

中国語を母語とする日本語学習者による感情音声の知覚

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

1. はじめに

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

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

2. 先行研究

2.1 日本語母語話者による感情音声の産出・知覚

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

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

2.2 非日本語母語話者による感情音声の産出・知覚

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

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

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

3. 日本語母語話者による判定実験

3.1 刺激音の収録

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

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

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

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

3.2 刺激音の評価と選択

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

4. 中国人日本語学習者を対象とした聴取実験

4.1 聴取実験

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

4.2 実験結果：中国人日本語学習者の知覚

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

た。

表 1 各学年の感情知覚 confusion matrix (%)

1年生	喜び	怒り	悲しみ	平静	2年生	喜び	怒り	悲しみ	平静
喜び	36	27	16	21	喜び	45	27	5	22
怒り	17	51	15	17	怒り	19	54	13	14
悲しみ	8	15	44	33	悲しみ	5	8	52	35
平静	18	7	22	53	平静	9	6	16	69
3年生	喜び	怒り	悲しみ	平静	院生	喜び	怒り	悲しみ	平静
喜び	55	24	6	15	喜び	59	29	5	7
怒り	12	69	7	12	怒り	17	69	8	6
悲しみ	2	3	64	31	悲しみ	1	3	65	31
平静	1	6	12	81	平静	1	2	14	83

表 1 の縦欄は刺激語の感情、横欄は回答された感情を示す。

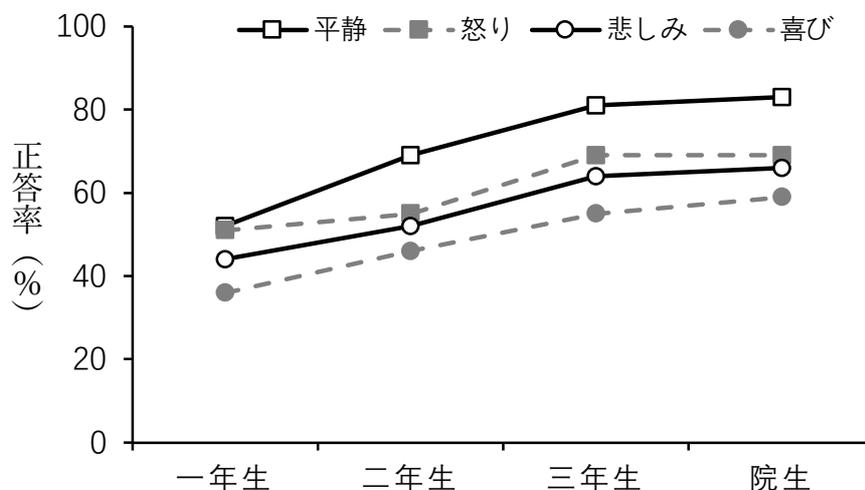


図 1 平均正答率

4.3 日本語感情音声の韻律的特徴

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

表 2 「小野さん」の音響的特徴

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

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

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

5. まとめ

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

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

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

謝辞

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

引用文献

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