P08

中国語を母語とする日本語学習者による感情音声の知覚と生成の関係

李　歆玥（神戸大学大学院）Aaron ALBIN・林　良子（神戸大学）

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

rhayashi@kobe-u.ac.jp

1. 背景と目的

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

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

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

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

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

2. 音響分析

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

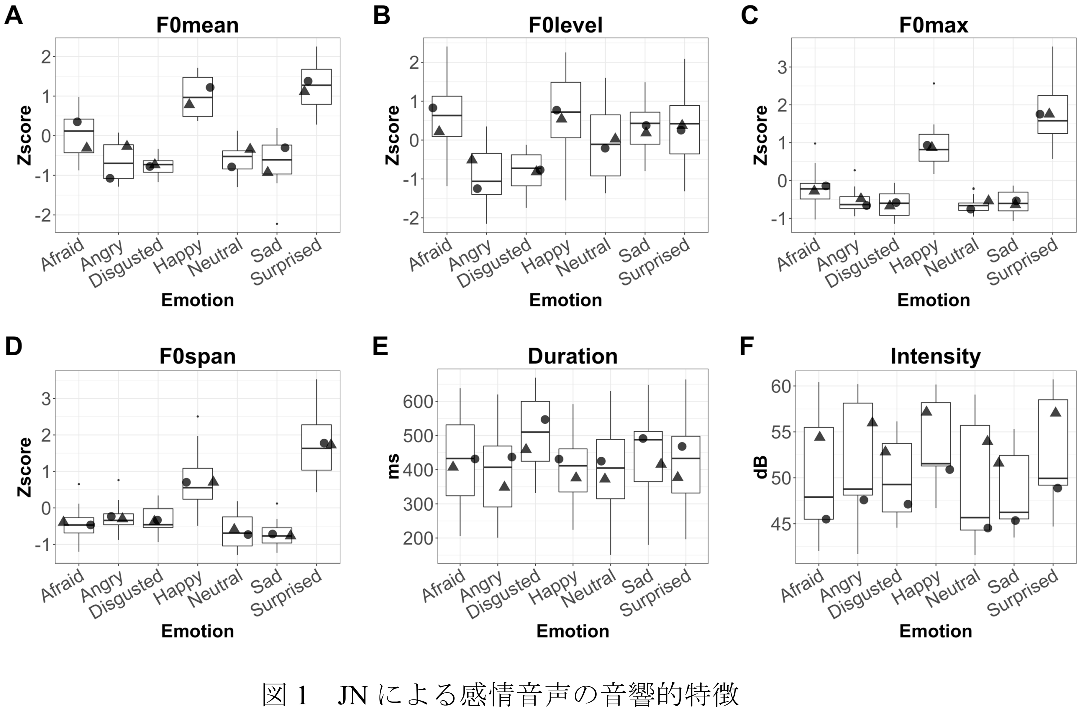

図 1　JN による感情音声の音響的特徴

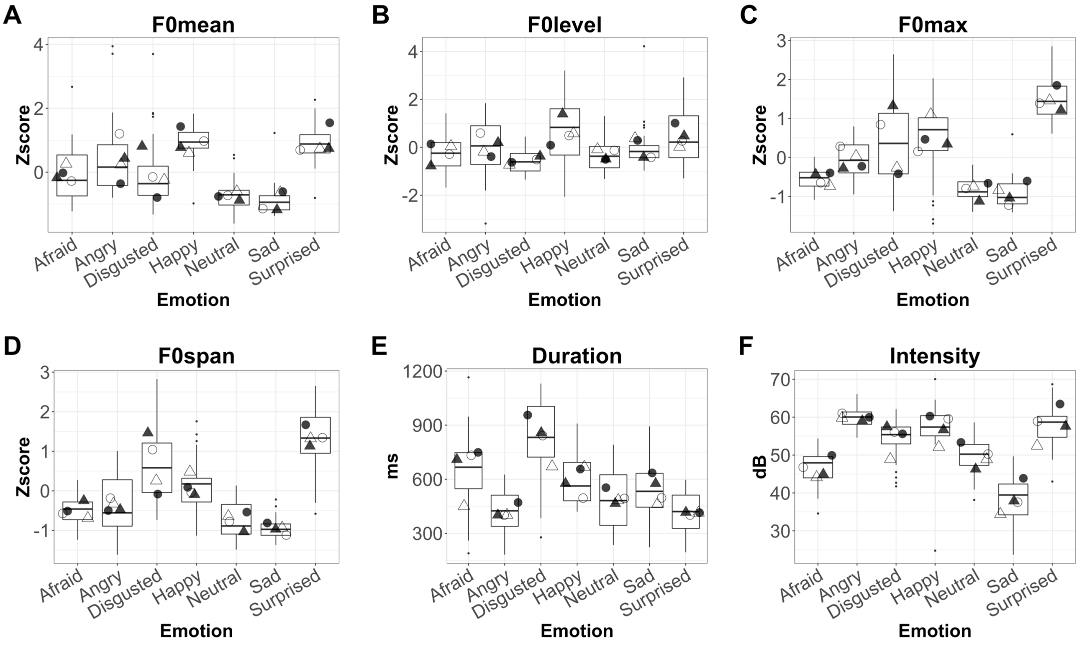

図 2　CJ による感情音声の音響的特徴

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

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

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

的特徴の違いが CJ に比べて顕著ではなかったためだと考えられる。

表 1　各音響的特徴における多重比較の結果

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

*$p < .05$

3. 生成と知覚の関係

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

CJ による発話では，「悲しみ」と「恐れ」に関して，「悲しみ」の同定率は F0mean が低いと高くなり（$r(132) = -.47, p < .001$）, Intensity が弱いと高くなる（$r(132) = -.38, p < .01$）。「恐れ」の同定率は F0mean が高いと高くなり（$r(132) = .57, p < .01$），Intensity が強いと高くなる（$r(132) = .23, p < .01$）。また，「喜び」と「驚き」に関して，「喜び」は Duration が長いと同定率が高くなる（$r(132) = .62, p < .001$）。さらに，「怒り」と「嫌悪」に関して，「怒り」は F0span が狭いほど同定率が高い（$r(132) = -.32, p < .05$）ことが観察された。

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

表 2　音響的特徴と同定率の相関（*r*）

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

$*p < .05$，$**p < .01$，$***p < .001$

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

4. おわりに

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

謝辞

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

参考文献

Ekman, P. (1992). An Argument for Basic Emotions. Cogniton And Emotion, 6, 169–200.

Erickson, D. (2005). Expressive Speech: Production, perception and application to speech synthesis. Acoust. Sci. & Tech, 26, 317–325.

エリクソン ドナ ・昇地 崇明 (2006). 性差，および母語が感情音声の知覚に与える影響—日本語，韓国語，英語母語話者を対象として— 音声文法研究会（編）音声文法（pp. 31–46）くろしお出版.

甲斐 朋子・田渕 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.

李 歆玥・Aaron Albin・林 良子 (2018). 中国語を母語とする学習者による日本語感情音声の知覚—母語話者および学習者による発話を対象に— 音声学会第 337 回例会要旨.

櫻庭 京子・今泉 敏・筧 一彦 (2001). 感情判断と言語判断—母語は外国語より感情知がしやすいか— 信学技報，20，33–38.

櫻庭 京子・今泉 敏・筧 一彦 (2004). 「ピカチュウ」にこめられた感性情報 音声研究，8，77–84.

Scherer, K. R., Banse, R., & Wallbott, H. (2001). Emotion Inferences From Vocal Expression Correlate Across Languages and Cultures. Journal of Cross-Cultural Psychology, 32, 76–92.