサーが発声した『怒り』の韻律的特徴と聴覚的印象の主成分分析を用いた比較 久留美信愛女学院短期大学研究紀要, 30, 65-70.
- 轟 真紀子・武田 昌一 (2004). 感情表現音声の韻律的特徴——聴取実験による検討——久留美信愛女学院短期大学研究紀要, 27, 67-73.
- 内田 伸子・繁榘 算男・杉山 憲司 責任編集委員 (2013). 最新心理学事典 東京：平凡社.
- Van Bezooijen, R., Otto, S., Heenan, T. (1983). Recognition of vocal expression of emotion: A three-nation study to identify universal characteristics. *Journal of Cross-Cultural Psychology*, 14, 387-406.
- Vidrascu, L., Devillers, L. (2008). Anger detection performances based on prosodic and acoustic cues in several corpora. *Workshop on Corpora for Research on Emotion and Affect*. 13-16.
- 若狭 健太・寺澤 洋子・河原 英紀・榊原 健一 (2019). 生理・音響的特徴量分析によるオペラ歌唱と合唱歌唱の比較検討 日本音響学会春季研究発表会講演論文集, 1121-1124.
- Wilson, D., Wharton, T. (2006). Relevance and prosody. *Journal of Pragmatics*, 38, 1559-1579.
- Xiong, W., Cui, D., Meng, F., Cai, L. (2008). Analysis of Prosodic Features of Emotional Speech and the Experiment of Conversion. *The 12th Phonetic Conference of China (PCC 2016). International Conference on Celebrating the Centenary Birthday Speech of Wu Zongji at the Frontier of Speech Science*. Retrieved from <http://hcsi.cs.tsinghua.edu.cn/static/Paper/paper08/04C-PCC.pdf> (November 12, 2017.)
- Xu, Y. (2013). ProsodyPro——A Tool for Large-scale Systematic Prosody Analysis. In *Proceeding of Tools and Resources for the Analysis of Speech Prosody (TRASP 2013)*, 7-10.

- Yang, H., Cai, L., Jiang, D. (2005). Analysis on Expressivity and Acoustic Correlation of Speech. Retrieved from <http://hcsi.cs.tsinghua.edu.cn/static/Paper/paper05/200516.pdf> (November 12, 2017.)
- Yildirim, S., Bulut, M., Lee, C. M., Kazemzadeh, A., Deng, Z., Lee, S., Narayanan, S., Busso, C. (2004). An acoustic study of emotions expressed in speech. *In Eighth International Conference on Spoken Language Processing.*
- Yoshinaga, I., Kong, J. (2011). Some Phonatory Characteristics of Tibetan Buddhist Chants (< Feature Articles> Laryngeal Features among Eastern Asian Languages). *Journal of the Phonetic Society of Japan*, 15(2), 83-90.
- Yoon, K. (2011). A study on the phonation characteristics in Mandarin prosody. PhD Thesis, Peking University.

付表 1

日本語感情音声における音響的特徴のまとめ（中立との比較）

先行研究	刺激語	発話者	音響特徴	感情				
				喜び	怒り	悲しみ	驚き	嫌悪
郡 (1989)	母音 「お」	男性 1 名	F0	>		>	>	
			F0range	>		<	> ^①	
			F0 の変化 方向	≤ (急下降 調ではな い)		下降傾向 が強い	関係ない	
			Duration	>		≤		
飯田・伊 賀・安村 (1997)	「ありが とう」 「こんに ちは」 「ちばた つや」 「何言っ てんの」	男女各 1 名	F0range	>			>	
			F0 の変化 方向 ^②					
			Duration		<			
			Intensity mean	>			>	<
重野 (2001, 2004)	「東京」 「河原崎 さん」 「11 時 半」 「さ ような ら」 「そ うです か」	男性 2 名	F0mean	>	<		>	<
			F0range		>			
			Intensity mean		>		>	

Erickson et al. (2003) Erickson et al. (2006)	「バナ ナ」	女性 1 名	F0	>		<		
			F0 の変化 方向 ^③					
			Duration		<			
霧・武田 (2004, 2007)	無意味語 「まなま な」	女性 1 名	F0max		<(度合 い)			
鈴木・田 村 (2006)	「こっち へいらっ しゃい」 「えー」 「今日は 1 日」	女性 1 名	F0			<		
			F0range		>			
			Duration			>		
			音圧	>	>		>	
中林 (2009, 2011)	「韓国」 「毎日」 「現金」 「片仮 名」 「カラオ ケ」 「温 泉」	女性 11 名	F0range				>	
			Duration					>
			F0 の変化 方向 ^④					

付表 1 注 :

①変化幅が大きければ, 変化の方向はあまり関係なさそう。変化幅が大きくななくても, 高い音域であれば, 短い「お」や疑問上昇調の「お」は驚きに聞こえている。

②単語を平板型と起伏型に分けて上昇調と下降調について検討した。

③F0 曲線の形について, 「驚き」と「喜び」は似ている(上がって下がる)。「怒り」も同様。

④「嫌悪」においては、無核語では発話のいずれかの部分で F0 の上昇が極めて小さいまたは下降している部分が見られる。「驚き」においては、有核語の場合にはアクセント核による下降の直前の F0 が高いこと、無核語の場合には終点の F0 が高いこと。

付表 2

中国語感情音声における音響的特徴のまとめ（中立との比較）

先行研究	刺激語	発話者	音響特徴	感情				
				喜び	怒り	悲しみ	驚き	恐れ
Yang et al. (2005)	感情ごとに200刺激文, 平均で7~13音節(陳述文, 疑問文)	女性1人	F0mean	>	>	<	>	>
			F0 変化率 (Hz/ms)	>	>	<	>	
			Duration	<	<	>	<	<
			Intensity	>	>	<	>	
			High-frequency energy	>	>	<	>	
			Frequency range	>	>	<	>	
			Spectral noise (非周期)					>
Spectral centroid (质心)		>	<		<			
Xiong et al. (2008)	5・6音節刺激文 48個	女性1人	F0mean	>	>		>	
			Duration			>		
			Intensity			<		
			句頭・句中・句末 ^①					

Liu (2011)	30 刺激文	男女各 8 人	F0mean		>			
			F0max		>		>	
			F0min		>		>	
			F0 変化率 (Hz/ms)		>(特定の 音節で急 激な上昇)			
			F0range			<	>	
			F0 の変化 方向 ^②					
			Duration	<	<	>	>	
			Pause	< (回数 も時間 も)		> (回数 も時間 も)		

付表 2 注 :

- ① 句頭の特徴：「怒り」のインテンシティは「喜び」と「驚き」より遥かに大きい；「喜び」の持続時間は「怒り」と「驚き」より長く，F0range は狭い。句中の特徴：F0max と F0min から見ると，「驚き」は混同されやすい「怒り」「喜び」と区別できる。句末の特徴：「驚き」の F0range は「中立」の 2.2 倍にあたるほど広い；「怒り」の持続時間は明らかに「喜び」より短い。また，句末の特徴が一番目立った。
- ② 「驚き」のピッチ曲線の上下起伏が激しく，特に句末のところで急激な上昇が見られ，ピッチももつとも高い。「驚き」のピッチ曲線は「怒り」と類似しており，フォーカスのあるところにピッチが一番高い。「怒り」と「喜び」のピッチ曲線は句末にて急激な下降が観察された。「悲しみ」の句末は緩やかな傾向があるが，「驚き」は激しく上昇する。「喜び」と「驚き」のピッチ曲線は似ており，混同されやすい原因になる。

付録 1

第 5 章生成実験収録に用いたシナリオ：

(1) 「え」

まず、教科書を音読するような口調で読み上げてください。

(中立：え)

①喜び

あなたの友だちが近況について話しています。

A：2年かけてやっと志望大学に合格した！

B：（喜び：え）！おめでとう！

②激しい怒り hot anger

あなたの部屋を借りた人がパーティーをして、部屋をめちゃくちゃにしました。あなたの友だちが電話をかけてきました。

A：あの人たちがソファを汚して、テレビも壊した！

B：（怒り：え）！ひどすぎる！

③押し殺した怒り cold anger

同僚が遅刻したために、会議が長引いています。あなたは大事なデートに遅れそうです。

A：さっき彼から電話があって、また 30 分遅れるって。

B：（押し殺した怒り：え）…！また 30 分遅れるなんて！

④悲しみ

あなたが子どものころよく一緒に遊んでくれたおじいちゃんが亡くなったことを親戚から聞きました。

A：おじいちゃんが先週亡くなりました…

B：（悲しみ：え）…おじいちゃんが…。

⑤驚き

あなたの友達があなたに話します。

A：昨日空港で君が好きなあの俳優を見かけたんだ！

B：（驚き：え）！どこの空港？

⑥恐れ

あなたが乗っている飛行機が急に激しく揺れ始めました。隣の友達がパニックになってあなたに話しかけます。

A：エンジンからけむりが出てる！

B：（恐れ：え）…私たちここで死ぬの？

⑦嫌悪

あなたと同僚が誰かが吐いた汚物を片付けなければなりません。

A：30分以内でここを片付けなければなりませんね。

B：（嫌悪：え）…見たくもないですね…

(2)「そうですか」

まず、教科書を音読するような口調で読み上げてください。

（中立：そうですか）

①喜び

新入社員のあなたのために、同じチームの人たちが今夜歓迎会を開こうとしています。ある同僚があなたに話しかけます。

A：今日、仕事のあと、地元の美味しいお店であなたの歓迎会を開こうと思っています。

B：（喜び：そうですか）！ありがとうございます！

②激しい怒り hot anger

あなたが友だちと会話していると、ある知り合いに何度も話の邪魔をされました。友だちもあなたもとてもうんざりしています。あなたは次回のお出かけについて、近くの川に釣りに行こうと話しているとき、またその知り合いに邪魔をされました。

A：ダメ！その川全然釣れないらしいよ。

B：（怒り：そうですか）！聞いてませんが！

③押し殺した怒り cold anger

2ヶ月かけて筆記試験と面接を突破して、やっと手に入りそうなポジションがコネのある人に取りられました。あなたははじめてこの情報を友達から聞きました。

A：あなたより面接の成績が低かった人が選ばれたって！

B：（押し殺した怒り：そうですか）…よりによって…！

④悲しみ

引越しのため、長年一緒にいたペットとお別れしなければなりません。新しいマンションの管理人さんがあなたに言います。

A：申し訳ありませんが、このマンションはペット禁止です。

B：（悲しみ：そうですか）…わかりました…

⑤驚き

あなたがフィールドリサーチで遠い田舎の村に行ったら、地元の人に言われました。

A：ここでは、夜になるとあの珍しい動物が見られますよ。

B：（驚き：そうですか）?!信じられません！

⑥恐れ

あなたと先輩は山奥の別荘で部活の合宿に参加しています。ある日の夜、あなたと先輩は止むを得ず暗い森を通過して別荘に帰らなければなりません。

A：この森、熊がよく出るらしいよ…

B：（恐れ：そうですか）…出ないように祈るしかないですね…

⑦嫌悪

あなたのルームメイトが冷蔵庫を開けて、消費期限切れのヨーグルトを取り出してあなたに話しかけます。

A：これ臭くないですか？

B：（嫌悪：そうですか）…くさっ！

(3)「オバマが当選した」

まず、教科書を音読するような口調で読み上げてください。

（中立：オバマが当選した）

①喜び

あなたはスミス大統領候補者を支持しています。彼の当選を心の底から期待しています。選挙結果を知ったあなたは飛び上がりそうになりました。

A：ニュースを観てください！

B：（喜び：オバマが当選した）！素晴らしい！

②激しい怒り hot anger

あなたはスミス大統領候補者の最大のスポンサーであり、多大な財産を使いました。しかし、オバマという名の候補者が当選しました。

A：ニュースを観てください！

B：（激しい怒り：オバマが当選した）！これまでの資金は全部無駄遣いだったか！

②押し殺した怒り cold anger

あなたはスミス大統領候補者の最大のスポンサーであり、多大な財産を使いました。しかし、オバマという名の候補者が当選しました。

A：ニュースを観てください！

B：（押し殺した怒り：オバマが当選した）！これまでの資金は全部無駄遣いだったか！

④悲しみ

冷血な移民政策で有名な大統領候補者が当選して、すぐその政策を出しました。あなたは移民であり、あなたとあなたの家族は長く住んでいる土地を離れなければなくなりました。

A：ニュースを観てください…

B：（悲しみ：オバマが当選した）…この町が大好きなのに…

⑤驚き

あなたは今回当選した候補者に注目したこともなければ、期待もなく、名前すら覚えていません。

A：ニュースを観てください！

B：（驚き：オバマが当選した）?!誰この人？

⑥恐れ

あなたは政治家として候補者を中傷する多くの発言をしました。今回の選挙で大統領に当選したその候補者があなたに復讐しにくると、あなたは判断しました。

A：ニュースを観てください！

B：（恐れ：オバマが当選した）…!大変なことになった!

⑦嫌悪

あなたはどうしてもあの候補者が嫌いです。

A：ニュースを観てください！

B：（嫌悪：オバマが当選した）!あんな男でも大統領になれるなんて、この国も終わりだ。

中国語シナリオ：

(1) “啊”

请像读教科书的单词表一样，无感情地读。

(中立：啊)

①高兴

你的弟弟跟你说起近况。

A：我考了2年，终于成功考上理想的大学啦！

B：（高兴：啊）!太棒啦！

②爆发式的愤怒 hot anger

租借你房屋的租户，在你的房屋里举行了聚会，把房屋弄得乌烟瘴气。你的妻子给你打电话。

A：他们把沙发弄的很脏，电视也坏掉了。

B：（爆发式的愤怒：啊）！欺人太甚！

③压抑的愤怒 cold anger

因为同事的又一次迟到，今天的会议拖了很久还没有结束，为此你不得已要取消一个重要的约会。

A：他刚刚来电话说还要再晚三十分钟。

B：（压抑的愤怒：啊）…！还要晚三十分钟！

④悲伤

你从亲戚那里听说，儿时经常陪伴自己玩耍的爷爷过世的消息。

A：听说爷爷走得很安详…

B：（悲伤：啊）…爷爷…

⑤惊奇

你的朋友对你说。

A：昨天我在机场看到你喜欢的那个演员了！

B：（惊奇：啊）！在哪个机场？！

⑥恐慌

你在飞机上，突然感觉飞机开始剧烈颠簸，这时候旁边的朋友很惊慌地跟你说。

A：飞机的左翼开始冒烟了！

B：（恐慌：啊）…，咱们要死在这里了吗…！

⑦嫌恶

你的室友打开冰箱，发现了一盒已经过期的酸奶，室友把它取了出来。

A：这盒酸奶已经臭了，你闻闻。

B：（嫌恶：啊）…臭死了！

(2) “是吗”

请像读教科书的单词表一样，无感情地读。

(中立：是吗)

①高兴

新入职的公司的同事们商量着为你开一个欢迎会，其中一个跟你年纪相近的同事来跟你说。

A：今天下班之后我们带你去吃一家当地的好馆子。

B：（高兴：是吗）！谢谢大家！

②爆发式的愤怒 hot anger

你和朋友正在愉快的聊天，突然来了个人三番五次的打断你们。你和你的朋友都很烦躁。正当你们聊到下次一起去河边钓鱼的事情，又被旁边的人打断了。

A：不行不行，那条河根本就钓不到鱼。

B：（爆发式的愤怒：是吗）！我们又没问你！

③压抑的愤怒 cold anger

你花了2个月的时间参加一家心仪公司的笔试和面试，眼看终于要拿到 offer 的时候，被一个走后门的家伙抢走了工作机会。你第一次听到这个消息是从朋友那里。

A：听说是那个面试成绩比你差很多的那个人入职了。

B：（压抑着愤怒：是吗）…！偏偏是他！

④悲伤

因为搬家，你不得不和在一起很久的宠物分别。房东对你说。

A：非常抱歉，这里的公寓禁止养宠物。

B：（悲伤：是吗）…，没想到会有一天要和它分别…

⑤惊奇

你要去一个偏僻的山村进行田野调查工作。到了那里之后，当地人对你说。

A：在我们这里一到晚上，可以看到一种很稀有的动物。

B：（惊奇：是吗）！那我可要看看！

⑥恐慌

你在山中的训练场进行集训。某天夜里，你和前辈不得不需要通过一条黑暗的森林小路返回集训地。

A：听说这片森林经常能碰到熊！

B：（恐慌：是吗）…，只能祈祷今天不要碰到了…

⑦嫌恶

你和同事不得不收拾一滩呕吐物。

A：30分钟之内我们要把这里清理干净。

B：（嫌恶：是吗）…，看都不想看一眼。

(3) “奥巴马当选了”

请像读教科书的单词表一样，无感情地读。

（中立：奥巴马当选了）

①高兴

你是支持那个候选人的选民，热切盼望他能当上总统，当你知道了竞选结果，兴奋得跳了起来。

A：你快看新闻！

B：（高兴：奥巴马当选了）！太棒了！

②爆发式的愤怒 hot anger

你是一名总统候选人的最大的赞助商，付出了巨大精力和财力，明明一直是优势，却在最后一刻败落给另一名候选人。

A：您快看新闻！

B1：（爆发式的愤怒：奥巴马当选了）！咱们的资金就这么打水漂了！

③爆发式的愤怒 hot anger

你是一名总统候选人的最大的赞助商，付出了巨大精力和财力，明明一直是优势，却在最后一刻败落给另一名候选人。

A：您快看新闻！

B：（压抑的愤怒：奥巴马当选了）…！咱们的资金就这么打水漂了！

④悲伤

一名总统候选人提出驱逐移民的强硬政策，使你和你的家人必须离开居住了多年的家园并且要和那里的老朋友们分别。

A：你快看新闻，咱们真的要离开这片土地了…

B：（悲伤：奥巴马当选了）…我明明很喜欢这片土地…

⑤惊奇

你觉得这名候选人无论如何也不会当选，根本没有给予关注。甚至连名字都没上心过。

A：你快看新闻！

B：（惊奇：奥巴马当选了）！这人是谁？！

⑥恐慌

你作为政客发表了很多严重攻击诋毁一名候选人的言论，谁知那名候选人当选了。这意味着他马上就会来铲除你了。

A：你快看新闻！

B：（恐慌：奥巴马当选了）…！快给我买出国的机票！快！

⑦嫌恶

你无比反感成为总统的那个候选人，觉得国家不能由这样的人领导。

A：你快看新闻！

B：（嫌恶：奥巴马当选了）！这种货色也能当总统，这个国家算是完了！

謝辞

本研究を遂行するあたり、ご指導を賜りました皆様に心より感謝の意を表します。指導教員の林良子先生には本研究の実施の機会を与えていただき、その遂行にあたって終始、ご指導をいただきました。心より厚く御礼申し上げます。論文の指導だけでなく、多くの学会参加や発表の機会を与えていただきました。厳しいご指導と優しい励ましにより、私自身の至らなさを実感できたことは、今後の努力の糧になるものです。本当にありがとうございました。

本論文の作成にあたり、大変貴重なご指導とご助言をいただいた株式会社国際電気通信基礎技術研究所（ATR）の石井カルロス寿憲先生に心より感謝申し上げます。石井先生からは実験計画や分析方法といった研究の基礎のみならず、研究に対する取り組み姿勢や考え方など多くのことを学びました。そして、温かく見守っていただくとともに、多くのご支援ご指導を賜りました山田玲子先生には深く感謝しております。お忙しい中お時間を割いていただきありがとうございます。

Aaron Lee Albin 先生からいただいた数多くのご助言によって本研究を発展させることができました。ここに深甚の謝意を表します。また、学内での発表のたびに有益なコメントをいただいた国際文化科学研究科の先生方に拝謝いたします。

感性コミュニケーションコースの皆さまには日頃より有益なご討論ご助言をいただきました。ここに感謝の意を表します。ATR カルロス研の皆さまには日頃から多岐にわたる相談に乗っていただきました。ありがとうございました。

これまで自分の思う道を進むことに対し、温かく見守りそして辛抱強く支援してくれた両親に深く感謝します。特に、どのような状況においても応援してくれた素晴らしい夫・川島朋也さんに心から感謝します。

最後に、本実験に参加いただいた実験参加者の皆さまに御礼申し上げます。