

「平静の答え」と「不満の答え」の弁別に対する イントネーションパタンの影響

田 川 恭 識

(平成18年 3月31日 提出)

要旨：音声は言語的情報のほかに様々な情報を伝達する。特に音声に含まれる感情情報やパラ言語情報は、状況によってはミスコミュニケーションを引き起こす原因ともなる。音声言語に含まれる感情情報およびパラ言語情報がどのように表現されるかについて明らかにすることは、音声による円滑なコミュニケーションを行う上で極めて重要であると考えられる。以上を踏まえ、本研究では音声に含まれるパラ言語情報、感情情報について、実験音声学的手法から分析を行った。具体的には、「あわないの」という発話を対象に仮想的な会話の場面を設定し、「平静」の音声と「不満」の音声について音響的分析を行った。その結果、「不満」の音声では、文末部分のイントネーションパターン（基本周波数パターン＝F0パターン）に特徴が見られた。分析の結果を踏まえ、合成音声を作成し聴取実験を行ったところ、「不満」の音声の弁別において、文末部分のイントネーションパターンが重要である事が明らかになった。

キーワード：音声コミュニケーション、イントネーション、感情情報・パラ言語情報、音響分析、聴取実験

1. はじめに

来日外国人の増加に伴い、普段の生活において外国人と接することは、もはや特別な事ではない。外国人との接触は、文化的背景を異にする者同士の接触場面として典型的な例であるが、そのような接触場面において、摩擦や軋轢が生じることも少なくない。摩擦の背景には、互いの文化の相違による習慣の違いや、日本語の運用能力に起因する意思疎通の滞りといった様々なレベルの問題が存在する（大坊・安藤・池田他1990）。日本語教育において、このような摩擦を回避することを目的とし、コミュニケーション能力の育成を重視した授業が志向されるようになって久しい。文法を中心とした学習事項の積み重ねだけではなく、会話の場面に即した表現の選択、円滑なコミュニケーションを行う為の方策の習得、さらには互いの文

化理解の促進といったものに焦点が当てられている（梶原2003）。また、近年では音読を推奨する書籍が出版市場を賑わしており、日本語教育以外の国語教育においても、専門誌で音読の特集が組まれるなど注目が高まっている。コミュニケーションにおける音声の重要性が広く一般に再認識され、音声表現についての関心が高まっていると言える。播磨（2001）では「相手と良好な関係を保ちつつコミュニケーションをとろうとする場合、使用する言葉についても配慮する必要がある。」として、以下の注意点を挙げている。

- （1）相手との立場や優しさなどの関係を把握して、ふさわしいことばづかいをする。
- （2）相手を傷つけたり不快にさせるようなことばを使用しない。
- （3）相手と共有している情報を用いる。つまり、

田 川 恭 識

相手が理解できないと思われることばや、誤解をまねく恐れのあることばを避ける。(播磨2001)

他者と円滑なコミュニケーションを行う上で、以上の指摘は重要であると考えられる。しかし、音声によるコミュニケーションを行う上で、「どんなことばで伝えるか」ということと同時に、それを「どのような言い方で言うか」ということも重要である。

学習者のコミュニケーション能力を育成するという目的において、音声教育は極めて重要である。しかしながら、従来の音声教育では、アクセントを始めとする言語的情報の表現に重点を置いたものが多かった。日本語学習者によって発せられた何気無い一言が、聞き手によっては批判的に聞き取られることがある。文法的に誤りは無く、使用場面も適切であるにも関わらず、このような摩擦が生じるのは何故であろうか。その理由の1つとして、学習者の発話に付随する音声的特徴の影響が挙げられる。日本語教育の現場では、「イントネーションやプロミネンスによって表される発話意図は具体的なものとではなく、心情的、感情的なものであるため、「誤解」が「誤解」と気づかれないままになりがちである」(水谷・鮎澤1991)という声が聞かれる。音声によって生じる摩擦を回避するためにも、従来の言語的情報の表現を中心とした音声教育に加え、コミュニケーション能力の育成を視野に入れた、より高度な音声教育の実践が望まれる。また、「イントネーションやプロミネンスの記述が進むならば、それは、日本語学習者のためだけでなく、日本人同士のコミュニケーションにも役立つものとなる」(水谷・鮎澤1991)との指摘からも明らかなように、日本語母語話者同士のコミュニケーションのあり方を捉える上でも極めて有益である。以上を踏まえ、本論文では、円滑なコミュニケーションを行う上で重

要な役割を果たす音声に注目し、音声に含まれる感情情報・パラ言語情報について、実験音声学的視点から考察を行う。

2. 先行研究

2. 1. 音声に含まれる情報

音声は、言語的成分によって構成された具体的情報(言語的情報)だけでなく、話し手の年齢、性別および心的状態といった、言語以外の情報をも伝達する。言語音声における、言語以外の情報についての分類は、パラ言語的情報、非言語的情報、言語外的情報、韻律の情報など、各研究者間においても様々な見解がある(前川・北川2002)。中でも藤崎(1994)では、言語音声によって伝達される情報について、①「言語情報」、②「パラ言語情報」、③「非言語情報」の3種を挙げている。以上の3種の情報について、前川(1998)によれば「言語的情報」とは、「言語記号が伝達する情報であり、主知的な意味の対立に関係する情報」であり、「パラ言語的情報」には「対話の制御に関わる情報のほか、発話の意図や話者の心的態度に関する情報が含まれる」としている。そして、「非言語的情報」とは「話者が意図的に制御できない情報であり、主として話者の身体性に関する情報」であり、さらにそれらには「話者の性別や個人的特徴のほか、体調の良否や、生理的反応としての感情(恐怖、驚きなど)も非言語的情報」に含まれるとしている。以上の分類は、「音声に特化した分類である」こと、また、「話し手が意図的に制御して生成する情報とそうでないもののとの区別を重視している」(前川・北川2002)という点から、実証的研究を行う上で有益であると考えられる。しかし、「パラ言語情報」に含まれる「心的態度」と「非言語情報」に含まれる「感情」は、どちらも話者の心理状態に関係する事象であり、両者は連続的であるとの見方も可能

「平静の答え」と「不満の答え」の弁別に対するイントネーションパタンの影響

である。以上について、相澤（2000）では、藤崎（1994）の3分類について「パラ言語的情報は、その時々話し手の意図によって比較的コントロールされやすいのに対して、非言語的情報は、話し手の意図とは別にむしろ無意識のうちに聞き手に伝わってしまうという点で大きな違いも見られる。パラ言語的情報は、話し手の発話時を含む比較的短時間の心的態度やその変動を伝えるのに対して、非言語的情報は、話し手のある程度恒常的な生理的・心理的状态を伝えるということも出来そうである」と述べている。本稿では、基本的に藤崎（1994）、および前川・北川（2002）の考えに従うが、「パラ言語情報」と、感情を含む「非言語情報」の間に明確な線引きを行うことは困難であるという考えから、両者は連続性を持つという立場で分析を行う。

2. 2. パラ言語情報とイントネーション研究

言語音声は、母音や子音といった個々の音の音響的特徴である分節的特徴と、声の高さ、強さ、長さによって表される韻律的特徴を持つ。韻律的特徴は分節的特徴に被さるため、超分節的特徴とも呼ばれる。日本語の韻律的特徴の研究については、特にアクセント研究において長い歴史を持ち、その過程において研究者間でアクセントについての捉え方、すなわちアクセント観について議論が展開するとともに、日本語諸方言のアクセント体系が記述されてきた（上野1977）。一方、イントネーションについては、イントネーションそのものをどのように捉えるべきかについて、アクセントの場合よりもさらに研究者間で意見の相違が見られる（国立国語研究所（1960））。国立国語研究所（1960）では、「話しことば」について構文と表現意図、さらには発話のイントネーションという三つの視点から「話しことばの文型」として体系的な記述を試みている。その中でイントネーショ

ンについては「句末や文頭のイントネーションも、イントネーション独自の問題としては興味ある問題ではあるが、いわゆる「文」の陳述の違いをあらわすと考えられる文末部分にあらわれるイントネーションは、句末や文頭に比べて、「文」に対する重みが極めて大きいと考えられる」という理由から、主として文末部分の高低変化を分析の対象としている。そして、文末のイントネーションについて、①「平調」、②「昇調1」、③「昇調2」、④「降調」、⑤「 α 型類（注：原文では、○に α ）」の5種に分類している。

また、国立国語研究所（1960）に続く国立国語研究所（1963）においても、文末のイントネーションの分類として、①「上昇調」、②「下降調」（意図表現のイントネーション）、③「高調」、④「低調」（卓立表現のイントネーション）の4種を挙げている。分類の枠組みや種数は異なるが、発話を構成する「文」の意味や機能から見て、文末部分のイントネーションが重要である、という視点に違いは見られない。同様の視点からの文末のイントネーションについての詳細な記述には、上村（1989）、郡（2004）などが挙げられる。

これに対し、文末ではなく、文頭のイントネーションについて注目したものに川上（1956）がある。これは、アクセント（東京方言）において、「非弁別的特徴」とであるとされている文頭（あるいは句頭）の声の上昇が、話し手の意図や感情状態を表現する上で重要な要素となることを指摘したものである。例として、川上（1956）では、「トンデモナイ」という発話について、以下のよう

【1】ト「ンデモナ' イ

【2】トン「デモナ' イ

【3】トンデ「モナ' イ

【4】トンデモ「ナ' イ

（「」は上昇点、「'」はアクセントの下げ）

東京方言の音調の原則として、頭高型の語を除き、語の第一拍目の高さは低く、第二拍目は高い(川上1956)。**【1】**の型は第一拍目が低く、第二拍目から高くなり、「ナ」から後続の拍にかけて高さの下降が生じる。**【1】**の型は東京方言の規範的な音調であると言えるが、このような音調に対し、「然しこれは、冷静な発話に於ける音調であることに注目しなければならない。或は寧ろ無表情、冷淡な発話とさえ云えるくらいである」と述べている。そして「生き生きとした、自由な感情に彩られた発話」では、**【1】**の型では表現できない音調がみられるとする。これに対し、**【2】**は**【1】**と異なり二拍目の後に上昇点があり、さらに**【3】**、**【4】**では上昇点が後ろの拍へずれている。上昇点が後方へずれた音声は、「驚きの気持・意外案外だとの気持・納得の行きかねる気持・あきれた気持・当惑した気持・または遠慮がちにものを云うといったような気持」を表し、後方へずれるごとにその意味が強くなるとしている。上昇点がずれることによって表現される「気持」とは、ある種のパラ言語的・非言語的情報を指すと考えられるが、以上の知見を実証的な研究に援用するためには、さらに整理が必要である。しかしながら、それまでのイントネーション研究において注目されることの少なかった「文頭のイントネーション」について焦点を当てたことで、後の研究に与えた影響は少なくない。

以上の研究に代表されるように、日本語音声学の分野においては比較的古い時代から示唆的な研究が見られる。ただ、そのような研究の多くが、内省によって行われているため、科学的裏付けが保証されていないのも事実である。

このような状況に対し、近年ではコンピュータを始めとする分析機器の飛躍的な向上もあり、音声による感情情報・パラ言語情報の表現について定量的な研究が可能となった。音声学以外の分野

においても、音声による感情表現についての研究が盛んに行われるようになり、中でも音響学などの工学分野では、音声認識や高品質な音声合成技術の開発を目的とし、活発な研究が行われるようになった(Williams and Stevens 1972、Ladd 1985、Murray and Arnott 1993)。そのような状況の中で、杉本(2000)では、人間とコンピュータとの声を用いたインターフェースの設計、合成音声の品質向上にとって、音声に含まれる感情情報の抽出および感情情報の制御が重要であるとし、「歓喜」、「悲哀」、「怒り」の調子で発話された音声と平静の音声について、物理的特徴の比較を行っている。比較分析によって得られた物理量変換ルールを平静の音声に適用して聴取実験を行った結果、声の高さである基本周波数の変化パターンと声の強さであるパワーの変化パターン、および持続時間長の操作が有効であることを述べている。またMaekawa(1998)では、Admiration、Disappointment、Indifference、Focused、Neutralとして発話された音声サンプルを用いて聴取実験を行い、それぞれの音声について音声分析を行っている。その結果、各発話において音声の基本周波数曲線であるF0曲線、時間長、音色に違いが見られることを報告している。以上の研究に代表されるように、感情音声の音響特徴量の記述については一定の成果が得られてきたと言える。しかしながら、日本語教育を始めとする言語教育的な観点からの実証的研究はほとんど無いのが現状である。外国人学習者の音声は、批判的に感じられることがあるが、発話者である学習者は、「批判」といった特定の感情の元で発話していないとすると、このような事例はミスコミュニケーションであると言わざるを得ない。コミュニケーションを重視した音声教育を目指す上では、まずこのような事態の改善を目指すべきである。その為には、これまでの感情情報・パラ言語情報が込められた音声につ

「平静の答え」と「不満の答え」の弁別に対するイントネーションパタンの影響

いての研究成果を援用しつつ、実際のコミュニケーション場面に即した観点で実証的研究を行っていくことが望まれる。

3. 研究の視点

実際に分析を行うにあたり、本研究で分析の対象とする感情と音響的パラメータについて整理して置きたい。音声による感情表現の研究において、怒りや悲しみ、喜びといった主要な感情を押さえて取り上げることは、各感情音声の物理的特徴の記述および音響パラメータ間の関連性を検討する上で有効であると考えられる (Pittam 1993)。また実際にそのような観点からの研究が多数を占めている。しかしながら、コミュニケーション上で現実的に問題となるのは先に述べた誤解が生じる場面であると考えられる。よって、コミュニケーション上摩擦が生じるとされる現象に焦点を絞り、研究を行うことが望ましいと考える。

以上を考慮し、これまでの筆者らによる一連の研究では、コミュニケーションの場면을仮想的に設け、そこで行われる発話について分析を行った (会話の場面については、5 節を参照)。先の研究 (田川他2000、田川他2001a) では疑問文を対象に、「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。本研究では、平叙文を対象に、「平静の答え」と不満の答えについて考察を加える。ここでは異なる「非難」や「不満」という異なる表現を用いているが、両者はパラ言語情報の上では同じカテゴリーに属すと考えられる。「非難」と「不満」という2つの表現を用いたのは、文として実現された場合の語用論的性格に差があると考えられるためである。音声に対し、どのようなパラ言語情報のラベルを貼るかについては研究者間でさまざまな意見の相違が見られる (Crystal 1969、森・相澤・粕谷2005)。以上の問題についての詳細な議論は、本稿では行わない

が、実験の参加者 (発話者及び聴取実験の被験者) が理解しやすいように、平易な表現を用いることを心がけた。

また音声の物理的性質として、①声の高さ (frequency)、②声の強さ (intensity)、③音色 (spectrum)、④時間長 (duration) などが挙げられる。「怒り」の音声の弁別において、声の高さの変化パターンや変化の幅が主要な手がかりになることが先行研究により明らかにされている (杉本2000、平館2002、Pittam 1994)。以上を踏まえ、本研究においても、声の高さの変化パターンに焦点を絞り考察を行う。なお、声の高さは、一秒あたりの声帯振動数によって決定され、その時間変化パターンは基本周波数曲線 (Fundamental frequency = F_0) 曲線によって表現される。本研究では、声の高さの物理的表現である F_0 曲線と、 F_0 値を中心に考察を行う。

4. 音声分析

4. 1. これまでの研究

先の研究 (田川他2000、田川他2001a) では、疑問文である「あわないの?」という発話を対象に、「問いかけ」と「非難の問いかけ」の弁別と F_0 パタンの関係について考察した。その結果、以下のことが明らかになった。

- ①発話開始部が低く、なおかつ発話開始から第二拍目にかけての高さの上昇幅が大きい場合、安定して「非難の問いかけ」として知覚される。
- ②第二拍目からアクセントによる下降部分までの高さの下降幅は、「非難の問いかけ」の判別に影響を与えにくい。
- ③文末部分の上昇幅は、「非難の問いかけ」の判別に大きな影響を与えない。

①から③の結果により、「非難」の音声の弁別において、文頭部分の高さや上昇部分までの上昇幅が重要であることが明らかになった。音声学的にはアクセントによる下降は語の弁別に中心的な役割を果たし、文末部分の上昇はその発話が「問いかけ」であることのマーカーとして機能する。従ってそのような役割を担っていない文頭部分が、バラ言語的情報を表現する一つのチャンネルとして機能すると考えられる。以上の結果は、川上(1956)の指摘と共通する部分もあり興味深いが、疑問文の場合の結果であり、平叙文においても同様の傾向を示すかについては定かではない。本研究では、先の実験と同様に「あわないの」という音声を取り上げ、それが「平静」として発話された場合を「平静の答え」、また「不満」を伴って表れた場合の音声を「不満の答え」とし、それぞれの音声のイントネーションパターン(F0パターン)と、その特徴が両音声の弁別に与える影響について考察する。

4. 2. 音声資料

分析の対象となる音声資料には、田渕他(1999)によって収録された音声の一部を使用した。収録は平静の問いかけと非難の問いかけの音声と同時に行った。本研究で取り上げるのは、波線で示した発話である。発話者は30代の静岡県出身の女性で、かつて劇団に所属した経験がある。収録に当たり、発話者に紙面で会話の場面を提示した。発話者は会話の場面をよく読み、発話の状況を理解した後、それぞれの音声の発話について練習を行った。収録はスタジオで行われ、収録した音声は標本化周波数44.1kHz、16bit量子化を行った。収録後、日本語母語話者を対象に確認聴取実験を行い、収録音声が発話者の意図通りに聴取者に知覚されるか確認を行った。その結果を踏まえ、高い割合で話者の意図通りに知覚された音声进行分析の

対象とした。

I) 佐藤さんと山田さんは同じ部署で働く仲の良い同僚です。月末の収支決算をしています。佐藤さんがたずねます。

佐藤：合わないの？（問いかけ）

山田：合わないの。（答え）

II) 佐藤さんと山田さんは同じ部署で働く仲の良い同僚です。月末の収支決算をしています。佐藤さんは計算が合いましたが、山田さんはなかなか合いません。計算が合うまで二人とも帰れません。佐藤さんは早く帰りたいので山田さんを責める口調でたずねます。

佐藤：合わないの？（非難の問いかけ）

山田：合わないの。（不満の答え）

図1 会話の場面

4. 3. 分析結果

本研究では音声の分析、合成にあたり、音声分析変換合成方式 STRAIGHT（河原1998）を用いた。STRAIGHTを用いることにより、高精度の音声分析が可能となる。また極めて自然性の高い合成音声を得られる。STRAIGHTはMath Works社が開発したインタープリタ型のプログラム言語 MATLABによって構築されているが、STRAIGHTがインターフェース上で提供する機能は、そのままではF0曲線、パワー、スペクトルなどの抽出と、声道長や発話速度の変換などのみである。F0やパワー、スペクトルを始めとする音声の物理量を操作するためには、目的とする処理に応じてSTRAIGHTの機能を拡張、変更しなければならない。本研究ではもともとのSTRAIGHTに早稲田大学の北原真冬氏によって開発されたプログラムを組み込み、さらにそれらを改良することで音声合成を行った。

「平静の答え」と「不満の答え」の弁別に対するイントネーションパタンの影響

図2に「答え」のF0曲線、図3に不満の答えのF0曲線を示す。縦軸はHzを、横軸は時間長（マイクロセカンド）を表す。

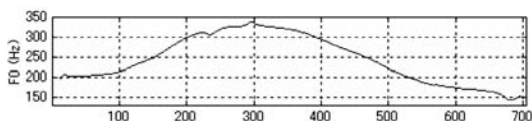


図2 「答え」のF0曲線

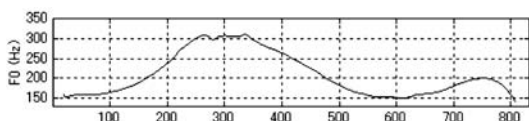


図3 不満の答えのF0曲線

図2、3から分かるように、不満の答え（図3）の方が、始まり部分のF0値が低い（「答え」の平均199.4Hzに対し、不満の答えの平均は170.4Hz）。また不満の答えでは、文末部分において「答え」にはない山型の起伏パターンが見られる。これらの起伏パターンは他の不満の答えにも共通して観察されたことから、不満の答えを特徴付けていることが窺える。先に見た国立国語研究所（1960）では、文末部分のイントネーションとして、①「平調」、②「昇調1」、③「昇調2」、④「降調」、⑤「α型類」の5種が挙げられていたが、ここで本研究と関連する文末のイントネーションパターンについて関連を見てみたい。

まず「平調」であるが、この音調は文末が特に上昇や急激な下降を伴うこともなくアクセントに従って言い収められる音調である。本研究では、「平静の答え」の音調が相当すると考えられる。次いで「昇調1」はいわゆる質問の際の、発話末における上昇の音調が相当する。これに対し、「昇調1」と区別して、「話し手が自分の思うところを強調して、文末部を卓立させることがある」とし、そのような音調として「昇調2」を立てて

いる。具体的には、以下のような例が挙げられている。

コレハ シャシ[△]ン[△]。
アナ[△]タハ ホントニ シリマセ[△]ン[△]△

（「△」印は、その音節全体が強く、高められていることを示す（国立国語研究所1960））

「昇調1」に対する「昇調2」について、「昇調2とする音調形式はいわゆるプロミネンスが文末部の最終音節だけに加えられた発話の音調をさすものである」としている。ここでのプロミネンスについては、「ふつう、プロミネンスは文全体、または文の一部に加えられる強勢の要素と考えられるが、この昇調2は、文末の最終音節に加えられたプロミネンスであるがため、文末のイントネーションとしての意味を持ってくるものである。」と説明される。

「△」は音調記号であるため、具体的にどのようなイントネーションパターンを取るかは定かでないが、本研究の不満の答えに見られた起伏型のパターンとの共通性を見ることができそうである。

文末部分にプロミネンスが置かれることについては、より文法的な観点から文末の「の」に着目することでさらに明確になる。終助詞とイントネーションの関連についての研究は、轟木（1992）片桐（1997）や小山（1997）に見られるように盛んに行われているが、文末の「の」について詳しく言及されたものは少ない。本研究の「あわないの」における文末の「の」は、いわゆる「の（だ）」と同じ文法カテゴリーに属する。文末を問わず、文法形式としての「の（だ）」についての論考は数多く見られる（橋本1948、佐治1981、田野村1990、国広1992など）。それらの中でも野田（1997）では、「の（だ）」の機能において、「対人的モード」の「の（だ）」について触れ、「対人的「のだ」の文は、たとえば、話し手だけが知っていること

を「告白」するような場合、聞き手の知らないことを言い聞かせる「教示」、聞き手に認識してほしいことを「強調」する場合などに、よく用いられる」としている。ただし、野田（1997）でも述べられているように、「告白」「教示」、「強調」、といった用法は連続的であり、明確に分離するのは困難である。本研究における「の」の用法について見ると、「平静」の発話の場合、「教示」や「告白」の意味が強く、「不満」においてはそれらに加えて強調の意味が強くなると考えられる。

本研究の「あわないの」という発話において、それが不満の答えとして用いられる場合、相手に自分の状況を理解してもらいたい、という心情から、文末に「の」に対し音声的な「強調」が行われると解釈される。

5. 聴取実験

ここまでの音声の分析結果から、平静の答えと比べて文頭部分の高さ（F 0 値）が低いこと、また不満の答えでは文末部分に起伏型のイントネーションパターンが見られることが分かった。以上から、文末部分の起伏パターンとパラ言語情報の表現に影響を及ぼす文頭部分のイントネーションパターンが、平静の答えと不満の答えの弁別に対して大きな影響を及ぼすと仮定される。本研究では、以上の仮定に基づき、文頭部分や文末部分におけるイントネーション上の特徴が、不満の答えの弁別に与える影響を明らかにするため、合成音声を作成し聴取実験を行った。音声合成の手順と聴取実験の結果について以下に述べる。

5. 1. 音声合成

前節で述べた仮定に基づき、異なる2つの刺激系列を設定し音声合成を行った。1つは文末部分に起伏パターンを付与した刺激群（csy）と、もう1つは音声未まで直線的に下降する刺激群（dsy）

である。それぞれの刺激群においてF 0 値を操作する箇所は、文頭部分（以下 start）、上昇部分（同 peak）、文末部分（同 end）である。音声合成に際し、原音のF 0 曲線を抽出し、各部分のF 0 値を求めた。続いて各部分のF 0 値を10Hz 刻みで上下に変化させ、直線補間を行った。文末に起伏パターンを持つ csy の系列では、はじめに文末に起伏パターンを付与した上で、各部分のF 0 値を操作した。起伏の開始点から頂点までの上昇幅は65Hz、頂点からF 0 曲線の終端までの下降幅は66Hz に固定した。

合成音声の自然性を考慮し、各部分におけるF 0 値の変化量は、start と peak で+40Hz から-40Hz、end で+30Hz から-30Hz までとした。作成した刺激数は、start を操作した刺激8音×2（系列）、peak を操作した刺激8音×2（系列）、end を操作した刺激6音×2（系列）の44音に、パラメータを操作していないコントロール刺激1×2（系列）を加えた計46音である。図4から9に合成音声のF 0 曲線を示す。いずれも、縦軸は高さ（Hz）、横軸は時間長（msec）である。また中心の破線はコントロール刺激のF 0 曲線、その上に+40Hz の刺激のF 0 曲線、下に-40Hz の刺激のF 0 曲線を重ねて表示している（end の場合、上+30Hz、下-30Hz）。表1に刺激における各部分のF 0 値を示す。

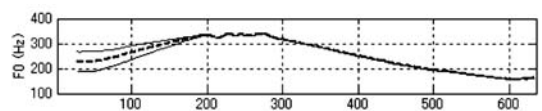


図4 startのF 0 値を操作した刺激のF 0 曲線

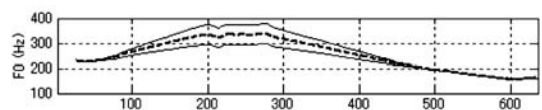


図5 peakのF 0 値を操作した刺激のF 0 曲線

「平静の答え」と「不満の答え」の弁別に対するイントネーションパタンの影響

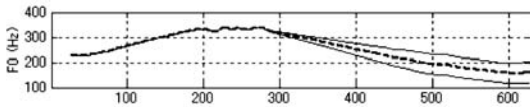


図6 endのF0値を操作した刺激のF0曲線

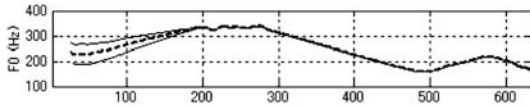


図7 文末に起伏パターンを付与し、startのF0値を操作した刺激のF0曲線

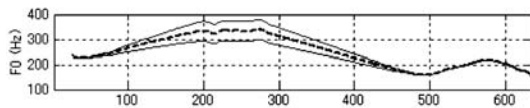


図8 文末に起伏パターンを付与し、peakのF0値を操作した刺激のF0曲線

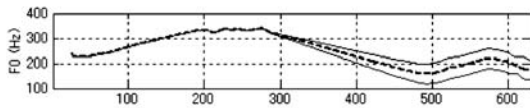


図9 文末に起伏パターンを付与し、endのF0値を操作した刺激のF0曲線

	start	peak	end
(+40)	240	372	—
(+30)	230	362	189
(+20)	220	352	179
(+10)	210	342	169
control	200	332	159
(-10)	190	322	149
(-20)	180	312	139
(-30)	170	302	129
(-40)	160	292	—

表1 刺激における各部分のF0値

5. 2. 被験者

聴取実験には、音響学、音声学の専門的な知識を有していない、10代から40代までの日本語母語話者17名を採用した。九州地方で育成した被験者が中心であった。

5. 3. 実験手順

作成した刺激の46音をランダム配列して1セットとし、計5セットをDATに収録して音声提示用のテープを作成した。実験は静謐な部屋で行われ、音声はヘッドホンを通して聴取しやすい音圧レベルで被験者の両耳に提示された。被験者はヘッドホンから流れる音声を聞いて、それが平静の答えか不満の答えのどちらに聞こえたかを、用意された回答用紙に記入した。

5. 4. 仮説

実験結果を述べる前に、本研究における仮説を示す。平静の答えと不満の答えの分析結果から、不満の答えは(1)「問いかけ」と比較してstartのF0値が低いこと、(2)endに起伏パターンが見られることが明らかになった。また田川他(2000)、田川他(2001a)では、「問いかけ」の音声を用いて実験を行った結果、startのF0値を下げた場合、安定して非難の問いかけと知覚された。一方で、peakのF0値を操作しても、安定して非難の問いかけとは知覚されなかった。以上から次の仮説が導かれる。以下の仮説を踏まえ、次節に聴取実験の結果を示す。

- i : startのF0値を原音よりも下降させた場合、不満の答えとして知覚される。
- ii : endに山型の起伏パターンを付与した場合、不満の答えとして知覚される。
- iii : endに起伏パターンを付与し、なおかつstartのF0値を下げた場合、不満の答えの選択率はさらに上昇する。反対にstartのF0値を

田 川 恭 識

上げた場合、選択率は減少する。

6. 実験結果

文末に起伏パタンを付与した csy の結果について図10から12に示す。いずれも全被験者の平均である。縦軸は、刺激に対して被験者が不満の答えを選択した割合、横軸は刺激における F 0 の変化量を示している。◆はプラス方向に F 0 値を変化させた刺激を、また■はマイナス方向に変化させた刺激を表している。図10は start の F 0 値、図11は peak の F 0 値、図12は end の F 0 値を操作

した刺激の結果である。

6. 1. csy の結果

全体の平均で見ると、文末に起伏パタンを付与した csy の系列では、いずれの刺激も不満の答えの選択率がほぼ90%以上である。start、peak、end を操作したいずれの刺激でも、(+)、(-) に関わらず、変化量との間に相関は見出せない。

続いて、図13から15に、文末に起伏パタンの無い dsy の実験結果を示す。図13は start の F 0 値、図14は peak の F 0 値、図15は end の F 0 値を操作した刺激の結果である。

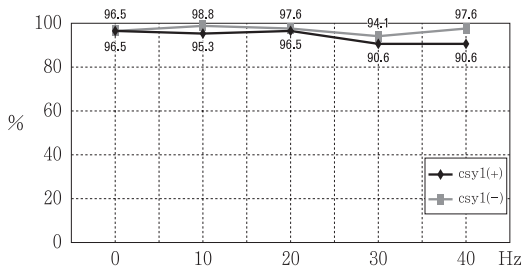


図10 起伏パタンを付与し、startの F 0 値を操作した刺激に対する実験結果 (csy1)

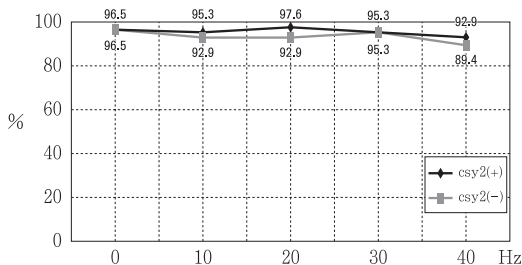


図11 起伏パタンを付与し、peakの F 0 値を操作した刺激に対する実験結果 (csy2)

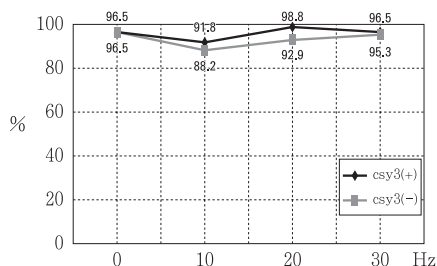


図12 起伏パタンを付与し、endの F 0 値を操作した刺激に対する実験結果 (csy3)

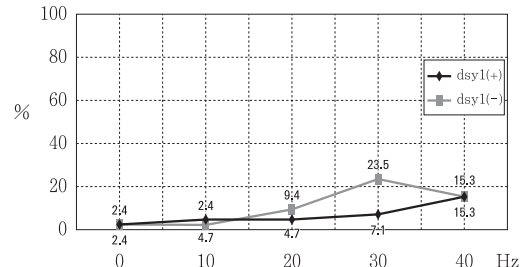


図13 起伏パタンを付与せずに、startの F 0 値を操作した刺激に対する実験結果 (dsy1)

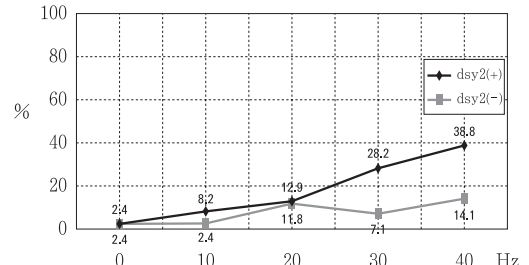


図14 起伏パタンを付与せずに、peakの F 0 値を操作した刺激に対する実験結果 (dsy2)

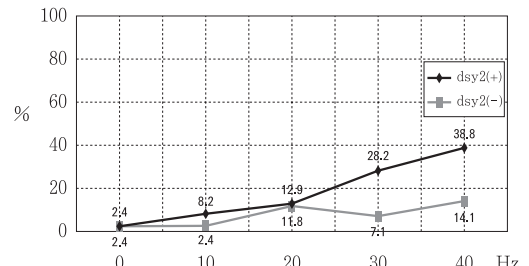


図15 起伏パタンを付与せずに、endの F 0 値を操作した刺激に対する実験結果 (dsy3)

6. 2. dsy の結果

start の F 0 値を操作した dys1 の場合、(－) と (+) の両系列とも、20Hz までは選択率はほぼ横這いであるが、(－) の系列では次の 30Hz から 40Hz にかけてわずかに選択率の上昇が見られる。(+) の系列においては、(－) の系列よりもやや早く、20Hz から 40Hz にかけてわずかに選択率の上昇が見られるが、その後、40Hz に至って逆に選択率が下がり、(－) の系列と同程度の選択率に収まっている。

続いて peak の F 0 値を操作した dsy2 の場合、(+) の系列において変化量が増加することに不満の答えの選択率が増加している。(－) の刺激では、変化量の増加に対して選択率はほぼ横這いで、すべて 20% 以下に収まっている。

最後に end の F 0 値を操作した dsy3 の場合、(+) と (－) の両系列とも、変化量が増加するに従い漸次上昇している。しかし同じ変化量の刺激では、(+) よりも (－) の方が倍以上の選択率となっている。

7. 考察

7. 1. csy の考察

本研究では、実験を行うにあたり、以下のような仮説を設定した。

- i : start の F 0 値を下降させた場合、不満の答えとして知覚される。
- ii : end に起伏パタンを付与した場合、不満の答えとして知覚される。
- iii : end に起伏のパタンを付与し、なおかつ start の F 0 値を下げた場合、不満の答えの選択率はさらに上昇する。反対に start の F 0 値を上げた場合、選択率は減少する。

end に起伏のパタンを付与した csy の実験結果は、ii の仮説を支持するものである。一方で start

の F 0 値を上下させても、両者の選択率は同程度の水準であり、このことから iii の仮説は支持されない結果となった。また start の以外の箇所、すなわち peak や end の F 0 値を操作しても選択率に大きな変化は見られない。以上の結果は、end における起伏パタンに対して被験者は、有か無かといったカテゴリカルな反応を示すことを表している。

7. 2. dsy の考察

文末が直線的に下降する dsy の刺激の場合、全体的に不満の答えの選択率が低い。しかしながら、peak の F 0 値を上昇させた刺激 (dsy2 (+)) や end の F 0 値を下降させた刺激 (dsy3 (－)) では、F 0 値の変化量が増加するに従い、不満の答えの選択率が上昇する傾向が窺える。上に示した結果はいずれも被験者全体の平均であるが、csy に比べて dsy の結果は被験者間でばらつきが大きかった。特に dsy2 (+) と dsy3 (－) に対しては、F 0 の変化量に応じて選択率が直線的に増加する被験者と、ほとんど反応を示さない被験者が見られた。これらのことから、peak や end の F 0 値が不満の答えの知覚に与える影響は、個人間の違いに左右され、安定した要因とならないことが窺える。

7. 3. 総合考察

本研究の実験結果から、文頭部分の高さは不満の答えの選択率に大きな影響を及ぼさないことが分かったが、これに対し、先行研究で対象とした疑問文の「あわないの？」では、「非難の問いかけ」と知覚されるには、文頭部分の高さが大きな要因となることが明らかになった。先の研究で作成した合成音声と、本研究での合成音声とを比較すると、本研究の方が F 0 値の変化量が少ないため、厳密な比較は困難である。しかし、全体的な

傾向として疑問文における「平静」と「非難」の弁別の場合、文頭部分の高さやパターンが重要な役割を果たすが、平叙文における「平静」と「不満」の弁別には、文末部分のイントネーションパターンが中心的な役割を果たすと考えられる。

疑問文ではパラ言語情報の有無に関わらず、文末にかけてF0が上昇する。文末にかけての上昇は、それが疑問であることを示す指標として要求され、形状そのもののバリエーションは少ない。一方、平叙文では、文末にかけて声の下降が生じるが、F0値という物理的表現からは、必ずしも「下降」とは言えず、むしろ「非上昇」とした方が現実的である。従って、聞き手が上昇と感じない程度の範囲内では、F0の下降幅や下降の際の動態について制約を受けにくい。結果的に文末におけるF0パターンのバリエーションは、疑問文よりも多くなり、これらのバリエーションによって話し手の感情を表現することが可能となるのではないか。一方、疑問文ではそれが不可能であるため、その不足を補うために、文頭部分の高さやイントネーションパターンが重要な意味を持つようになると考えられる。これに対し、不満の答えでは文末の起伏パターンによって「不満」の意図が表現されるため、文頭部分や上昇部分の高さやイントネーションパターンは大きな影響を及ぼしにくいと推測される。

8. まとめと今後の課題

本研究では平叙文の「あわないの」という発話を対象に、「答え」と不満の答えの弁別に対するイントネーションパターンの影響を明らかにするため、合成音声を用いて聴取実験を行った。その結果、「あわないの」という発話の場合、それが不満の答えとして知覚されるには、文末部分における起伏パターンの存在が重要であることなどが明らかになった。

文末に起伏パターンの無い刺激の系列では、上昇部分や文末部分の高さ（F0値）を操作した場合、被験者間で異なる傾向が見られた。被験者間で傾向差が生じる理由としては、方言差や性差、年代差といった多種の要因の影響が考えられるが、研究方法についても考慮の対象とすべきであろう。本研究の聴取実験では、提示される刺激音声に対し平静の答えと不満の答えという2つの選択肢からどちらかを強制選択させるという手法を採用した。しかし一連の刺激には、被験者の知覚領域においてどちらにも当てはまらない音声があった可能性も否定できない。以上については、筆者らが最近行った調査で、上記の2つの選択肢に「分からない」という項目を加えて、東京方言話者を中心に判断の傾向を見た。その結果、本稿での結果とほぼ同様の傾向が見られた。以上は、「平静」と「不満」の判断について範疇的な知覚を行なっていることを裏付けるものである。しかし、前川・北川（2002）で述べられているように、「パラ言語情報の強度」という問題もある。つまり被験者の知覚領域の中では「平静」と「不満」のカテゴリーが確立されているものの、各カテゴリーの中でより「平静らしい」、あるいはより「不満らしい」音声の存在している可能性もある。これらの問題の解決には、聴取実験において、段階評価等を行う、などの改善策が期待される。また本研究においては、10代から40代までの被験者を対象とした。より厳密な結果を導くには、40代以上の被験者についても対象とすべきである。最後に、本研究では対象としたのは「あわないの」という発話であり、他の発話についても検証して行く必要がある。いずれも今後の課題とするものである。

付記：本稿の執筆に当たり、2名の査読者の方々に有益なご指摘を頂きました。記して感謝致します。

「平静の答え」と「不満の答え」の弁別に対するイントネーションパタンの影響

※本稿は、2004年4月4日から9日に京都国際会館で開催された The 18th International Congress on Acoustics (ICA 2004) で行った口頭発表の原稿をもとに加筆・修正を加えたものである。

参考文献

- 上村幸雄「日本語のイントネーション」『ことばの科学 3』言語学研究会編, 193-220, 東京: むぎ書房.
- 上野善道 (1977)「日本語のアクセント」『岩波講座日本語 5 音韻』, 281-321, 東京: 岩波書店.
- 梶原綾乃 (2003)「留学生と日本人学生との交流促進を目的としたコミュニケーション教育の実践」『日本語教育』117, 93-132.
- 片桐恭弘 (1997)「終助詞とイントネーション」, 『文法と音声 I』音声文法研究会編, 235-256, 東京: くろしお出版.
- 川上夔 (1956)「文頭のイントネーション」『日本語アクセント論集』所収, 61-75, 東京: 汲古書院.
- 河原英紀 (1998)「自然性の極めて高い音声分析合成変換法」『音声研究』, 28-36.
- 国広哲弥 (1992)「「のだ」から「のに」・「ので」へー「の」の共通性ー」『日本語研究と日本語教育』, 名古屋大学出版会.
- 郡史郎 (1988)「強調とイントネーション」『講座日本語と日本語教育 2 日本語の音声・音韻 (上)』, 316-341, 東京: 明治書院.
- 郡史郎 (2004)「イントネーション」『朝倉日本語講座 3 音声・音韻』, 109-131, 東京: 朝倉書店.
- 国立国語研究所 (1960)『話しことばの文型 (1) 一対話資料による研究一』, 東京: 秀英出版.
- 国立国語研究所 (1963)『話しことばの文型 (2) 一独和資料による研究一』, 東京: 秀英出版.
- 小山哲春 (1997)「文末詞と文末イントネーション」『文法と音声 I』音声文法研究会編, 97-119, 東京: くろしお出版.
- 佐治圭三 (1981)『『～のだ』の本質』『日語学習与研究』所収『日本語の文法の研究』1991, 東京: ひつじ書房.
- 杉藤美代子 (1990)「プロソディーとは何か」『月刊言語』9: 8, 16-21.
- 杉本隆 (2000)「音声中の感情表現に関連する物理量とその制御に関する研究」, 北陸先端科学技術大学院大学情報科学研究科修士論文.
- 橋本進吉 (1948)「国語法要説」『国語法研究』, 東京: 岩波書店.
- 平館郁雄・赤木正人 (2002)「怒りの感情音声における音響特徴量の分析」『日本音響学会聴覚研究会資料』32: 2, H-2002-17.
- 藤崎博也 (1996)「韻律研究の諸側面とその課題」『音響学会講演論文集』, 287-290.
- 大坊郁夫・安藤清志・池田謙一他編 (1990)『社会心理学パースペクティブ 2 一人と人を結ぶとき』東京: 誠信書房.
- 田川恭識 (2002)「イントネーションと感情表現ー文末詞「の」の機能とイントネーションー」『国文研究』48, 82(1)-66(19), 熊本県立大学日本語日本文学会.
- 田川恭識・田淵咲子・甲斐朋子・宮園博光・馬場良二 (2000)「発話の感情知覚に及ぼす韻律的要因」『日本音響学会聴覚研究会資料』H-2000-113.
- 田川恭識・田淵咲子・甲斐朋子・宮園博光・馬場良二 (2001a)「発話の感情知覚に及ぼす韻律的要因の影響」『第15回日本音声学会全国大会予稿集』, 25-26.
- 田野村忠温 (1990)『現代日本語の文法 I 「のだ」の意味と用法』, 東京: 和泉書院.
- 田淵咲子・甲斐朋子・馬場良二 (1999)「文末「ノ」の音声的特徴」『日本音響学会聴覚研究会資料』, MA99-54.
- 轟木靖子 (1992)「東京語の文末詞の音調と形容詞・動

田 川 恭 識

- 詞のアクセントについて」『STUDIUM』, 大阪外国語大学大学院研究室.
- 野田春美 (1997) 『Frontier series 日本語研究叢書 9 「の(だ)」の機能』, 東京: くろしお出版.
- 播磨桂子 (2001) 「対話コミュニケーション」『音声言語とコミュニケーション』中村萬里・永淵道彦編, 81-95, 東京: 双文社出版.
- 前川喜久雄 (1998) 『岩波講座 言語の科学 2 音声』, 東京: 岩波書店.
- 前川喜久雄・北川智利 (2002) 「音声はパラ言語情報をいかに伝えるか」『認知科学』9: 1.
- 水谷修・鮎澤孝子 (1991) 『シンポジウム日本語音声ー韻律の研究と教育をめぐるー』, 東京: 凡人社.
- 森大毅・相澤宏・粕谷英樹 (2005) 「対話音声のパラ言語情報ラベリングの安定性」『日本音響学会誌』61: 12, 690-697.
- Carl E. Willams and Kenneth N. Stevens (1972) "Emotion and Speech: Some Acoustical Correlates" The Journal of the Acoustical Society of America, 52:2, 1238-1250.
- Crystal, David (1969) "Prosodic Systems and Intonation in English", Cambridge: Cambridge University Press.
- D. Robert Ladd (1985) "Evidence for independent function of intonation contour type, voice quality, and F0 range in signaling speaker affect." The Journal of the Acoustical Society of America, 78:2, 435-444.
- Iain R. Murray and John L. Arnott (1993) "Toward the simulation of emotion in synthetic speech: A review of the literature on human vocal emotion." The Journal of the Acoustical Society of America, 93:2, 1097-1108.
- Maekawa, Kikuo (1998) "Phonetic and Phonological Characteristics of Paralinguistic Information in Spoken Japanese.", Proceedings of the 5th International Conference on Spoken Language Processing (ICSLP), 635-638.
- Pittam, Jeffery (1994) "Voice in social interaction: An interdisciplinary approach", Language and Language Behaviors volume 5, SAGE Publications.