

自由対話における発話の共感状態の音声分析 Speech Analysis of Empathic States in Spontaneous Utterances

高亜罕[†] 大野ゆう子[†] 林良彦[‡]
Gaoyahan[†] Yuko Ohno[†] Yoshihiko Hayashi[‡]

論文概要

共感状態に関する評定が付与された自由対話の発話データの統計的な分析を行い、共感状態を規定する要素と音声特徴との関係を検討した結果を報告する。本研究における共感状態とは、対話中の先行する発話に反応する発話に対して定義され、(a) 反応対象の発話の話題への態度、(b) 反応対象の発話への感情の一致状況、という2つの情報の組として表現される。分析の結果、対話音声から発話者の態度や感情を推定するために有用と考えられるいくつかの音声特徴が明らかとなった。一方で、発話の反応先を同定することの必要性も明らかとなり、音声以外のモードからの情報を統合する必要性が示唆された。これらの結果は、対話の音声データから発話者の共感状態を推定する計算機構の基盤を与える。

Abstract

This paper discusses characteristics of empathic states expressed in spontaneous utterances. We basically define the empathic state of an utterance as a combination of attitude and emotional state that are expressed toward the preceding utterance to which the target utterance is directed. Among the possible combinations, we define "empathy" as a special empathic state, where the attitude and the emotional state are identical. Through a series of statistical analyses of speech data collected from three-party spontaneous dialogues, we show that some empathic states are identifiable by using a range of speech parameters. These results serve as solid foundation on which a computational mechanism for estimating a speaker's empathic states can be developed.

1. はじめに

円滑な対話の進行のためには、適切な場面やタイミングで相手への「共感」を示すことが重要である。対話における共感とは、表情、身振り、話し方、発話内容などによって表出され、相手に伝達される。共感の捉え方や定義はさまざまであるが、本研究における共感とは以下に述べる「共感状態」に基づいて定義される。

すなわち、本研究における共感状態とは、対話中の先行する発話 X に反応する発話 Y に対して定義され、(a) X の話題内容への「態度」、(b) X と Y の発話に込められた「感情」の一致状況、という2つの情報の組として表現される。この中で、特に態度と感情の両方が一致している共感状態を「共感」と呼び、他の共感状態と区別する。

本論文では、共感とそれ以外の共感状態における発話の音声特徴の差異について比較検討した。より具体的には、以下の2段階の分析を行った。

- **単一の発話を単位とする分析:** これにより、各態度と感情の表出の音声的な特徴を比較した。
- **発話の反応対象を考慮に入れた分析:** 分析の対象である発話区間 Y とその反応対象である先行する発話 X における感情の一致状況から、共感状態を「共感」のほか、「同意」、「同感情」、「感情一致」、「感情不一致」に分類し、これらの間の音声的な特徴を比較分析した。

これらに対する分析結果の詳細は本文中に述べるが、音声情報から共感状態を推定する計算機構 [高亜罕 2010b]

[†] 大阪大学大学院医学系研究科 Graduate School of Medicine, Osaka University.

[‡] 大阪大学大学院言語文化研究科 Graduate School of Language and Culture, Osaka University

の基盤を与えるものである。

本論文の構成は以下のとおりである。まず、2節で本研究における共感状態と共感の定義を示し、3節で対話における共感状態について説明する。次に、4節で今回の分析対象である対話データとその共感状態の評定、音声分析について述べる。以上の準備に基づき、5節では単一の発話を単位とする分析結果を示す。すなわち、各発話における態度と感情の間の依存関係の分析結果、音声的特徴の分析結果を示す。さらに、6節では発話の反応対象を考慮に入れた分析の結果を示す。すなわち、共感状態と話題の展開との関係、「共感」とその他の共感状態の音声的特徴の比較について論じる。7節では関連研究を参照しながら、今回得られた知見や可能性について議論し、8節でまとめを行う。

2. 共感状態と共感

2.1 共感状態の下位分類としての共感

「共感」の捉え方や定義はさまざまである。良く知られている共感の定義には、(1) 他人の考え、感情、行動を正確に認知、予測ができること、(2) 他人の感情的な経験を自分のことのように感情的に反応すること、の2つがある [Mehrabian 1972]。前者は認知的側面、後者は感情的側面に重きをおいた定義であるが、いずれも内的な心的状態として捉えられている。

本研究では特に、コミュニケーションの円滑な進行という側面から、コミュニケーションの相手・発話に対する態度、および、自己の感情の表出に着目する。すなわち、対話参加者の心的な状態には少なくとも、話題に対する自身の受け止め方(態度)と、それをどのように感じているか(感情)という側面があり、これを適切に表出することにより、コミュニケーションの相手に共感、または、

非共感の「共感状態」を伝えていると考える。すなわち、2.2に示すように、共感を共感状態の一つの下位分類として定義する。

共感状態の表出はさまざまなモードにおいてなされることが考えられるが、本研究では特に対話参加者の発話音声における特徴に着目する。音声認識・言語処理により自由対話における発話の言語的内容を得ることは現状では困難であるのに対し、音声特徴は比較的頑健に取得することが可能であると考えられ、対話参加者の音声情報からその共感状態を推定する[高亜罕 2010b]ことができれば、他のモードからの情報と統合することにより、対話の「場の活性度」の推定[前田 2010]に寄与することが期待される。また、患者を対象とする病院の相談業務[高亜罕 2009]などでは、必要に応じて患者の話に共感し、それを適切に伝えることが重要であるが、そのための指針を得られることが期待される。

2.2 共感状態の定義

本研究における「共感状態」とは、対話中の先行する発話 X に反応する発話 Y に対して定義され、(a) X の話題内容への「態度」、(b) X と Y の発話に込められた「感情」の一致状況、という2つの情報の組として表現される。表1に本研究における共感状態の定義を示す¹。

表1: 共感状態の定義

態度	感情状態の一致状況	共感状態
ポジティブ (一致)	一致	共感
	不一致	同意
ネガティブ (不一致)	一致	同感情
	不一致	完全不一致
中立	一致	感情一致
	不一致	感情不一致

この中で特に、態度と感情の両方が一致している共感状態を「共感」と呼び、他の共感状態と区別する。ここで、共感に感情の一致まで条件付けることに関しては、議論がありえる。例えば、ある発話に対して、それに喚起された別の感情を表出する場合(例: 悲しみ→怒り)も共感に含めるべきかもしれない。ただし現状では、このような感情相互の反応関係は明らかではないため、現行の共感の定義を用いることとする。

2.3 態度と感情の分類カテゴリ

ここで、本研究における態度と感情の分類カテゴリについて整理しておく。

まず「態度」とは、対話中の先行する発話 X に反応する発話 Y に対して定義され、発話 Y の発話者が発話 X の話題内容に対して同じ意見・考えを持つかどうかによって以下のように下位分類される。

- **ポジティブ**: Y の発話者の意見・考えが発話 X の話題内容における意見・考えと同じ
- **ネガティブ**: Y の発話者の意見・考えが発話 X の話題内容における意見・考えと異なる

¹[高亜罕 2010a]では、本稿における「同意」を「同感」としていた。同様に、「感情」は「感情状態」としていた。

- **中立**: Y の発話者が発話 X の話題内容に対して、特別な意見・考えを示さない

さらに、全く新規の話題を提示する場合には、反応対象の発話が存在しないと考えるので、態度を定義することができない。この場合の態度を「Φ」と表す。

次に、「感情」とは、各発話に表出された発話者の感情である。本研究では、[Ekman 1992]の提唱した6種類の基本感情分類カテゴリに「平静」を加えた7種の分類カテゴリを用いる。

- **Ekman の基本感情分類**: 「喜」、「怒」、「悲」、「恐」、「嫌悪」、「驚」
- **平静**: 発話に特定の感情が表出されていない

3. 対話における共感状態

3.1 分析の単位と話題の流れ

対話音声データの分析を行うためには、その基本的な単位を定める必要がある。本研究においては、何らかの意図を表していると判断できる発話の時間区間を分析の基本単位とし、これを「発話区間」と呼ぶ。

対話中の各発話は、対話の話題の流れという観点から分類すると、以下のように分類できる。これを発話の「話題機能」と呼ぶ。

- **新規話題提示**: 対話の冒頭で新しい話題を提示する。あるいは、対話の途中で現在の話題に関連しない話題に推移する。
- **話題展開**: 現在の話題に関連した話題に推移する。
- **話題継続**: 現在の話題を維持する。

同一の話題が連続する区間を「話題区間」と呼ぶ。すなわち、新規話題提示/話題展開の発話から次の新規話題提示/話題展開の発話が現れる区間における発話区間の集合が話題区間である。

3.2 対話における話題の流れと共感状態

以下、図1に示す対話進行を例に、対話における共感状態について説明する。ここでは、ある話題区間が4つの発話から構成されている。

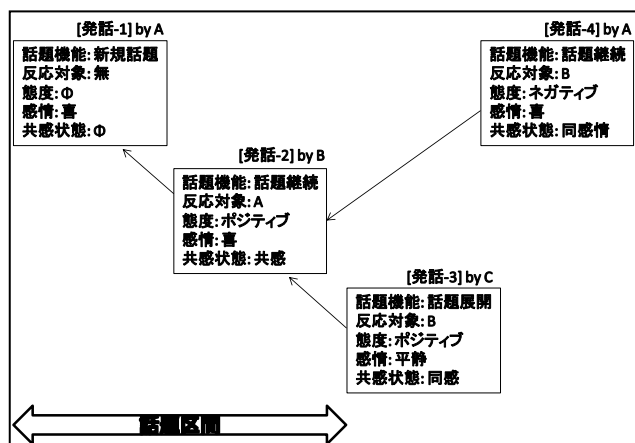


図1: 対話進行の例

- **発話-1:** まず話者 A が新規の話題を提示する。新規話題の提示は、先行する発話への反応ではないため、態度は定義されず、従って共感状態も定義されない。
- **発話-2:** 次に話者 B がこの話題に反応して特に新たな話題を提示することなく発話を行う。すなわち、話題継続の発話である。この発話の態度は「ポジティブ」(A の先行する発話と同一の考え・意見)であり、さらに感情は「喜」である。この感情は A の先行発話と一致する。よって、共感状態は「共感」となる。
- **発話-3:** C の発話は先の B の発話への反応で、態度は「ポジティブ」であるが、感情が「平静」であり B とは異なるため、その共感状態は「同意」となる。
- **発話-4:** 先の B の発話には、A も反応しているが、B の発話とは異なる考え・意見を表明したものであり、態度は「ネガティブ」となる。一方で感情は「喜」で一致する。このため、共感状態は「同感情」となる。

4. 対話データの評定と分析

4.1 対話音声データ

今回の分析の対象としたのは、「場の活性度」に関する研究のために行われた3人会話実験[前田 2010]により得られた対話音声データである。会話実験の参加者は初対面の大学生計24名(男性15名: 21.6±1歳, 女性9名: 20.9±1.6歳)である。図2に対話状況を示す。



図2: 会話実験の様子

対話参加者は、ヘッドセットを装着し、比較的自由に動ける状況下で会話をを行った。

これら24名の参加者を同性3名による8つの組に分け、各組について、4分間の対話を2回行った。これにより、各組8分間、計64分間の対話音声データを得た。

4.2 共感状態に関する評定

4.2.1 評定作業

音声分析ツール Praat²を用いて対話音声データの評定を行った。まず、音声データの発話区間へのセグメンテーションを行い、発話内容の書き起こしを付与した。ここでは、ひとつの意見や考えを表出していると考えられる時間区間を発話区間として認定した。図3に音声分析フリーソフト Praat 上のタグ付けられた発話区間の例を示す。

次に、得られた各発話区間に話題機能と共感状態に関する評定を付与した。3人の評定者による評定作業を行い、

多数決により分類カテゴリを決定した。多数決の取れないものは、「不一致」とした。また、態度と感情の双方について、あらかじめ用意したカテゴリに適切なものがない場合は「不明」、音声の状態によって判定が困難な場合は「判定不能」というラベルを付与した。

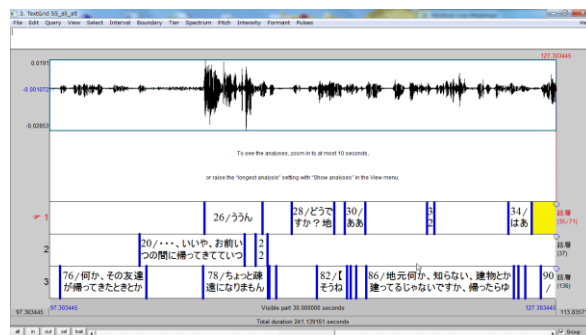


図3: Praat 上のタグ付き発話区間

なお、評定は音声のみに基づいて行い、対話中の録音データは用いていない。また評定者は、評定作業の事前

に評定基準を合わせるための意識合わせを行っている。共感状態に関する評定項目は、話題提出の有無、反応対象者、態度の表出、感情の表出の4種類である。話題提出の有無と反応対象者を合わせることで、3.1節に示した発話の話題機能を定めることができる。

今回の対話は三人対話であるため、ある発話がどの発話に対する反応であるかを一意に明確に定めることが困難な場合も想定される。例えば、Aの発話を契機にBがCへ向けて発話を行うような場合が有り得る。今回の評定においては、発話が誰に向けられているかではなく、どの発話を契機とするものかという観点から反応対象を評定した。なお、[藤本 2006]では、「質問・応答」の直後に注目し、質問者が応答に対して行う態度表明を分析したのに対し、本研究では、そのような限定は行っていない。

4.2.2 評定結果の概要

今回の対話音声データからは、合計2,241³の発話区間、448の話題区間が得られた。発話区間の平均継続長は1.933 sec (標準偏差: ±2.511), 話題区間の平均継続長は8.549 sec (標準偏差: ±6.598)であった。各評定項目に関する評定結果の概要を以下に示す。

- **話題提出の有無:** 全発話区間で話題提出「無」の発話区間が79%、「有」が20%であった。平均すると5つの発話ごとに1つの話題提出が行われていることになる。判定の不一致は少なく、話題提出の有無は安定的に判定できた。話題提出「有」の内訳は、新規話題提示: 4%、話題展開: 16%であった。

- **態度の表出:** 観察された態度は、「中立」が71%、ポジティブが19%であったのに対し、「ネガティブ」が3%と少なかった。これは、今回の対話では賛否を問う議論になるような話題は少なかったためと考えられる。不一致は2%と比較的少なく、態度は安定的に判定できるものと考えられる。

³ただし、この中の137の発話区間については、反応先が不明となったため、6節の分析では、全2,104の発話区間を分析対象とした。

² <http://www.fon.hum.uva.nl/praat/>

● **感情の表出:** 観察された主な感情は、「平静」(61%)、「喜」(23%)、「驚」(12%)の3種類であり、これら以外の出現はごく少なかった。これも、初対面の対話者が当たり障りのない話題について話すという今回の対話の特性によると考えられる。不一致は2%と比較的少なく、感情もかなり安定して判定できると考えられる。

4.3 発話区間の分析に用いる音声的特徴

本研究の最終的な目的は、観測できる音声データから発話に表出された共感状態を推定する枠組みを確立することにある。このため、評価が付与された発話区間における音声的特徴を求めため、先行研究([Banse 1996], [Cornelius 1996], [Murray 1993])を参考にして基本的な音声分析を行った。今回の分析においては、各音声区間について、以下の3種別、7つの音声パラメータに対する値を求めた。この中で、基本周波数と音圧は、Praatの音声信号分析機能により求めた。

- **声の高さ(基本周波数 F0):** 最大値(以下, Max_F0), 最小値(Min_F0), 平均値(Mean_F0)
- **声の大きさ(音圧):** 最大値(Max_db), 最小値(Min_db), 平均値(Mean_db)
- **発話速度(1秒あたりのモーラ数):** 発話の継続時間と書き起こしのローマ字表記より求めた

先行研究[西尾 2005]によると、発話音声は個人、年齢、性別の影響を受けるため、今回は個人ごとの態度=「中立」、感情=「平静」の場合の各パラメータの値を基準値とし、各発話区間における測定値の基準値からの変動量を算出して統計分析に用いた。ここで、個人の特徴からの影響を最小限にすることを目的とし、基本周波数では比、音圧では差、発話速度では比を利用した。

以上により得られた対話データに対して統計分析⁴を行うことにより、単一の発話を単位とする分析(5節:各態度と感情の表出の音声的な特徴の比較)、および、発話の反応対象を考慮に入れた分析(6節:共感状態間での音声的な特徴の比較)を行った。

なお以下の分析では、声の高さや大きさといった音声パラメータの値以外に男女で共感状態の表出の仕方が異なる可能性を考慮し、男女のグループに分けて比較を行った。

5. 単一の発話単位を対象とする分析

5.1 態度・感情の間の依存性

感情と態度の組み合わせにより共感状態を推定できると仮定したが、これらは独立なものか、あるいは、どのような相互依存性があるかをカイ2乗検定により調べた。ただし感情は、ある程度の出現回数を得られた「喜」と「驚」に限定した。その結果を表2に示す。表2におけるカッコ外の数字は実頻度、カッコ内の数字は頻度の期待値を示す。

表2からは以下のことがいえる。

- 態度=「ポジティブ」の場合に感情=「喜」が有意に多く(このため、感情=「驚」が有意に少ない)、態度=「中

立」の場合には有意に少ない。前者は、主に相手の意見・考えと同じ態度を持った時に表出されやすい傾向を示している。

- 態度=「ネガティブ」の場合に感情=「平静」が有意に少ない。これは、相手と異なる態度・考えを示す時に何らかの感情が伴いやすくなるという通常の会話の特徴の現れと考えられる。

- 感情=「驚」が態度=「中立」のときに有意に多い。これは、意外な情報を受けた時には態度を示さない、あるいは、示せないからと考えられる。後で議論するように、「驚」という感情には、発話者の知識状態などの心的状況との関連性が含まれていると考えられる。

- 態度=Φは新規の話題が提示される時であり、その時、感情=「平静」で提示する傾向がみられる。また、態度=Φのときは感情=「驚」が有意に少ない。これは、新規の話題を驚きながら提示することはまれであるという通常の会話の特徴を反映していると考えられる。

表2: 感情と態度のカイ2乗検定の結果

態度	平静	喜	驚
中立	979(989.5)	327(372.1)**	256(200.4)**
ポジ	276(262.3)	131(98.6)**	7(53.1)**
ネガ	28(41.8)**	29(15.7)**	9(8.5)
φ	65(54.5)*	20(20.5)	1(11)**

(有意水準: *0.05, **0.01)

以上をまとめると、感情と態度は必ずしも独立ではなく、態度=「ポジティブ」の時、感情=「喜」が多く、感情=「驚」は少ないという偏りが見られた。「ポジティブ」という態度は、当事者が話題提出者の意見・考えに同意している場合であるから、上記のような傾向になる。

5.2 態度・感情の音声的な特徴

態度・感情の組み合わせによって音声の表出が異なっているかを確認するために、性別の影響を考えた2元配置分散分析を行った。結果を表3と表4に示す。

表3: 音声特徴量を感情と性別を因子とする2元配置分散分析した結果(態度別) (有意水準: *0.05, **0.01, ***0.001)

態度	従属変数	固定因子		交互作用
		感情(P値)	性別(P値)	感情*性別(P値)
中立	Max hz	5E-13***	1.E-05***	0.051*
	Mean hz	4.E-07***	0.413	1.E-04***
	Min hz	<2E-16***	5.E-11***	1.E-06***
	Max db	<2E-16***	0.196	3.E-05***
	Mean db	<2E-16***	4.E-05***	1.E-07***
	Min db	0.002**	0.010**	2.E-04***
	Mora/s	<2E-16***	0.012*	7.E-04***
ポジ	Max hz	1.E-06***	0.791	0.004**
	Mean hz	1.E-05***	1.E-05***	0.005**
	Min hz	0.088*	0.536	0.164
	Max db	<2E-16***	0.988	0.573
	Mean db	<2E-16***	0.013*	0.523
	Min db	0.001***	0.092	0.219
	Mora/s	0.0187*	2.E-06***	0.200
ネガ	Max hz	0.719	0.033*	0.880
	Mean hz	0.377	0.042*	0.242
	Min hz	0.030*	4.E-05***	8.E-04***
	Max db	0.002**	0.688	0.481
	Mean db	0.005**	0.303	0.248
	Min db	0.322	0.939	0.940
	Mora/s	0.009**	0.018*	0.523
無 (新規話題 提出)	Max hz	0.829	0.272	0.931
	Mean hz	0.199	0.024*	0.872
	Min hz	0.266	0.827	0.374
	Max db	0.031*	0.012*	0.379
	Mean db	0.111	0.722	0.639
	Min db	0.342	0.163	0.187
	Mora/s	0.603	0.415	0.877

⁴ R (<http://www.r-project.org/>) を利用した。

表4: 音声特徴量を態度と性別を因子とする2元配置分散分析の結果(感情別) (有意水準: *0.05, **0.01, ***0.001)

感情	従属変数	固定因子		交互作用
		態度 (P値)	性別 (P値)	態度*性別 (P値)
平静	Max_hz	2E-06***	0.107	0.067
	Mean_hz	0.001***	2E-05***	0.003**
	Min_hz	0.229	0.244	0.104
	Max_db	1E-06***	0.105	0.528
	Mean_db	6E-06***	0.149	0.401
	Min_db	2E-05***	0.886	0.931
喜	Mora/s	0.003**	0.060	0.001***
	Max_hz	0.330	2E-06***	0.998
	Mean_hz	0.021*	0.050*	0.183
	Min_hz	9E-05***	9E-16***	5E-04***
	Max_db	0.023*	0.617	0.332
	Mean_db	4E-08***	0.013*	0.919
驚	Min_db	0.002**	0.009**	0.581
	Mora/s	0.284	3E-06***	0.867
	Max_hz	0.778	0.040*	0.129
	Mean_hz	0.834	0.013*	0.180
	Min_hz	0.437	0.003**	0.848
	Max_db	0.219	6E-06***	0.912
驚	Mean_db	0.427	1E-08***	0.880
	Min_db	0.757	8E-06***	0.905
	Mora/s	0.805	0.007**	0.195

表3と表4から、(a) 態度ごとに音声パラメータの値が感情間で差がある、(b) 感情ごとに音声パラメータの値が態度間で差がある場合が多いことが分かる。また、(c) 性別からの影響が大きく、さらに、(d) 態度と性別、感情と性別間に交互作用が確認される場合も存在する。よって、男女で分けて音声の表出を比較する必要がある。

そこで男女別に、どの共感状態にどのような差異があるかを明らかにするため、多重比較分析 (Games-Howell法⁵)を行った。表5に態度を固定した時の感情間の音声特徴量の有意差に関する分析結果を示す。表において、比較される感情は感情1、感情2で示されている。頻度がある一定数以下の組み合わせ、および、統計的有意差が確認できなかった組み合わせは表に含まれていない。カッコ外の数字は平均、カッコ内の数字は標準偏差である。

表5: 音声特徴量を感情間で多重比較した結果(態度別)

性別	態度	感情状態 ₁	感情状態 ₂	音声パラメータ	感情状態 ₁ Mean(±SD)	感情状態 ₂ Mean(±SD)	T値	P値
男	中立	平静	喜	Max F0	1.25(0.33)	1.44(0.37)	5.93	2E-06
				Min F0	0.81(0.10)	0.76(0.10)	5.84	3E-06
				Mean db	-0.03(4.21)	4.51(4.89)	11.10	1E-13
		驚	Max db	11.85(5.84)	20.77(7.61)	14.27	0.000	
			Mora/s	1.00(0.44)	1.21(0.40)	5.93	2E-06	
			Mean F0	0.99(0.13)	1.08(0.20)	4.55	0.003	
	ポジ	平静	喜	Min F0	0.81(0.10)	0.87(0.13)	3.89	0.028
				Mora/s	1.00(0.44)	0.80(0.50)	3.85	0.030
				Min F0	0.76(0.10)	0.87(0.13)	7.02	1E-08
		驚	Mean db	4.51(4.89)	0.92(5.37)	5.68	1E-05	
			Max db	20.77(7.61)	12.09(6.81)	9.98	3E-13	
			Mora/s	1.21(0.40)	0.80(0.50)	7.14	4E-09	
女	中立	平静	喜	Mean F0	0.94(0.11)	1.03(0.14)	5.00	0.001
				Max F0	1.17(0.30)	1.39(0.37)	4.60	0.002
				Mean db	0.56(4.48)	6.67(4.38)	9.98	2E-13
		驚	Max db	12.68(5.86)	20.29(6.81)	8.41	2E-11	
			Max F0	1.21(0.19)	1.31(0.16)	5.93	1E-06	
			Min F0	0.79(0.23)	0.62(0.20)	8.18	1E-12	
	ポジ	平静	喜	Mean db	0.15(4.90)	5.60(6.01)	9.97	2E-13
				Max db	11.29(6.09)	21.57(7.56)	15.00	0.000
				Mean db	0.15(4.90)	5.07(6.03)	8.80	2E-13
		驚	Max db	11.29(6.09)	15.67(6.09)	7.39	3E-10	
			Mora/s	1.01(0.51)	0.64(0.44)	8.36	1E-12	
			Min F0	0.62(0.20)	0.79(0.24)	6.68	2E-08	
ポジ	平静	喜	Max db	21.57(7.56)	15.67(6.09)	7.46	2E-10	
			Mora/s	1.03(0.42)	0.64(0.44)	7.95	8E-12	
			Mean db	1.74(5.36)	7.71(4.80)	7.25	8E-09	
	驚	Max db	12.71(5.81)	19.98(7.11)	6.58	5E-07		

⁵分析対象データに特定の分布や分散に関する制約を仮定できなくても適用できる。

表5からは以下のことがいえる。

- 態度=「中立」においては、感情間で音声特徴が有意に異なる。
- 態度=「ポジティブ」では、感情=「驚」の頻度が少ないので、意味ある比較ができなかった。
- 態度=「ネガティブ」の場合は、各感情とも頻度が少なく、有意差が確認できなかった。
- 女性のデータ数が少ないため、感情間の有意差が確認できたケースは少なく、さらに、音声特徴量で男女の比較はできなかった。
- 同じ態度と感情の組み合わせでも、男女で有意差のある音声パラメータが異なる場合がある。例えば、態度=「中立」のときに、感情=「驚」と感情=「平静」を比較すると、男性が基本周波数と発話速度で異なるのに対し、女性では音圧と発話速度で異なっている。すなわち、同じ態度でも感情の表出の仕方に男女差がある場合があることが分かった。

表6は感情を固定した時の態度間の音声特徴量の有意差に関する分析結果を示す。表の見方は表5と同様である。

表6: 音声特徴量を態度間で多重比較した結果(感情別)

性別	感情状態	態度 ₁	態度 ₂	音声パラメータ	態度 ₁ Mean(±SD)	態度 ₂ Mean(±SD)	T値	P値
男	平静	中立	ポジ	Mean F0	0.99(0.13)	0.94(0.11)	5.03	2E-04
			ネガ	Max db	11.85(5.84)	17.48(3.31)	6.35	0.001
		ポジ	ネガ	Mean F0	0.94(0.11)	1.07(0.11)	4.43	0.036
			中立	Max db	12.68(5.86)	17.48(3.31)	4.95	0.007
		中立	Φ	Min db	-16.28(3.94)	-18.37(3.18)	4.13	0.020
			Φ	Min db	-15.56(4.17)	-18.37(3.18)	4.84	0.001
	喜	中立	Φ	Mean db	2.75(2.49)	-1.56(3.20)	5.36	0.001
			Φ	Mean db	4.51(4.89)	-0.31(3.16)	4.72	0.031
		ポジ	Φ	Mean db	6.67(4.38)	-0.31(3.16)	6.49	9E-04
			Φ	Mean db	5.55(3.12)	-0.31(3.16)	4.76	0.013
		中立	ポジ	Mora/s	1.01(0.51)	0.85(0.39)	3.70	0.031
			ポジ	Mora/s	0.85(0.39)	1.13(0.28)	3.89	0.043
喜	中立	ポジ	Min F0	0.62(0.20)	0.75(0.22)	3.84	0.024	

表6からは以下のことがいえる。

- 感情=「平静」においては、態度の違いによって音声の特徴が有意に異なる場合が多いが、これ以外の感情では有意差が確認できるケースは少ない。
- 同じ感情と態度の組み合わせでも、男女で有意に異なる音声パラメータが異なる場合がある。例えば、感情=「平静」において、態度=「中立」と態度=「ポジティブ」を比較すると、男性は基本周波数に有意差があり、女性は発話速度で有意差が確認された。すなわち、同じ感情においても態度の表出の仕方が男女で異なる場合があることが分かった。

5.3 本節のまとめ

以上の結果、(1) 態度と感情の表出の異なりが音声特徴に現れること、(2) 態度よりも感情に音声的な特徴が現れやすいこと、(3) いくつかの態度と感情の表出においては男女間で相違がある可能性があること、(4) 感情は声の高さと大きさによって表出されているが、(5) 態度は主に声の大きさにより表出されている傾向があることが分かった。

6. 発話の反応対象を考慮に入れた分析

前節の分析では単一の発話区間を対象としたが、本研究における共感状態は、注目している発話の反応対象である先行発話を含む概念である。そこで、各共感状態の発話の音声特徴を明らかにするため、注目している発話の態度・感情をその話題機能と組み合わせて分析した。

以下では、まず新規話題提示発話(反応対象となる先行発話がないため共感状態は定義されない)における音声特徴を新規話題提示以外の発話との比較から分析し、次に共感とその他の各共感状態の発話の音声特徴を比較することにより、音声特徴をもとに共感状態を推定する可能性について議論する。

6.1 新規話題提示発話の音声特徴

今回の分析対象の2,104の発話区間のうち、話題機能=「新規話題提示」である発話区間は、86件(3.9%)であった。また、これらの発話における感情の分布は、平静:63, 喜:20, その他:3であった。

まず、話題機能(新規話題提示, 話題展開, 話題継続)の間で音声特徴が異なっているかを確認するために、2元配置分散分析を行った。分析においては、感情ごとに性別の影響を考えた。分析結果を表7に示す。

表7: 音声特徴量を話題機能と性別を因子とする2元配置分散分析した結果(感情別) (有意水準: *0.05, **0.01, ***0.001)

感情	従属変数	固定因子		交互作用
		話題機能 (P値)	性別 (P値)	話題機能*性別 (P値)
平静	Max hz	<2E-16***	0.053	1E-05***
	Mean hz	0.001***	3E-05***	2E-05***
	Min hz	1E-11***	0.212	2E-04***
	Max db	<2E-16***	0.113	0.008***
	Mean db	0.013*	0.148	0.024*
	Min db	1E-10***	0.870	0.735
	Mora/s	<2E-16***	0.034*	0.654
喜	Max hz	0.110	2E-06***	0.706
	Mean hz	0.039*	0.075	0.323
	Min hz	0.003**	2E-15***	0.337
	Max db	0.094	0.614	0.052
	Mean db	3E-09***	0.009**	0.378
	Min db	1E-06***	0.006**	0.374
	Mora/s	7E-04***	5E-07***	0.929
驚	Max hz	0.407	0.049	0.597
	Mean hz	0.957	0.011*	0.428
	Min hz	0.163	0.002**	0.427
	Max db	0.393	2E-05***	0.705
	Mean db	0.661	9E-09***	0.603
	Min db	0.059	3E-06***	0.240
	Mora/s	4E-04***	0.004**	0.803

表7からは、(a) 多くの音声特徴が話題機能間でも性別間(男女間)でも異なっていること、(b) 話題機能の因子と性別の因子の間に交互作用があることが分かる。以上から、話題機能=「新規話題提示」の発話の音声特徴とそれ以外の話題機能を持つ発話の音声特徴を比較するには、男女別に分析を行う必要があることが分かる。

表8に音声特徴量を新規話題提示と共感状態間で多重比較した結果を示す。この表には、比較対象の発話区間において感情が同一である場合に有意差のある音声パラメータのみが示されている。例えば、表8の一番上の行から分かるのは共感状態=「共感」の時、男性では感情=「平静」における音圧最小値の変動量(Min_db)が共感状態=「共感」の場合(-16.06)に比べて新規話題の場合(-18.37)

に有意に低い。すなわち、相対的に小さな声で新規の話題を提示する傾向がみられる。このことは、各共感状態間の音圧最小値の変動量を比較する図4の箱髷図によって、より明確に示される。

表8: 音声特徴量を新規話題提示と共感状態間で多重比較した結果(感情別)

性別	音声パラメータ	感情状態	共感状態の発話		新規話題の発話	T値	P値
			共感状態の種類	Mean(±SD)			
男	Min db	平静	共感	-16.06(3.53)	-18.37(3.18)	3.95	0.045
	Mean db	喜	共感	7.20(3.77)	-0.31(3.16)	6.87	0.001
	Mean db	平静	共感	2.00(4.45)	-1.56(3.20)	4.60	0.005
	Min db	喜	同感	-14.55(5.14)	-18.37(3.18)	4.50	0.007
	Mean db	喜	同感	5.70(5.25)	-0.31(3.16)	4.33	0.033
	Mean db	平静	同感	3.10(2.34)	-1.56(3.20)	5.33	0.007
	Mean db	喜	同感情	5.84(3.24)	-0.31(3.16)	4.60	0.029
	Mean db	喜	感情一致	5.31(4.73)	-0.31(3.16)	5.09	0.015
	Min db	平静	感情不一致	-15.86(4.07)	-18.37(3.18)	4.47	0.009
	女	Mora/s	平静	同感	0.71(0.33)	1.13(0.28)	5.18

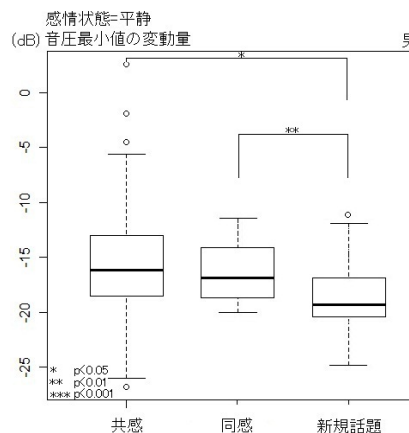


図4: 新規話題提示と共感状態の音声特徴の比較

以上の結果から、新規話題提示の発話区間は、声の大きさに着目することにより、話題展開または話題継続の話題機能を持つ発話区間から区別できる可能性があることが分かる。これは、新規話題提示の発話においては特定の発話に反応することがないため、基準値(態度=「中立」かつ感情=「平静」)に比べ声を特に大きくする必要がないことによると考えられる。

6.2 共感状態=「共感」の発話の音声特徴

すでに述べているように、発話区間における共感状態は、その発話が先行する発話のいずれかに対する反応である場合に定義される。そこでまず、各共感状態と話題機能(「話題展開」または「話題継続」)の依存関係をカイ2乗検定により調べた。その結果を表9に示す。表9におけるカッコ外の数字は実頻度、カッコ内の数字は頻度の期待値を示す。

表9からは以下のことがいえる。

- 共感状態=「共感」または「同意」(すなわち態度=「ポジティブ」)において話題継続が有意に多いのに対し、共感状態=「感情一致」「感情不一致」(すなわち態度=「中立」)では話題展開が有意に多い。すなわち、話題を維持する場合、先行発話と同じ意見・考えを持つことを表明することが多い、逆に話題を転換する場合、反応対象の発話に対する意見・考え方を示さない傾向がある。

- 共感状態=「感情不一致」における話題継続において、感情=「驚」が有意に多く出現している。これは、態度=「中立」かつ先行発話の感情が「驚」以外の場合であるので、態度=「中立」で話題展開までには至らない場合(話題継続)に「驚」の感情を表出しながら話題を維持している可能性が示唆される。
- 話題機能は態度により偏りが見られるが、感情では偏りが見られない(「驚」の場合を除く)。これから、話題機能は感情よりも態度と深い依存関係があると推測される。
- 共感状態=「同感情」、および「完全不一致」では、明らかな特徴は見られなかった。これは、そもそも頻度が低いためである。

表9: 共感状態と話題機能のカイ2乗検定の結果

(有意水準: *0.05, **0.01)

共感状態	感情	話題展開	話題継続
共感	平静	4(30.3)**	174(147.7)**
	喜	2(14.8)**	85(72.2)**
同感	平静	4(16.5)**	93(80.5)**
	喜	1(7.5)**	43(36.5)**
	驚	0(1.2)	7(5.8)
同感情	平静	1(3.7)	21(18.3)
	喜	2(2.7)	14(13.3)
	驚	0(0.3)	2(1.7)
感情一致	平静	157(107.3)**	473(522.7)**
	喜	36(26.2)*	118(127.8)*
	驚	0(0.3)	2(1.7)
感情不一致	平静	75(57.0)**	260(278.0)**
	喜	51(28.1)**	114(136.9)**
	驚	8(43.3)**	246(210.7)**
完全不一致	平静	2(1.2)	5(5.8)
	喜	0(2.2)	13(10.8)
	驚	1(1.2)	6(5.8)

次に、共感状態=「共感」の発話とそれ以外の共感状態の発話の間に見られる音声特徴の差異を多重比較により分析する。その結果を表10に示す。表8と同様に有意差が確認できた音声パラメータのみが示されており、比較対象の発話区間において有意差のある音声パラメータのみが示されている。

表10: 共感状態における音声特徴を多重比較した結果

性別	感情状態	音声パラメータ	その他の共感状態の発話		T値	P値	
			共感状態の種類	Mean(±SD)			Mean(±SD)
男	平静	Max db	同感情	17.55(3.38)	12.20(6.27)	4.79	0.023
		Mean F0	感情一致	1.00(0.13)	0.94(0.12)	4.19	0.015
	喜	感情不一致	3.95(4.94)	7.20(3.77)	4.48	0.006	
女	喜	Mean db	感情不一致	8.43(4.88)	4.65(4.25)	4.19	0.019
		Min db	感情不一致	-12.72(4.43)	-16.43(4.09)	4.38	0.011

まず、「共感」とそれ以外の共感状態を区別する音声特徴は少なく、特に「共感」と「同意」の間で有意差のある音声特徴は見出すことができなかった。共感と同意の差は、先行発話に対する感情の一致の有無であることから、個々の発話では感情の表出が見られるにせよ、先行発話に対して同じ感情を示すか・示さないかで顕著な音声特徴の差異はないことになる。このことは、図5に示す箱髷図によって、より明確に示される。

表10からは以下のことがいえる。これらの結果は、共感を同意以外の共感状態から区別するための手掛かりを与えるものである。

- 共感と同意情では男性の場合、感情=「平静」の時に最大音圧で有意差が見られる。共感状態=「同感情」とは態度=「ネガティブ」の場合には加ならず、この場合に声が相対的に大きくなっていることを示している。これより、平静な感情であっても、相手に意見・考えの違いを示すために声の大きさをコントロールしている可能性が示唆される。
- 共感と「その他(感情一致)」では男性の場合、感情=「平静」の時に基本周波数の変動量の平均値で有意差が見られる。後者では基準値から変化がない(比が1.00)であるのに対し、前者では相対的に声が低く(比が0.94)なっている。よって、感情=「平静」において相手へ共感を示す時には、声を低くする傾向が確認できる。
- 共感と「その他(感情不一致)」では感情=「喜」において男女差が見られた。すなわち、男性では「共感」において声が大きくなる傾向があるのに対し、女性では逆の傾向が観察された。

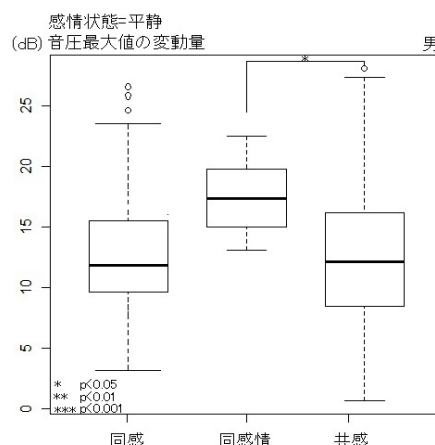


図5: 共感とその他の共感状態の音声特徴の比較

以上の分析結果からは、分析対象の発話 Y と反応対象である先行発話 X との感情の一致状況を、発話 Y の音声特徴から推定することは困難であるため、共感と同意を現状の分析の枠組みの範囲で弁別することは困難であることが分かった。しかし、単一の発話に対してはその感情を推定する手掛かりが得られているので、反応対象である先行発話 Y を特定することにより、共感状態を推定することが可能となると考えられる。反応対象を特定するためには、発話区間の時間間隔や韻律などの音声における同調傾向[大坊 1985]、発話内容(言語情報)、他のモード(例:視線、顔の向き)からの情報を統合する方法を検討する必要がある。

6.3 本節のまとめ

以上の結果、(1) 音声的特徴から共感を「同感情」、 「感情一致」、 「感情不一致」と区別できる見通しが得られたが、(2) 現在の分析の枠組みでは、共感と同意を有意に区別する特徴量は得られなかった。また、(3) 感情のカテゴリの一つである「驚」は、「喜」などの他の感情のカテゴリとは異質である可能性を示唆するという結果が得られた。

7. 議論

7.1 共感状態の手がかりとしての音声の有効性

今回の評定においては、音声のみを手がかりに共感状態の評定作業を行ってもらった。音声共感状態の規定や推定に有効であることを検証するには、さらなる検討が必要と考えられる。例えば書き起こされた言語情報や録画から得られる総合的な情報を手がかりとして評定した結果との整合性を調べるのが考えられる。あるいは、今回の結果に基づく音声情報からの共感状態の推定結果が、何らかの応用(例: 場の活性度の推定)において有用であることを示していく必要がある。

一方で、評定者は発話の言語的内容に影響されて評定作業を行っている可能性も否定することはできない。もしこの程度が大きければ、共感状態の推定メカニズムは発話の言語的内容をも考慮することが必要であることになる。ただし、現状の音声認識・話し言葉の処理技術が有効に適用できる範囲は限定的であると考えられる。

7.2 「驚」の感情カテゴリ

感情の表出は人、物、事、文化背景など様々影響を受け、表出も媒体が様々である。また、対象分野・視点の違いにより感情に関する認識は多様であり[Mehrabian 1972]、神経生理学的な研究も行われている[Ekman 1983]。

多くの感情の分類体系において、「驚」というカテゴリが設定されている。例えば、感情研究の端緒となった[Darwin 1872]においても、「心配」や「恐れ」といった情緒的なカテゴリと同列に「驚」(surprise)の感情が取り上げられている。しかし、「驚」というカテゴリは、少なくとも対話における発話を記述するという文脈においては、これらの情緒的なカテゴリにはない特徴を含んでいると考えられる。例えば、共感状態に基づく本研究の分析からも、態度を陽に表さない発話において「驚」が多く観察され(意外な情報を得た可能性を示唆)、話題の継続と深い関係を持つことが明らかとなるなど、「驚」の他のカテゴリには見られない特徴が浮き彫りとなった。すなわち、「驚」では少なくともそれが向けられる発話内容、あるいは、それを発話する発話者の社会的状況などに関する話者の知識状態や推論が深く関与していると考えられる。よって今後の検討の方向性の一つとして、「驚」の感情の本質や対話進行上の機能について、[藤本 2006]で議論されている「知識の有無と応答の予測性」という観点を参考にしつつ、発話内容も含めた分析を進めることが有用であると考えられる。

8. おわりに

発話音声における感情や態度に関する従来研究は、コントロールされた発話を分析対象としていたのに対し、本研究では自由対話を対象とし、発話に表出された態度と感情をもとに発話の共感状態という新しい概念を定義したうえで、その音声特徴について分析した。その結果、以下のことが明確となった。

1. 個々の発話に表出される態度と感情は独立ではないこと、
2. 態度よりも感情に音声的な特徴が現れやすいこと、
3. 発話の話題機能は態度と深く関係していること、

4. 共感状態を推定するためには、個々の発話の態度や感情の推定だけでなく、注目している発話がどの先行発話に反応したものかを明らかにする必要があること。

なお、上記(4)の課題は、発話の同調傾向、発話内容、視線・顔の向きといった多様な情報を統合する新たな研究課題につながるものである。また本研究からは、基本感情カテゴリにおける「驚」の異質性が浮き彫りとなった。例えば、態度=「中立」において話題継続を図る場合、「驚」が他の感情よりも有意に多く現れることなどが確認された。「驚」という感情の表出が持つ話題機能について、対話参加者の心的状態(知識・信念、推論)との関連を考慮しながら分析を進めていくことが望まれる。

謝辞

日頃ご議論いただく岸野文郎教授(関西学院大学)、北村喜文教授(東北大学)、大坊郁夫教授(大阪大学)、および、これらの研究室の方々に感謝する。また、本稿に関して有益なコメントを頂いた査読者の方に感謝する。

本研究の一部は、厚労省科研費(H19・がん臨床・一般009)、および、文科省グローバルCOEプログラム(アンビエント情報社会基盤創成拠点)の援助を受けた。

参考文献

- Banse, R., Scherer, K. R. Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology*, Vol.70, No.3, pp.614-636 (1996).
- Cornelius, R.R. *The Science of Emotion: Research and Tradition in the Psychology of Emotions*. Pearson Education (1996).
- Darwin, C. *On the Expression of the Emotions in Man and Animals*. John Murray (1872).
- Ekman, P., Levenson, R.W, and Friesen, W. V. Autonomic nervous system activity distinguishes among emotion. *Science*, 221, pp.1208-1210 (1983).
- Ekman, P. An argument for basic emotion. *Cognition and Emotion*, Vol. 6, No.3/4, pp.169-200, (1992).
- Mehrabian, A., and Epstein, N. A measure of emotional empathy. *Journal of Personality*, Vol.40, No.3, pp. 525-543 (1972).
- Murray, I.R. and Arnott, J.L. Toward the simulation of emotion in synthetic speech: A review of the literature on human vocal emotion. *Journal of the Acoustic Society of America*, Vol.93, no.2, pp.1097-1108 (1993).
- 大坊郁夫. “対人コミュニケーションにおける同調傾向—主に音声的行動について—”, 山形心理学レポート 4, pp1-15 (1985).
- 藤本英輝, 高梨克也, 竹内和広, 河野恭之, 木戸出正継. “自由発話データに含まれる態度表明発話の整理”, 電子情報通信学会技術研究報告, 巻 105, 542 号, pp. 27-32 (2006).
- 高亜罕, 大野ゆう子, 岡田志麻, “感情表出の特性に関連する音声関連パラメータの抽出—看護実務経験などの気配り度を考えて—”, ITヘルスケア, 巻 4, 1 号, pp. 46-49 (2009).
- 高亜罕, 大野ゆう子, 林 良彦, “対話における共感状態の分析”, 情報処理学会創立 50 周年記念全国大会, 2D-6, (2010a).
- 高亜罕, 大野ゆう子, 林 良彦, “対話における共感状態の音声情報からの推定”, 電子情報通信学会ヒューマンコミュニケーション基礎研究会, (to appear, 2010 年 8 月), (2010b).
- 前田貴司, 前嶋和毅, 梶村康祐, 山口徳郎, 北村喜文, 岸野文郎, 前田奈穂, 大坊郁夫, 林 良彦, “3 人会話における非言語情報と「場の活性度」に関する検討”, 電子情報通信学会ヒューマンコミュニケーション基礎研究会, Vol.2009, No.64-88 2010, pp.73-78 (2010).