

音声による感情表出とその音響的特徴について —問い返し疑問文に表れる「嫌」「驚き」の感情を例として—

中 林 律 子

1 はじめに

音声は言語的情報だけではなく、発話意図、感情といったパラ言語的情報も伝達する。感情に関する情報の伝達は、音声の F0、強さ、持続時間、フォルマントなどの変化により実現する。従って、音声によって感情を伝達・受容するためには、これらの音響的パラメータの変化を知覚し、かつ産出することが必要となる。しかし、日本語を学習する非日本語母語話者にとっては音声からの感情の伝達・受容が大きな問題となりうる。感情を伝える音声はある程度普遍性があると言われているが(Ladd, 1980)、非日本語母語話者が日本語音声から感情を推測する際、日本語母語話者と異なる推測を行う場合のあることが Campbell & Erickson(2004)、エリクソン・昇地(2006)において指摘されている。しかし、非日本語母語話者が日本語音声から感情を推測する際、日本語音声のどの音響的パラメータが手がかりとなっているのか、またはなりにくいのかについては十分に明らかになっていない。これらを明らかにするためには、まず、日本語音声による感情表出が音響的にどのように実現しているのかを明らかにする必要がある。本研究では問い返し疑問文を用い、話者が感情を込めて発話した場合の音響的特徴を分析する。

問い返し疑問文は上昇イントネーションを伴って発話される疑問文の一種であり(郡1997、上村1989)、「相手のことばそのままか、それに近い形を繰り返して、それを相手に確かめるもの(南1985)」であるが、森山(1989)は相手の発話に対して否定的である場合にも問い返し疑問文が用いられることを指摘している。さらに、近藤(2001)は、相手の発話内容の意外性が高い場合、その発話内容に対する驚きを表すのに問い返し疑問文が用いられることを述べている。これらのことから、問い返し疑問文は、勧誘や依頼といった相手からの働きかけの発話に対して用いられる場合、相手の発話内容に対

中林律子

して否定的である場合には「嫌(いや)」という感情¹を、相手の発話内容が意外性の高いものである場合には「驚き」という感情を表出する働きを持つと考えられる。日本語学習者が問い返し疑問文からこのような感情を推測できない場合、例えば続けて依頼や勧誘を行ってしまい「押しの強い人」といった印象を相手に与えてしまうなど、コミュニケーションに支障が生じる可能性がある。

2 先行研究

日本語音声感情を伝達する際の音響的特徴を扱った研究には、北原・東倉(1988)、郡(1989)、西端他(1996)、重野(2004)が挙げられる。このうち、郡(1989)、西端他(1996)は単音節を対象とし、感情を伝達する音声の音響的特徴を分析している。

郡(1989)は単音「お」の F0・持続時間長を段階的に操作した合成音を作成し、「喜び」「恐怖」「驚き」「悲しみ」「嫌悪」「怒り」の感情をどの程度表しているかを日本語母語話者12名に5段階で評価させた。その結果、「驚き」「悲しみ」「喜び」の判定には F0及び持続時間長の変化が大きく関与していることが明らかになった。一方、「恐怖」「嫌悪」「怒り」の感情を表すと判定された音声が少ないことから、これらの感情の表出には F0・持続時間長以外の音響的パラメータが関与している可能性を示唆している。

西端他(1996)は日本語母語話者4名が「中立」「喜び」「驚き」「落胆」「疑い」の5種類の感情・心的態度を込めて発話した間投詞「あ」「え」の音響的特徴を分析した。それぞれの感情・心的態度については、「強い驚き」「弱い驚き」など、感情・心的態度の強弱についても表現し分けさせている。F0パタン・F0変動幅・パワー・持続時間長について分析が行われた結果、感情・心的態度の種類だけでなく、それらの強弱の違いも主に F0パタンによって表されていることが明らかにされている。

単語及び文レベルでの音声による感情伝達を扱った研究としては北原・東倉(1988)、

¹ 感情の生起に認知プロセスが関わるとする立場では、人がある状況や出来事に対しどのような感情を抱くかは、感情を引き起こす元となる対象をどのように評価するかにより決定するとしている(コーネリアス 1999)。土田(1996)は、感情は対象や状況に対し望ましいと思う「快感情」と、望ましくないと思う「不快感情」とに大別され、「良い」「好きだ」など正の評価を表す言語表現は「その評価の対象を受容し、それに接近していく」という心理状態を、「悪い」「嫌いだ」など負の評価を表す言語表現は「その評価の対象を拒否し、回避する」という心理状態を表すとしている。本研究では「嫌」という感情を「相手からの働きかけの内容が自分にとって望ましくないため、拒否・回避したい」と感じる感情とした。

重野(2004)がある。北原・東倉(1988)では「船まで迎えに行きますよ」という一文が「平静」「怒り」「歓喜」「悲哀」の4種類の感情を伴って発話された発話資料を収集し、「平静」の発話に、「怒り」「歓喜」「悲哀」の発話の F0、パワー、持続時間長の情報を付与した上で日本語母語話者11名に感情を判定させた。その結果、「怒り」の判定には持続時間長が、「歓喜」「悲哀」の判定にはF0が重要な手がかりとなっていることが明らかになっている。

重野(2004)は「東京」「さようなら」など単語及び短文を用い、「幸福」「驚き」「怒り」「嫌悪」「恐れ」「悲しみ」の感情を表していると判定された日本語母語話者2名による音声の音響的特徴を分析した。その結果、「快感情」と判定された発話は平均 F0が高く、「不快感情」と判定された発話は平均 F0が低いことを明らかにしている。

これらの研究から、日本語音声による感情表出には F0及び持続時間長が大きく関わっていると考えられる。しかし、例えば F0の変動幅が広いという場合に、F0の上昇・下降が語中または文中のどの位置で顕著に生じているのかなどについては、これまでの研究ではほとんど言及されていない。感情による F0の変化はアクセント型によって多様であると予想されることから、異なるアクセント型間で感情表出の実現にどのような共通点・相違点があるのかについての分析が必要であると考えられる。前川・北川(2002)ではパラ言語的情報による音響的特徴の変化は主に句頭及び句末に顕著に現れることが指摘されているが、感情の表出に焦点を当てた研究において句頭・句末の変化に着目した分析はほとんど行われていない。音声による感情表出について明らかにするためには、句頭・句末の変化に着目した分析も必要だと思われる。郡(1989)、西端他(1996)、重野(2004)では「驚き」の感情が扱われている。意外性の高い事態の内容によっては好意的な驚きも否定的な驚きも生起し得るが、上記の研究においては好意的な驚きか否定的な驚きかといった区別はなされていない。「嫌」という感情は相手の発話内容の意外性の高低に関わらず生起すると考えられる。本稿では、話し手が驚いているかどうかに関わらず「嫌」という感情が音響的にどのように実現するのか、また、「驚き」という感情が加わった場合とそうでない場合とでは「嫌」という感情を表出する音響的特徴がどのように異なっているのかをアクセント型別に明らかにする。

3 方法

3-1 音声資料

本研究では、アクセント型の違いにより感情表出にどのような共通点・相違点があるかについて検討するため、「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つの語を選び、² これらが表1の①～④として発話される24のダイアログ(6種類の語×4種類の問い返し疑問文)を作成した(巻末資料参照)。

表1 分析対象とする発話意図

①	意外性の低い発話に対し、単に確認を行っただけの発話	驚いていない・嫌がっていない
②	意外性が低く、望ましくない発話に対して嫌だという感情を表す発話	驚いていない・嫌がっている
③	意外性の高い発話に対して驚きを表す発話	驚いている・嫌がっていない
④	意外性が高く、望ましくない発話に対して驚き、嫌だという感情を表す発話	驚いている・嫌がっている

音声資料提供者は20代～40代の女性で、演劇・アナウンスの経験者、日本語教師、音声学を専門とする大学院生の計11名である。対話者は女性アナウンサーに依頼した。録音では、音声資料提供者と対話者として一つのダイアログにつき3回演じるよう依頼した。その際、森山(1990)では、断りを行う際、親しくない間柄では直接的な表現を避ける傾向があることを明らかにしていることから、親しい間柄の方が「嫌」等の感情も表出されやすいと考え、ダイアログ中の人物設定を親しい友人同士とした。

収録では、問い返し疑問文の文末が非上昇となったり、言い間違いが起こったりする場合があった。文末が非上昇となった場合、その発話は典型的な「問い返し疑問文」ではなく、感情による文末のF0の変動に一定の傾向を見出しにくくなる恐れもある。また、言い間違いのある発話は持続時間長の計測に影響するだけでなく、「か、韓国」などのような言い淀みはそれ自身が「恐れ」「躊躇」などといった感情や心的態度を伝達する可能性がある。そのため、これらの発話が生じた場合には音声資料提供者に演じ直させた。音声資料提供者が演じ直しを要求した発話も含め、音声資料として用いた発話は

² 資料語は初級レベルの学習者にとって既習の語であること、外来語でないこと、長音・促音が含まれていないこと、ダイアログを作成しやすい語であることを基準として選定した。

828例である(24のダイアログ×11名の音声資料提供者×3回=792例に、音声資料提供者が演じ直しを要求した発話36例を追加)。

3-2 聴取実験

感情を適切に表出している音声資料のみを音響分析の対象とするために、名古屋・東京の2箇所での20代・30代の日本語母語話者を対象として聴取実験を行い、音声資料の選別を行った。³ まず名古屋市及びその近郊で生育した日本語母語話者39名を対象として聴取実験を行った。収集した828例の音声資料を話者・感情共にランダムにし、各判定者に約128例を判定させ、一つの音声資料につき6名が判定を行うようにした。まず、約128例の音声資料を2回ずつ聞き「驚いているか驚いていないか」を強制選択させ、その後、もう一度同じ約128例の音声資料を2回ずつ聞き「嫌がっているか嫌がっていないか」を強制選択させた。聴取実験はなるべく騒音のない室内を選び、基本的に3名ずつで行った。説明等の時間も含め、実験に要した時間は約45分である。音声資料の呈示方法を図1に示す。

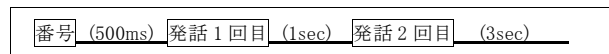


図1 音声資料の呈示方法

聴取実験の結果、「驚いているかどうか」「嫌がっているかどうか」の両項目で6名中5名以上の判定が一致した音声資料(判定一致率83.3%以上)⁴ は332例であった。これら332例を話者・感情共に再度ランダムにし、東京及びその近郊で生育した日本語母語話者18名を対象として再び聴取実験を行った。各判定者に約110例を判定させ、一つの音声資料につき6名が判定を行うようにした。聴取実験は名古屋での実験と同様、なるべく騒音のない室内を選び、基本的に3名ずつで行った。説明等の時間も含め、実験に要した時間は約40分である。その結果、「驚いているかどうか」「嫌がっているかどうか」それぞれの項目で6名中5名以上の判定が一致した音声資料は203例であった(表2)。

³ 本来ならば東京1箇所のみで実験を行うのが理想的であるが、本研究者が東京近郊在住でないことから、時間的にも人数的にも東京のみで実験を行うのが困難であった。そのため、本研究者の居住地に近い名古屋及びその近郊での第1回目の聴取実験で発話意図がより適切に実現されている発話を選定し、その後、判定一致度の高かった発話について改めて東京及びその近郊で実験を行うこととした。

⁴ 6名中5名以上の判定が一致した発話は、発話意図が適切に実現されているものと判断した。

中林律子

表 2 名古屋・東京での聴取実験で 6 名中 5 名以上の判定が一致した発話(一致率 83.3%以上)

アクセント型	頭高型		中高型		平板型		計
語	韓国	毎日	現金	片仮名	カラオケ	温泉	
①驚いていない・嫌がっていない	25	16	4	22	23	15	105
②驚いていない・嫌がっている	9	6	2	6	11	7	41
③驚いている・嫌がっていない	7	0	4	8	4	6	29
④驚いている・嫌がっている	2	12	4	5	4	1	28
計	43	34	14	41	42	29	203

「毎日」では、名古屋・東京の2箇所では6名中5名以上が③「驚いている・嫌がっていない」と判定した発話がなく、「温泉」の④など判定の一致した発話が少ない語もあった。判定実験において判定者から「「現金」を嫌がる場面が想像しにくい」などの意見があったことから、判定者が語に持っているイメージが「嫌がっているかどうか」「驚いているかどうか」の判定に影響した可能性が考えられる。

3-3 分析項目

音声資料はサンプリング周波数44.1kHz、16bit で量子化してパソコンに取り込み、音声解析ソフトpraatにより分析を行った。図2は頭高型語「韓国」、中高型語「現金」、平板型語「カラオケ」のF0曲線である。

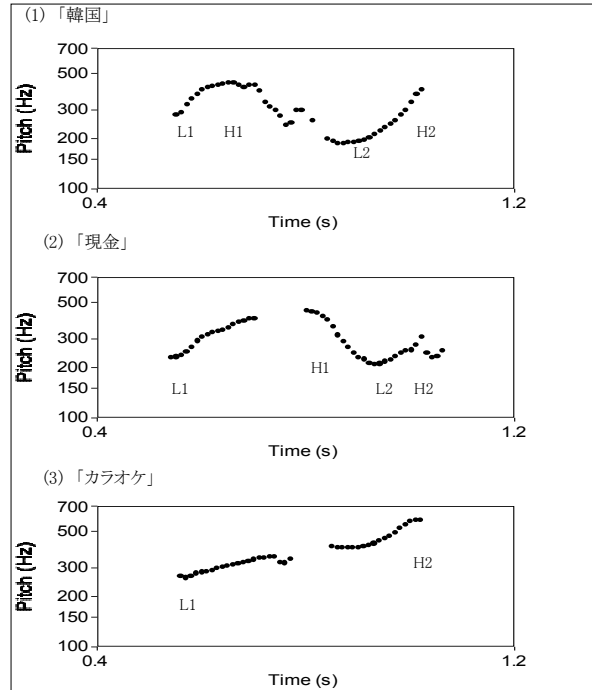


図2 praat による F0曲線

分析項目は持続時間長とF0である。

持続時間長については、発話全体及び各拍の計測を行った。持続時間長は音声波形の振幅(最大振幅時の70%以下は切り捨て)及び広域帯スペクトログラムで高次フォルマントの開始点・終了点を基準として計測した。子音の計測基準は Peterson and Lehiste(1960)、母音の計測基準は鹿島(1992)を主に参照したが、母音の連続部分など計測が困難な部分については、聴覚印象及び第1フォルマント・第2フォルマントの推移から、後続音が安定したと見られる部分を境界とした。

F0については、Hz で計測し、100Hz ベースの半音値(st)に変換した。⁵

有核語の F0は、発話の始点(以下「L1」)、アクセント核による下降が開始する直前の

⁵ Hz から st への変換は、Fant et al.(2002)を参考にし、次の式により行った。

$$st = 12[\log(x/100)/\log 2]$$

中林律子

部分において最も F0が高い点(以下「H1」)、語末の上昇が開始する直前の部分において最も F0が低い点(以下「L2」)、発話の終点(以下「H2」)の4点を計測した。土岐(1998)は頭高型語について、発話開始時に際立った声門閉鎖が伴わない限り語頭に短い上昇が起こることを指摘しており、本研究においても頭高型語の音声資料に語頭での上昇が見られる(図2(1)参照)。これら頭高型語の上昇部分にも感情による変化があることが予想されるため、本研究では頭高型語も中高型語と同様に語頭の上昇部分の分析を行った。前川・北川(2002)においては、バラ言語的情報の表出により F0の上昇・下降のタイミングが遅れることが指摘されている。このことから、L1から H1(以下「語頭上昇部」)、H1から L2(以下「下降部」)、L2から H2(以下「語末上昇部」)の持続時間長についても計測した。

無核語の F0は、上昇イントネーションの影響で、発話の開始から終点まで直線に近い形で上昇している。発話のどの部分で上昇が急激なのか、または緩やかなのかを調べるため、発話の始点・終点の F0に加え、拍と拍の境界部分の F0を計測した。ただし、「カラオケ」の始点は無声子音であるため、有声区間の始まりの F0を始点の F0とした。

有核語・無核語共に、発話の最高 F0 と最低 F0 の差をピッチレンジとして算出した。

4 音響分析

表3～5は頭高型語・中高型語・平板型語それぞれの持続時間長、F0の平均値を示したものである。図3はそれぞれの語について持続時間長、F0の平均値を基にしたピッチ曲線のモデルである。実線は「驚いていない」、点線は「驚いている」ことを表し、白いマーカーは「嫌がっていない」、黒いマーカーは「嫌がっている」ことを表している。

音声による感情表出とその音響的特徴について

表 3 頭高型語の持続時間長(sec)、F0(st)の平均値(①「驚いていない・嫌がっていない」、②「驚いていない・嫌がっている」、③「驚いている・嫌がっていない」、④「驚いている・嫌がっている」)

頭高型		「韓国」				「毎日」			
発話意図		①	②	③	④	①	②	③	④
持続時間長 (sec)	発話全体	0.626	1.00	0.559	0.693	0.574	0.909	無し	0.818
	第 1 拍	0.123	0.205	0.103	0.179	0.107	0.192	無し	0.213
	第 2 拍	0.063	0.098	0.071	0.061	0.057	0.067	無し	0.073
	第 3 拍	0.123	0.179	0.121	0.148	0.125	0.172	無し	0.153
	第 4 拍	0.317	0.521	0.264	0.307	0.286	0.479	無し	0.379
	語頭上昇部	0.096	0.237	0.109	0.139	0.100	0.233	無し	0.259
	下降部	0.299	0.338	0.229	0.332	0.335	0.359	無し	0.325
	語末上昇部	0.200	0.369	0.158	0.181	0.139	0.317	無し	0.234
F0 (100Hz ベースの st)	レンジ	13.7	14.2	17.1	15.4	12.3	13.3	無し	18.0
	L1	18.8	10.8	21.5	15.8	16.9	11.8	無し	13.6
	H1	23.8	22.9	30.7	30.7	22.1	22.8	無し	30.5
	L2	10.9	9.7	13.6	11.5	10.5	9.8	無し	13.3
	H2	21.6	22.4	25.9	25.9	20.3	21.1	無し	23.6

中林律子

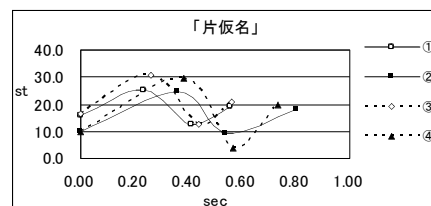
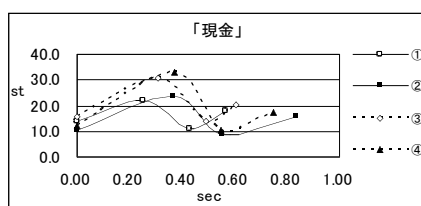
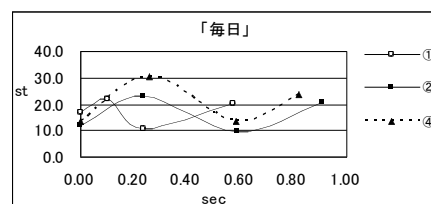
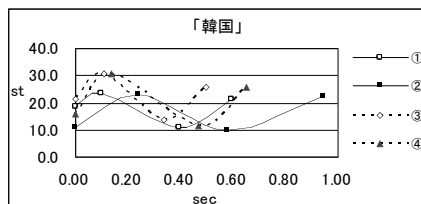
表 4 中高型語の持続時間長(sec)、F0(st)の平均値(①「驚いていない・嫌がっていない」、②「驚いていない・嫌がっている」、③「驚いている・嫌がっていない」、④「驚いている・嫌がっている」)

中高型		「現金」				「片仮名」			
発話意図		①	②	③	④	①	②	③	④
持続時間長 (sec)	発話全体	0.573	0.865	0.631	0.767	0.611	0.863	0.618	0.791
	第 1 拍	0.126	0.156	0.160	0.188	0.077	0.105	0.099	0.112
	第 2 拍	0.057	0.085	0.078	0.068	0.126	0.164	0.128	0.167
	第 3 拍	0.160	0.176	0.158	0.186	0.140	0.174	0.139	0.200
	第 4 拍	0.232	0.448	0.236	0.326	0.268	0.421	0.252	0.312
	語頭上昇部	0.254	0.367	0.307	0.369	0.234	0.359	0.263	0.386
	下降部	0.176	0.185	0.184	0.178	0.178	0.181	0.175	0.184
	語末上昇部	0.133	0.284	0.113	0.200	0.145	0.262	0.125	0.167
F0 (100Hz ベースの st)	レンジ	11.9	14.9	16.7	22.5	12.5	15.3	18.1	28.4
	L1	13.9	10.5	15.7	12.0	16.1	9.8	16.5	10.0
	H1	22.1	23.6	30.8	33.1	25.0	24.5	30.7	29.5
	L2	11.1	8.7	14.1	10.6	12.7	9.4	12.6	3.9
	H2	18.2	15.8	20.0	17.2	19.1	18.2	20.9	19.9

音声による感情表出とその音響的特徴について

表 5 平板型語の持続時間長(sec)、F0(st)の平均値(①「驚いていない・嫌がっていない」、②「驚いていない・嫌がっている」、③「驚いている・嫌がっていない」、④「驚いている・嫌がっている」)

平板型		「カラオケ」				「温泉」			
発話意図		①	②	③	④	①	②	③	④
持続時間長 (sec)	発話全体	0.478	0.902	0.498	0.711	0.446	0.745	0.463	0.587
	第 1 拍	0.096	0.186	0.127	0.170	0.110	0.175	0.131	0.185
	第 2 拍	0.096	0.138	0.109	0.132	0.069	0.099	0.057	0.079
	第 3 拍	0.063	0.102	0.051	0.093	0.220	0.415	0.215	0.289
	第 4 拍	0.223	0.476	0.212	0.315	0.048	0.056	0.060	0.034
F0 (100Hz ベースの st)	レンジ	13.9	14.0	19.7	19.1	13.2	13.3	16.2	12.6
	始点	13.7	11.3	12.2	12.2	12.3	11.7	14.4	12.7
	1・2 拍境界	14.1	10.8	13.8	14.2	17.1	13.1	22.6	21.8
	2・3 拍境界	18.7	14.6	19.8	23.2	18.3	15.9	25.7	26.5
	3・4 拍境界	20.3	16.9	21.4	26.3	23.2	23.8	28.9	23.9
	終点	27.4	24.3	31.8	31.3	25.5	25.0	30.6	25.3



中林律子

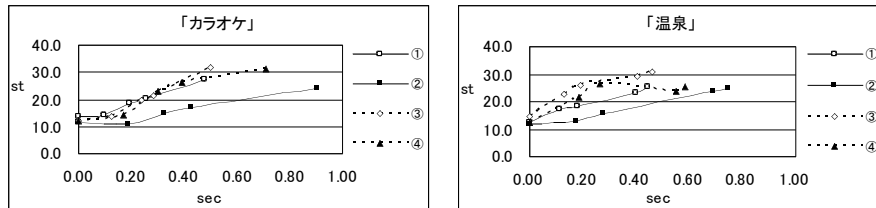


図3 持続時間長・F0 の平均値を基にしたピッチ曲線のモデル(①「驚いていない・嫌がっていない」、②「驚いていない・嫌がっている」、③「驚いている・嫌がっていない」、④「驚いている・嫌がっている」)

4-1 持続時間長

4-1-1 発話全体及び各拍の持続時間長

発話の全体長には全ての語で感情の違いによる有意な差があった(「韓国」 $F(3,39)=16.090(p<.001)$ 、「毎日」 $F(2,31)=49.188(p<.001)$ 、「現金」 $F(3,10)=6.514(p<.05)$ 、「片仮名」 $F(3,37)=12.070(p<.001)$ 、「カラオケ」 $F(3,38)=29.415(p<.001)$ 、「温泉」 $F(2,25)=42.512(p<.001)$)。特定の感情が込められていない①の発話と比較すると、頭高型語では、「韓国」で②、「毎日」で②、④と判定された発話の持続時間長が有意に長かった。中高型語では、「現金」で②、「片仮名」で②、④と判定された発話が有意に長かった。「現金」で④と判定された発話も、有意差はないものの②と同様に長い傾向がある。平板型語では、「カラオケ」で②、④、「温泉」で②と判定された発話が有意に長かった。「温泉」の④は1例のみであるが、①の平均値と比較すると、その全体長はやや長い。②、④の発話は共に「嫌がっている」と判定された発話であることから、アクセント型の違いに関わらず、「嫌がっている」と判定された発話には持続時間長が長い傾向があることがわかる。

「嫌がっている」と判定された②と④を比較すると、全ての語に共通して、「嫌がっている」とだけ判定された②よりも、「嫌がっている」とも「驚いている」とも判定された④の方が短い傾向がある。

しかし、④と判定された発話のうち、撥音/N/が含まれていない語では①と比較して有意に長かったのに対し、いずれのアクセント型においても撥音/N/を含む語では②のみが有意に長く、④では有意差が生じなかった。このことは、撥音の有無が持続時間長

に大きく影響する可能性を示しているが、なぜそのような影響が生じるのかについては今後の課題とする。

図4は各拍の持続時間長の平均値を表している。

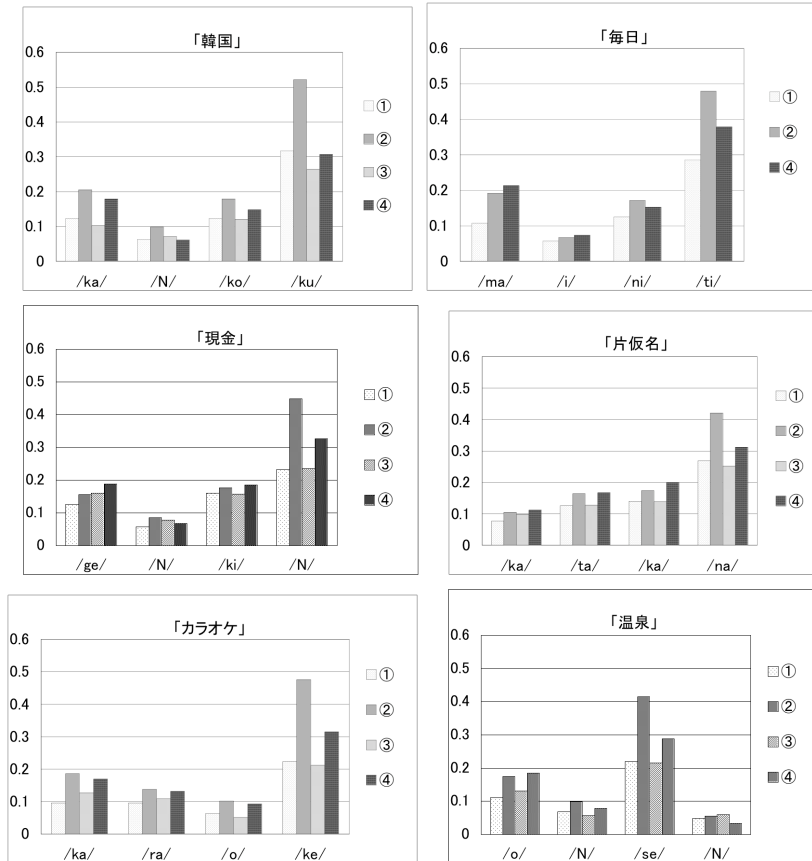


図4 各拍の平均持続時間長(sec) (①「驚いていない・嫌がっていない」、②「驚いていない・嫌がっている」、③「驚いている・嫌がっていない」、④「驚いている・嫌がっている」)

感情による伸長の差は特に第4拍で顕著に見られるが、平板型語「温泉」では第3拍で感情の違いによる差が大きい。頭高型語では第1拍の伸長にも差が目立ち、平板型語においてもその傾向があるが、中高型語ではその傾向はほとんど見られない。共に「嫌がっている」と判定され、発話全体の持続時間長の伸長の目立った②と④を比較す

中林律子

ると、伸長の顕著な第4拍(「温泉」では第3拍)においては、全ての語に共通して、「嫌がっている」とだけ判定された②に比べ、「嫌がっている」とも「驚いている」とも判定された④の方が持続時間長の短い傾向がある。

4-1-2 語頭上昇部・下降部・語末上昇部の持続時間長

有核語の語頭上昇部・下降部・語末上昇部の持続時間長を計測した結果、語頭上昇部、語末上昇部の持続時間長に有意な差が見られた(語頭上昇部:「韓国」 $F(3,39)=19.130(p<.001)$ 、「毎日」 $F(2,31)=62.756(p<.001)$ 、「現金」 $F(3,10)=5.411(p<.05)$ 、「片仮名」 $F(3,37)=15.569(p<.001)$ 、語末上昇部:「韓国」 $F(3,39)=13.549(p<.001)$ 、「毎日」 $F(2,31)=27.470(p<.001)$ 、「現金」 $F(3,10)=3.877(p<.05)$ 、「片仮名」 $F(3,37)=18.843(p<.001)$)。一方、下降部では4種類全ての語において感情による有意な差が生じず、平均値においても目立った差は見られなかった。

語頭上昇部については、特定の感情が込められていない①の発話と比較して、「韓国」で②、「毎日」で②、④、「現金」で④、「片仮名」で②、④と判定された発話で有意に長かった。「現金」で②と判定された発話の語頭上昇部も、有意差はなかったが④と同様に長い傾向がある。②、④の発話は共に「嫌がっている」と判定された発話であることから、アクセント核の位置に関わらず有核語で「嫌がっている」と判定された発話には、語頭上昇部の持続時間長が長い傾向があると言える。

発話全体及び第4拍(「温泉」では第3拍)の分析においては、「嫌がっている」とだけ判定された②と比較して、「嫌がっている」とも「驚いている」とも判定された④の持続時間長が短い傾向があった。しかし、この語頭上昇部については「驚き」の感情の有無による伸長の違いはほとんど見られなかった。

語末上昇部については、特定の感情が込められていない①と比較して、「韓国」で②、「毎日」で②、④、「片仮名」で②と判定された発話で有意に長かった。「現金」では多重比較による有意差は見られなかったが、平均値では②、④の語末上昇部が長い傾向がある。これらのことから、語末上昇部においても「嫌がっている」と判定された②、④の持続時間長が長い傾向があると言える。

「嫌がっている」と判定された②と④の語末上昇部を比較すると、「嫌がっている」とだけ判定された②に比べ、「嫌がっている」とも「驚いている」とも判定された④では持続時間長が短い傾向がある。語末上昇部は感情による伸長の差が顕著であった第4拍を含

んでいるため、第4拍の計測で見られたのと同様の傾向が表れていると考えられる。

以上のことから、語頭上昇部・下降部・語末上昇部それぞれの部分において「嫌」という感情が含まれたことによる持続時間長の伸長は語頭上昇部及び語末上昇部に表れる傾向があること、さらに、「驚き」の感情の有無は語頭上昇部には影響しないが語末上昇部の伸長には影響し、「驚き」の感情が含まれることにより持続時間長がやや短くなる傾向があることが考えられる。

4-2 F0(基本周波数)

4-2-1 ピッチレンジ

ピッチレンジには「温泉」以外の語で感情の違いによる有意な差があった(「韓国」 $F(3,39)=3.487(p<.05)$ 、「毎日」 $F(2,31)=12.817(p<.001)$ 、「現金」 $F(3,10)=8.242(p<.01)$ 、「片仮名」 $F(3,37)=26.006(p<.001)$ 、「カラオケ」 $F(3,38)=9.590(p<.001)$)。特定の感情が込められていない①の発話と比較すると、頭高型語では、「韓国」で③、「毎日」で④と判定された発話のピッチレンジが有意に広がった。有意差はなかったが、「韓国」で④と判定された発話のピッチレンジも①と比較してやや広い傾向がある。中高型語では、「現金」で④、「片仮名」で③、④と判定された発話のピッチレンジが有意に広がった。「現金」で③と判定された発話のピッチレンジも、有意差はなかったが、①と比較して広い傾向がある。平板型語では、「カラオケ」で③、④と判定された発話のピッチレンジが有意に広がった。一方、分散分析において有意差が見られなかった「温泉」では、平均値では③と判定された発話のピッチレンジがやや広い傾向があった。③、④の発話は共に「驚いている」と判定された発話であることから、有意差の出なかったケースがあるものの、アクセント型の違いに関わらず、「驚いている」と判定された発話にはピッチレンジが広い傾向があると考えられる。

「驚いている」と判定された③と④を比較すると、中高型語では、「驚いている」とだけ判定された③よりも、「驚いている」とも「嫌がっている」とも判定された④のピッチレンジのほうが顕著に広い傾向があった。しかし、頭高型語・平板型語ではこの特徴は見られなかった。

図3を見ると、有核語では、特に H1の F0に「驚き」の感情による差が見られることから、アクセント核による下降がどの高さで始まるかがピッチレンジの広狭に大きく関与していることがわかる。一方、無核語である平板型語では、語や感情の種類によってF0の上昇

中林律子

の仕方にばらつきがあるが、発話の終点の F0に「驚き」の感情による差が目立つことから、発話がどの高さで終わるかがピッチレンジの広狭に関与していると考えられる。

4-2-2 発話開始点(L1)の F0

有核語の L1～H2の F0を見ると、ピッチレンジの広狭に大きく関与している H1だけでなく、発話の始点である L1の F0に違いが目立つ。頭高型語では、「韓国」「毎日」共に感情の違いによる有意な差があった(「韓国」 $F(3,39)=14.079$ ($p<.001$)、「毎日」 $F(2,31)=10.873$ ($p<.001$))。特定の感情が込められていない①の発話と比較すると、「韓国」で②、「毎日」で②、④と判定された発話の L1が有意に低かった。中高型語では、「現金」の L1には感情の違いによる有意差は見られなかったが、「片仮名」では有意な差があり($F(3,37)=9.307$ ($p<.001$))、②、④と判定された発話の L1が有意に低かった。「片仮名」では④と判定された発話で L2が低いことも目立つが、他の語ではこの傾向は表れていない。L1の F0が低い傾向の見られた②、④の発話は共に「嫌がっている」と判定された発話であることから、有核語で「嫌がっている」と判定された発話には、発話開始点の F0が低い傾向があると考えられる。これら②と④の発話を比較すると、差の見られなかった「韓国」の④を除き、「嫌がっている」とも「驚いている」とも判定された④の L1よりも、「嫌がっている」とだけ判定された②の L1の方がわずかに低い傾向が見られる。

平板型語「カラオケ」「温泉」では、発話開始点の F0に感情の違いによる有意差は見られなかった。

5 まとめと考察

音響分析により、聴取実験判定者が「(話し手は)嫌がっている」と判定した②、④の発話には、全てのアクセント型に共通して、発話の持続時間が長く、特に第4拍を含む(「温泉」では第3拍)語末での伸長が顕著であるという特徴があった。しかし、「温泉」で最も伸長が顕著だったのは第3拍であった。「温泉」では第4拍が撥音であることが関係していると思われるが、同様に第4拍が撥音である「現金」ではその傾向は見られなかった。このことは、発話の音響的特徴が感情によって変化する際に、音節構造、アクセント型によってその変化の仕方が多様である可能性を示している。

有核語では、拍の伸長が顕著であった語末だけでなく、語頭上昇部の持続時間長も

伸長する傾向があった。川上(1956)では句頭の上昇の遅れが話し手の感情を反映することが指摘されているが、本研究で「嫌」と判定された発話の語頭においても上昇の遅れが生じていると考えられる。有核語においては、「(話し手は)嫌がっている」と判定された発話において発話開始点の F0が低い傾向も見られた。以上の結果から、(a) 発話の持続時間長が長いことに加え、有核語では(b)発話が低く始まること、(c) 発話の始めの上昇が遅れることが、判定者に「(話し手は)嫌がっている」と知覚させたと考えられる。

「嫌がっている」と判定された②、④を比較すると、全体的に「嫌がっている」とだけ判定された②に比べ、「嫌がっている」とも「驚いている」とも判定された④の持続時間長が短い傾向があった。この差は有核語の語頭上昇部ではほとんど見られなかったが、伸長の差が顕著であった第4拍(「温泉」では第3拍)で顕著に見られた。「驚いている」と判定された発話にはピッチレンジが広い傾向があるため、「嫌」「驚き」両方の感情を表現するにはピッチレンジを広く、持続時間長を長く発話する必要がある。そのため、「嫌」「驚き」両方の感情を表す発話にはより労力がかかると考えられ、このことが④の持続時間長がやや短くなることに影響しているのではないと思われる。

「韓国」において④「驚いている・嫌がっている」と判定された発話では、「嫌がっている」と判定された発話に見られた上記(a)～(c)のいずれの傾向もほとんど見られなかった。また、平板型語においても(b)、(c)の傾向がほとんど見られなかったことを併せると、持続時間長・F0以外の音響的パラメータも「嫌」という感情を推測するための手がかりとなっていることが考えられる。Teshigawara(2004)、Scherer et al. (1984)では、音声による感情の推測に声質が影響することが指摘されており、今後、声質等の影響についても検討が必要であると思われる。今回の平板型語の分析において、(b)、(c)の傾向が見られなかったことについては、このことが平板型語の特徴と言えるのか、または今回の音声資料にこれらの傾向が表れなかっただけなのかは、現時点では判断が難しい。今後、音声資料を増やし、無核語による感情表出についてさらに検討したい。

一方、聴取実験判定者が「(話し手は)驚いている」と判定した発話には、全てのアクセント型に共通してピッチレンジが広いという特徴が見られた。ピッチレンジが広いという特徴は、有核語ではアクセント核による下降の始まりが高いこと、無核語では発話の終点が高いことによるものであった。無核語では音節構造や感情の違いによって F0の上昇の仕方や持続時間長の伸長の仕方にばらつきが目立った。このことは、アクセント核がないことにより、感情を表現するための F0の上昇が比較的自由に行われることを示唆

していると思われる。

「驚いている」と判定された③、④を比較すると、中高型語においては、「驚いている」とだけ判定された③に比べ、「驚いている」とも「嫌がっている」とも判定された④のピッチレンジの方が広い傾向があった。中高型語「現金」「片仮名」の④では、H1の部分が一セットになるほどの高音で発話されていることがあったが、他のアクセント型ではその傾向はほとんど見られなかった。③と比較して④のピッチレンジが広いことは、④と同様に「嫌がっている」感情を表している②との区別を明確にするためではないかと思われるが、アクセント型によってはその区別のつけやすさが異なっている可能性が考えられる。

本研究では、「嫌」「驚き」の感情を表す音響的特徴として、ほぼ全ての語に共通して見られる特徴と、アクセント型や音節構造の違いにより異なる特徴とが見られた。日常のコミュニケーションにおいて音声から感情を推測する際、非日本語母語話者にとってはこれらの違いが推測に影響する可能性がある。本研究においては F0のピーク(H1)の遅れは「(話し手は)嫌がっている」という判定に影響する可能性が示されているが、Chen et al. (2004)では、英語とオランダ語において「驚いている」と判定された発話では、F0のピークが遅れるという特徴があることが報告されている。今後、F0のピークを始めとし、日本語音声に表れる音響的特徴が非日本語母語話者にどのような感情を知覚させるのかを詳細に検討する必要があると考えられる。

6 おわりに

本研究では、「嫌」「驚き」の感情を変数とした問い返し疑問文の音響分析を行い、これらの感情に伴い持続時間長及び F0が多様に変化することが示された。しかし、聴覚実験において判定一致度の高い発話が少ないグループがあったため、有意差が生じなかった項目も多くあった。アクセント型や音節構造による音響的特徴の違いをより明らかにするためにも、今後、音声資料を増やして分析を行う必要がある。さらに、聞き手が話し手の感情を推測する際にこれらの音響的特徴をどの程度手がかりとして利用しているのかについても、合成音声を用いた知覚実験により確認する必要がある。その上で、これらの結果を元に、非日本語母語話者が日本語音声から感情を推測する際に、どのような音響的パラメータを手がかりとしているのか、また、日本語母語話者が推測の手がかりとしている音響的特徴に対しどのように反応するのかといった問題を明らかにしていきたい。

[参考文献]

- Campbell, N. and D. Erickson (2004) “What do People Hear? A Study of the Perception of Non-verbal Affection Information in Conversational Speech,”『音声研究』8(1), pp.9-28.
- Chen, A., C. Gussenhoven, & T. Rietvelt, (2004) “Language-specificity in the perception of paralinguistic intonational meaning,” *Language and Speech* 47(4), pp.311-349.
- コーネリアス, ランドルフ・R.(1999) 『感情の科学—心理学は感情をどこまで理解できたか』齋藤勇監訳, 誠心書房
- 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『文法と音声Ⅴ』くろしお出版, pp.31-46.
- Fant, G., A. Kruckenberg, K. Gustafson, & J. Liljencrants (2002) “A new approach to intonation analysis and synthesis of Swedish,” *Speech Prosody 2002, Aix en Provance*, <http://www.lpl.univ-aix.fr/sp2002/pdf/fant-et-al.pdf>.
- 鹿島 央(1992) 「イントネーションパターンとアクセント型に関する実験音声学的考察—長さ基本周波数(F0)の変化を中心として—」『日本語論究1 言語学とその周辺』和泉書院, pp.21-44.
- 川上 泰(1956) 「文頭のイントネーション」『国語学』25, pp.21-30.
- 北原義典・東倉洋一(1988) 「音声の韻律情報と感情表現」『電子情報通信学会技術研究報告』SP88-158, pp.27-32.
- 近藤研至(2001) 「「意外である」ということと「問い返し疑問文」について」『言語と文化』14, 文教大学言語文化研究所, pp.16-27.
- 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』吉沢典男教授追悼論文集編集委員会編, pp.21-30.
- 郡 史郎(1997) 「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, pp.169-202.
- Ladd, D.R.(1980) *The Structure of Intonational Meaning : Evidence from English*,

中林律子

Indiana U. P., Bloomington, IN.

前川喜久雄・北川智利(2002)「音声はパラ言語をいかに伝えるか」『認知科学』9(1), pp.46-66.

南不二男(1985)「質問文の構造」『朝倉日本語新講座4 文法と意味Ⅱ』朝倉書店, pp.39-74.

森山卓郎(1989)「内容判断の一貫性の原則」仁田義雄・益岡隆志編『日本語のモダリティ』くろしお出版, pp.75-94.

森山卓郎(1990)「『断り』の方略—対人関係調整とコミュニケーション」『言語』19(8), pp.13-30.

西端千香子・浅田健太郎・篠崎大司・徐愛紅・茅田美有紀・細田和雄(1996)「「ア」と「エ」で表出される感情の韻律的特徴」『広島大学日本語教育学科紀要』6, pp.63-72.

Peterson, G. E. and I. Lehiste (1960) "Duration of syllable nuclei in English," In I. Lehiste, (ed.)(1967) *Readings in Acoustic Phonetics*, The M. I. P. Press, pp.191-201.

Scherer, K. R., D. R. Ladd, & K. E. A. Silverman (1984) "Vocal cues to speaker affect : Testing two models," *The Journal of the Acoustic Society of America*, 76(5), pp.1346-1356.

重野 純(2004)「感情を表現した音声の認知と音響的性質」『心理学研究』74(6), pp.540-546.

Teshigawara, M.(2004) "Vocally expressed emotions and stereotypes in Japanese animation: Voice qualities of the bad guys compared to those of the good guys," 『音声研究』8(1), pp.60-76.

土岐 哲(1998)「アクセントの下げとイントネーションの下げ」『阪大日本語研究』10, pp.53-66.

土田昭司(1996)「感情の社会的判断—意思決定と態度構造」土田昭司・竹林和久編『対人行動学研究シリーズ4 感情と行動・認知・生理—感情の社会心理学』誠信書房, pp.103-126.

上村幸雄(1989)「日本語のイントネーション」『ことばの科学3』言語学研究会, むぎ書房, pp.193-220.

音声による感情表出とその音響的特徴について

〔資料〕 音声資料収集に用いたダイアログ(例「韓国」①④)

<p>① 「驚いていない・嫌がっていない」</p> <p>あなたは友達と夏休みの海外旅行の行き先について話しています。二人ともヨーロッパへ旅行するつもりでしたが、友達は韓国もいいのではないかと提案してきます。確かに、韓国旅行は安く済むし、今人気もあるので、あなたは韓国へ行くのも悪くないと思います。</p> <p>友達：ヨーロッパもいいけど、韓国へも行ってみたいと思わない？</p> <p>あなた：韓国？それもいいね。</p>
<p>④ 「驚いている・嫌がっている」</p> <p>友達が突然、来週末に韓国へ行かないかと誘ってきました。あなたは急な提案にとっても驚いたし、今から旅行の準備をするのは大変そうだし、韓国にもあまり興味が無いので友達の誘いを断りたいと思います。</p> <p>友達：来週の週末、空いてる？</p> <p>あなた：うん、空いてるけど。</p> <p>友達：一緒に韓国に行かない？</p> <p>あなた：（あまりに急な話なので驚いたし、準備も大変そうだから行きたくないと思って）韓国！？</p> <p>友達：うん、一緒に行くはずの子が、急に行けなくなっちゃって。無理？</p> <p>あなた：うーん。ちょっと急すぎるから・・・。</p>

