

国立国語研究所学術情報リポジトリ

What Speech Conveys

メタデータ	言語: jpn 出版者: 公開日: 2015-10-30 キーワード (Ja): キーワード (En): 作成者: 森, 大毅, MORI, Hiroki メールアドレス: 所属:
URL	https://doi.org/10.15084/00000751

話し言葉が伝えるもの

What Speech Conveys

森 大毅 (MORI Hiroki)

1. はじめに

話し言葉の本質は、それによって伝達されるものが言語的情報だけではないところにある。逆の見方をすれば、話し言葉を文字言語に書き起こした時に失われる情報、それが話し言葉が持つ書き言葉にはない特質であるとも言える。本共同研究プロジェクトは、話し言葉が伝達する情報に (1) 何が含まれ、(2) それらの本質は何か、を整理し体系化することを目的としたものである。後述するように、話し言葉が伝達する情報は Fujisaki (1996) によって言語的情報・パラ言語的情報・非言語的情報の3つに分類された。我が国の音声研究ではこの分類が標準的である。本共同研究プロジェクトの対象は、話し言葉が伝達する言語的情報以外の情報であるが、プロジェクト名に「パラ言語情報および非言語情報」と入っているのはそういう意味である。

話し言葉は言語的情報以外に何を伝えているのか。この基本的な問いに対し、誰にも広く受け入れられる答は、実のところ存在していないと言ってよい。前川 (2002) は「パラ言語情報研究の課題」という講演の中で、この研究分野における「意味論」、すなわちパラ言語情報を体系づける理論の欠如を指摘した。この状況は現在においてもほとんど変わっていない。「話し言葉が伝えるもの」の基本的認識を確立しないままのパラ言語情報・非言語情報に関する研究は、音韻論の知識がないのに音声認識や音声合成をやるうとしているがごとき危うい状況であり、一刻も早い混乱の収束が必要である。

本稿は、本共同研究プロジェクトにおける議論を通してまとめ上げた知見のうち、話し言葉が伝えるものとは何かを理解する上で最も基本となる、話し手のメッセージと聞き手が得る情報に関係する部分を整理したものである。本稿はまた、話し手のメッセージおよび聞き手が得る情報に関連する概念の一部についても説明し、本共同研究プロジェクトの目的である基本概念の体系化にあたって何が問題となったのかを紹介する。

2. 概念の整理

2.1 言語的情報・パラ言語的情報・非言語的情報

Fujisaki (1996) は、音声に含まれる情報を言語的情報・パラ言語的情報・非言語的情報の3つに分類した。文字により転記できる単語や統語構造などを含む言語的情報と異なり、断

定／疑問などの意図による韻律の違い、あるいは丁寧／ぞんざいなどの態度による韻律の違いは文字により転記することができず、これらをパラ言語的情報と定義した。また、話者の個性や身体性、さらに気質や感情などの情報は談話全体を通して不変で話者が意識的に制御できない情報であり、これらを非言語的情報と定義した。

藤崎の定義では、転記可能性と話者の意識的な制御の有無が分類の要になっている点に注目すべきである。このため、話者の意識的な制御の有無が明確でない現象に関しては分類上の問題を生ずる可能性がある。しばしば問題となるのは感情の扱いである（2.2 参照）。

パラ言語的情報の研究では、言語的情報を統制した実験パラダイムが採用されることがある。これは一種の演技ということである。演技によるパラ言語的表現は話者が意識的に制御していることがある程度明白であるが、統制された実験室音声ではなく、社会的インタラクションにおける自然な対話音声を研究の対象とする場合には、意識的な制御によるものか否かを判別するのは一般に困難である。

2.2 感情

感情（あるいは情動, emotion, affect）という語を明確に定義することは困難である。心理学においても、感情を定義する上での基本的合意は長らく欠如している。このことは、“Almost everyone except the psychologist knows what an emotion is” (Young 1973) という言葉によく現れている。本共同研究プロジェクトでは、話し言葉が伝える話者の感情に関する概念の整理も重要な課題の1つと位置付けられているが、このような背景から、主として感情と話者の意識的制御との関係に研究の焦点をしばった。

藤崎の図式では、感情は非言語的情報に分類されている。この分類は、感情は意識的な制御によるものではないという考えに基づく。この考えは、強い感情体験に伴う感情表出に関しては一定の妥当性があるが、常に妥当な仮定とは言えない。感情表出には、自動的・不随意的なものだけでなく、話者の意志により制御されたものもある。他人が喜んでいるときに自分も喜んでみせることや、逆に個人的な不幸があっても人前では悲しんでいるようには見せないことはよくある。

Ekman などのネオダーウィニズムを信奉する心理学者は、基本感情説を唱えている。感情を進化の産物であるとするネオダーウィニズムでは、感情には先天的かつ普遍的なもの（＝基本感情）があると考えられる。Ekman などによって多く取り上げられた基本感情には、怒り・嫌悪・驚き・喜び・恐れ・悲しみのいわゆる6大感情が含まれ、各基本感情は固有の生理反応パターンを持つと考えられている。一方で、表情によって表出された感情を読み取る実験の結果から、感情表出には文化依存性があることが知られている。この矛盾を説明するために、Ekman and Friesen (1969) は表示規則という概念を導入している。表示規則は、感情表出を社会的約束事に従って消したり弱めたり強めたり置き換えたりする行動のことで、生得的なものではなく成長の過程で社会的に獲得されるものである。基本感情と表示規則の存在を認めると、感情は社会的文脈に従って多かれ少なかれ常に制御されているということになる。

欧米ではパラ言語には感情を伝える機能があるとされることが多い (Crystal 1975, Ladd 1996)。これは感情表出の社会的側面を考えれば妥当とも考えられるが、一方で、あまりの恐怖に声が震えるなど、感情表出には制御しがたい側面があることもまた事実である。音声伝える感情の概念を整理する上では、意識的制御の有無に関する両面性を考慮すべきである。

2.3 態度

態度という用語の使われ方は文脈により異なる。心理学では、人の社会的行動に影響を及ぼす、ある特定の対象に対する評価や好悪のことを態度という。「肯定的態度」「否定的態度」という表現が指しているのがこれである。一方言語学では、例えば「行く」という命題に対して、「行こう」という勧誘、「行くな」という禁止、「行きたい」という願望のように、話者の態度の違いが言語形式(法)として現れると説明される。しかしながら、後者は次節で述べる発話行為の違い、すなわち意図に分類した方が適切であろう。

話し言葉における態度の違いは、言語的情報と同時にパラ言語的情報としても伝達される。丁寧さ (politeness) と親しみ (friendliness) は、態度としてよく取り上げられる現象の例である。

態度と感情の間には密接な関係がある。感情との大きな違いは、態度には対象がある点である。因果関係としては、ある人物との関係の中で生じた感情がその人物に対する態度を作り出す、と考えるとわかりやすい。態度の表出は主として意識的な制御によるが、不満な態度が思わず顔に出るといった場面もよくあることを考えると、制御しがたい側面も存在すると言えるだろう。

2.4 意図, 発話行為

発話行為論 (speech act theory) では、音声を発することは、何らかの行為をしようとすることであると解釈される。意図は行為の目的であり、発話行為は何らかの意図を達成するための行動である。例えば、「行くな」という発話は、相手に行くという行動を取らせないことを目的としている。

発話の意図は主として言語によって伝達されるが、パラ言語的情報として伝達される場合もある。「行くの」という発話は、句末音調の違いにより、陳述／質問／強調のいずれの発話行為ともなり得る。

態度と意図との境界は曖昧である。強調の「行くの！」は、あまりにしつこく行くなと言われたため不愉快だという態度が反映されたものと解釈する方が自然な場面もあり得る。

態度と異なる点としては、意図の表出はいつも話者の意志によって選択されたものであるという点がある。上述した「行くの」の句末音調 (パラ言語) は、「行くの」という文 (言語形式) の選択と同時に、話者自身が選択したものと考えられる。

2.5 話者個人性

音声は、男女の別や年齢など話者の個人性に関する情報を伝達する。個人性情報のキャリアとなる声質には、生理学的・物理的な制約に由来するもの（＝個性）と、後天的に獲得した言語能力や言語運用に由来するものがある。男性に比べて女性の声が高く子供はさらに高いとか、声道が長い人のフォルマント周波数は低いとかいった傾向は、前者の生理学的・物理的な制約に由来する声質である。また、話し言葉に現れるジェンダーや、若者らしい話し方・子供らしい話し方などは後者に属する。このどちらも、一般には簡単に制御できないし、このような声質の特徴は談話全体を通して持続するので、個人性情報はパラ言語的情報とは別の種類の情報である、という考え方があり得る。

しかし、これらの声質は確かに個人にある程度固有のものだが、自らの意志で制御することが全くできないというわけでもない。ニューハーフと呼ばれる人々は、生物学的には男性であるが女性として認識されることを望み、自らの自然な音域よりも高い声で話そうとする。物真似は、話し手と聞き手が共有しているある特定の人物の個人性の演技である。物売り、幼稚園の保育士、駅員などに特有の話し方（職業口調）も、その仕事をしている間だけのものであるから個人性の演技の一種と言える。

このように、一口に個人性と言っても、その中には純粋な生理学的・物理的な制約に由来する個性だけでなく、話者の意志により制御される側面もある。

3. 話し言葉が伝えるものの分類

3.1 話し手のメッセージ

音声コミュニケーションには、話し手と聞き手が関わっている。話し手は、何らかの意図を完遂するための行動として話し言葉を生成する。話者が話し言葉によって聞き手に伝えることを意図した意味内容のことを、本稿ではメッセージと呼ぶ。

例えば、Aが暑いので窓際にいるBに窓を開けてほしい場面を考える。最終的な意図は「窓が開かれる」であるが、その達成のために、Aはまず「Bが窓を開けることを意図すること」を意図し、その達成のために、Bが「Aが『Bが窓を開けることを意図すること』を意図すること」を知ることが意図する（Grosz and Sidner 1986）。Aはその達成のために、「Aが『Bが窓を開けることを意図すること』を意図すること」をメッセージ化する。

話し手の意図を完遂するための主要な手段は言語的メッセージである。上の例では、例えば「窓開けてくれない」などのメッセージが作られ得る。

同時に、話し手の意図はノンバーバルな行動にも反映される。「窓開けてくれない」という発話は、異なる句末音調を伴うことができる。下降調は、何らかの命題（例えば、誰かが窓を開けてくれなかった、という報告）を陳述的に伝える場合に選ばれる可能性がある句末音調である。しかし、上の例の意図を達成するためには、上昇調が選ばれるであろう。このように、話し手の意図を完遂するために生成されるノンバーバルなメッセージを、**パラ言語的メッセージ**と呼ぶことにする。重要なのは、どちらのメッセージも話者の意志により選択されるということである。

言語的メッセージとパラ言語的メッセージのどちらも、発話行動の前に話し手の脳内で生成されるものである。したがって、これらは直接観察することができない。

3.2 話者個人性と感情

2.5 で述べた話者の個性性は、好むと好まざるとに関わらず音声に反映される。この意味で、個性性は話し手の意図とは無関係であり、メッセージ性がない。個性性は話者個性の全てではなく、個性性に話者の意志により制御される側面があるのは 2.5 で述べた通りであり、これは話し手が聞き手に伝える意図に基づくものであるからメッセージの一部であると考えられる。

2.2 で述べたように、感情の表出には意識的に制御されたものと意識的な制御によらないものの両方が存在する。前者は話し手が聞き手に伝えることを意図したメッセージであり、言語的メッセージとパラ言語的メッセージの両方が同時に使用される。一方、後者は話し手の意図とは無関係な心理状態であり、メッセージではない。このように、本研究における整理では、メッセージ性をもって生成された感情表出と不随意的に生成された感情表出とは理論上区別される。藤崎の三分法との最も顕著な違いはこの点にある。この区別は音声研究の方法上でも非常に重要である。なぜなら、感情音声の研究のために用いられている音声資料の多くは、発話者が研究者の指示に従って自らの意志により生成した音声だからである。話し手により意識的に制御された感情表出は、藤崎の分類では適切に位置付けることが困難であった。本研究における整理は、話し言葉の中でそのような種類の感情が占める位置を認め、研究の価値を積極的に評価できる点で優れていると言える。

3.3 聞き手が得る情報

3.1 で、話者が聞き手に伝えることを意図したものをメッセージと定義した。これに対し、発話から聞き手が得たもののことを本稿では**情報**と呼ぶことにする。

聞き手は話し手の発話を知覚して言語的メッセージおよびパラ言語的メッセージを復号化する。もちろん、復号化されたメッセージは話し手のメッセージといつも一致するわけではない。聞き手は、復号化されたメッセージから異なる種類の情報を得る。これら 2 種類のメッセージを復号化した結果得られる情報を、それぞれ**言語的情報**および**パラ言語的情報**と呼ぶことにする。3.1 に挙げた例では、言語的情報は「窓開けてくれない」というテキストが表す命題的情報であり、パラ言語的情報は上昇調音調に関連した話し手の意図や、それ以外の韻律的情報に関連した話し手の態度などを含む情報である。

さらに、聞き手の注意が向けられるのは、話し手のメッセージだけではない。言語的情報およびパラ言語的情報以外に聞き手が得る情報として、話者の意識的制御によらない感情表出から知覚される感情状態が挙げられる。ここでは話者の意識的制御によらない態度の情報まで含めて広く考え、**心理状態情報**と呼ぶことにする。

これらの情報以外に聞き手が得る別の種類の情報には、話者の個性性がある。これは藤崎の分類では非言語的情報として分類されていた種類の情報の一部である。

3.4 話し手と聞き手

3.1に挙げた例のように、話し手は意図を達成するために言語的メッセージとパラ言語的メッセージの両方を手段とする。一方、言語的情報とかパラ言語的情報という分類は、聞き手が話し手の意図を解釈するための手がかりを、復号化されたメッセージの種類によって区別したものである。重要なのは、ここで定義された言語的情報およびパラ言語的情報は、聞き手が得る情報であり、話し手がメッセージに託した意味内容とは区別されることである。これらが同じでないことは以下の点から説明される。

第一に、聞き手は話し手のメッセージの全てに注意を向けるのではない。暑いのを我慢していたのに気づいてくれなかったという恨めしい態度をパラ言語的メッセージとして伝えようとしても、聞き手には伝わらず、「窓開けてくれない」を単なる依頼と解釈するかもしれない。

第二に、復号化されたメッセージの聞き手による解釈は、メッセージの背景となる話し手の意図と必ずしも一致しない。話し手にはそのつもりがなくても、聞き手は「窓開けてくれない」という発話から話し手のいらついた態度を勝手に感じ取るかもしれない。

第三に、3.3で述べたように、聞き手の注意が向けられるのは、話し手のメッセージに限定されていない。誰が話しているか、どのような心理状態で話しているかなどの情報は、話し手が伝達を意図したものではないが、それらもまた聞き手が得ることができる情報である。

3.5 音声コミュニケーションの構図

以上で述べた話し手のメッセージ、話者の個性性と感情、聞き手が得る情報、話し手と聞き手との関係を整理して図1に示す。

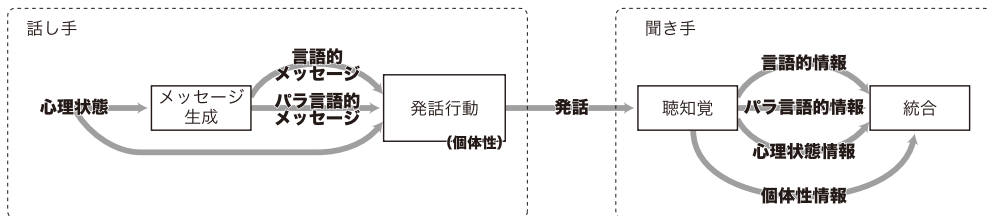


図1 メッセージと情報

話し手における発話生成の出発点には心理状態がある。心理状態は、発話行動の背景となる話者の認知的状態、および非認知的な感情状態を含む。心理状態は、メッセージ生成に影響を及ぼす。2.4で示したような、不愉快な態度の発話において「行くの!」という単語列(言語的メッセージ)を選んだり、上昇調音調(パラ言語的メッセージ)を選んだりする例は、不快な心理状態がメッセージ生成に影響を及ぼす例である。心理状態はまた、メッセージ生成を経由せず、発話行動に直接影響を及ぼす。感情が声質への影響という形で不随意的に表出する現象はこのパスで表現されている。

メッセージ生成の箱によって表されている行為には、言語的メッセージの生成とパラ言語

的メッセージの生成とが含まれる。言語的メッセージの生成とは、語彙と統語構造の決定である。一方、パラ言語的メッセージは意識的に制御可能な韻律パターンの違いや声質の違いとして、言語的メッセージとは独立に生成される。

続く発話行動の段階では、言語的メッセージとパラ言語的メッセージが統合され、さらに不随意的な感情と個性の要素も統合されて、実際に観察可能な音声生成される。

聞き手は、発話から知覚された言語的情報、パラ言語的情報に加え、話者の意志とは無関係に伝達される心理状態情報および個性情報を統合し、話し手の発話行為を総合的に解釈する。

4. おわりに

本稿では、話し言葉が伝えるものに関連した、話し手のメッセージと聞き手が得る情報に関して本共同研究プロジェクトでまとめ上げた内容を整理して述べた。

話し手のメッセージおよび聞き手が得る情報を構成する具体的な要素については、その大部分が未だ不明のままである。本共同研究プロジェクトでは、定義が曖昧または矛盾した概念についての共通認識を得る方法として、個別研究で使用されている概念に属性を付与する試みを実施した。付与した属性には、使用されているカテゴリ、データの性質（自発／朗読）、記述対象（形式／機能、原因／効果）、時間スケール（発話未満／発話／談話／永続…）などがある。多数の関連研究に対してこのような属性を付与することで、類似した概念の抽出や矛盾した用語の使用例の発見などが可能になると期待される。

● 参考文献 ●

- Crystal, David (1975) Paralinguistics. In: Jonathan Benthall and Ted Polhemus (eds.) *The body as a medium of expression*, 162-174, London: Institute of Contemporary Arts.
- Ekman, Paul and Wallace V. Friesen (1969) The repertoire of nonverbal behavior: Categories, origins, usage, and coding. *Semiotica* 1: 49-98.
- Fujisaki, Hiroya (1996) Prosody, models, and spontaneous speech. In: Yoshinori Sagisaka, Nick Campbell and Norio Higuchi (eds.) *Computing prosody*, 27-42, New York: Springer-Verlag.
- Grosz, Barbara J. and Candance L. Sidner (1986) Attentions, intentions, and the structure of discourse. *Computational Linguistics* 12: 175-204.
- Ladd, D. Robert (1996) *Intonational phonology*. Cambridge: Cambridge University Press.
- 前川喜久雄 (2002) 「パラ言語情報研究の課題」『日本音響学会 2002 年秋季研究発表会講演論文集』 247-250.
- Young, Paul T. (1973) Feeling and emotion. In: Benjamin B. Wolman (ed.) *Handbook of general psychology*, 749-771, New Jersey: Prentice-Hall.

《要旨》 Fujisaki (1996) は、音声に含まれる情報を言語的情報・パラ言語的情報・非言語的情報の3つに分類した。藤崎の定義では、転記可能性と話者の意識的な制御の有無が分

類の要になっている。このため、話者の意識的な制御の有無が明確でない現象に関しては分類上の問題を生ずる可能性がある。特に、感情の扱いはしばしば問題となっていた。本研究では音声によるコミュニケーションの図式を整理し、話し手により意識的に制御された感情表出を適切に位置付けるために、メッセージ性をもって生成された感情表出と不随意的に生成された感情表出とを区別した。また、話者の言語的メッセージおよびパラ言語的メッセージと、聞き手が得る言語的情報およびパラ言語的情報とを区別し、それらの違いを明確に述べた。

Abstract: Fujisaki (1996) defined three classes of information delivered by speech: linguistic information, paralinguistic information, and non-linguistic information. The key points in this classification are whether transcription is possible and whether the speaker's intentional control is involved. Classification problems can arise when it is unclear whether or not the speaker's intentional control is involved in some phenomenon. In particular, the treatment of emotion has been problematic. This review aims to reorganize the schematic view of speech communication and position emotional expression appropriately by drawing a distinction between emotion as message and involuntarily expressed emotion. It also distinguishes the linguistic and paralinguistic messages of the speaker from the linguistic and paralinguistic information received by the listener.

森 大毅 (もり・ひろき)

宇都宮大学大学院工学研究科准教授。博士（工学）（東北大学）。東北大学助手、宇都宮大学助手を経て、2006年2月より現職。

主な著書・論文：Some considerations for designing spoken dialogue database from the viewpoint of paralinguistic information (with Hideki Kasuya, Makoto Nakamura, and Minoru Amanuma, *Acoustical Science and Technology* 24(6), 2003), 「対話音声のパラ言語情報ラベリングの安定性」(共著, 『日本音響学会誌』61(12), 2005), Facial expression generation from speaker's emotional states in daily conversation (with Koh Ohshima, *IEICE Transactions on Information and Systems* E91-D(6), 2008), Constructing a spoken dialogue corpus for studying paralinguistic information in expressive conversation and analyzing its statistical/acoustic characteristics (with Tomoyuki Satake, Makoto Nakamura, and Hideki Kasuya, *Speech Communication* 53(1), 2011).

領域指定型共同研究プロジェクト

「パラ言語情報および非言語情報の研究における基本概念の体系化」
プロジェクトリーダー 森 大毅 (宇都宮大学 大学院工学研究科 准教授)

プロジェクトの概要

発話の意図・態度、話者の感情状態、話者の個性など、パラ言語情報・非言語情報のアノテーションに関連した研究は国内外で盛んに行われるようになったが、これまでは研究者間で共通の基本概念が確立されていなかったため、用語も問題設定も独自のものとなる傾向があった。本研究は、パラ言語情報および非言語情報に (1) 何が含まれ、(2) それらの本質は何か、を整理し体系化することを目的とする。これにより、音声コーパスの設計者・作成者・利用者間で、アノテーションに対する認識を誤解なく共有できるようになる。