

音声伝えるものとは？*

○森 大毅 (宇都宮大)

1 はじめに

音声は言語のほかに何を伝えているのか。この一見あまりにも基本的な問いに対し、誰にも広く受け入れられる答は、実は存在していない。この話題を最も専門的に扱っているはずの本学会においても、音声伝えるものに対する基本的な概念や用語法についての統一がないまま、研究グループごとに独自の問題設定がなされて来た。

ちょうど10年前、前川 [1] は本学会研究発表会において「パラ言語情報研究の課題」という講演を行った。前川はその中で、この研究分野における「意味論」、すなわち音声伝えるものを体系づける理論の欠如を指摘している。その後もパラ言語情報／非言語情報に関連する研究が数多く発表されてきたが、「音声伝えるものは何か」はずっと明らかでないままにされて来た。これは、さしずめ音韻論の知識がないのに音声認識や音声合成をやろうとしているのにも似て、議論の一般性の点で甚だ心もとない。

本稿では、これまでパラ言語情報／非言語情報と呼ばれてきた「音声伝えるもの」に関する研究を振り返り、研究の方向性に関わる様々な視点を提供するとともに、改めてこの問題に対する考え方を整理したい。これは、パラ言語情報／非言語情報という分類を含むこれまでの整理の再検討をも含む。さらに、現時点での未解決問題と今後研究すべき課題を明らかにすることを目指す。

2 音声伝えるもの：研究の歴史

2.1 藤崎の三分法

これまでの多くの研究は藤崎 [2] による分類を参照している。藤崎は、音声に含まれる情報を言語的 (linguistic) 情報・パラ言語的 (paralinguistic) 情報・非言語的 (nonlinguistic) 情報の3つに大別した (表1)。

注目すべきは以下のポイントである。

- (1) パラ言語的情報は、言語的情報と同様、話者の意識的な制御の下に選択される。非言語的情報は話者が意識的に制御していない。

- (2) 感情は非言語的情報に分類される。

我が国ではこの分類が広く受け入れられているが、しばしば問題となるのは感情¹の扱いである。

後述するように、音声学／音声言語情報処理の分野では、paralinguistic という用語は話者の感情を含んだ文脈で用いられるのが近年では大勢になっている。ここに用語の不統一の問題がある。

また、感情は意識的な制御によるものではないから非言語的情報に分類されているのだが、ここにも問題がある。感情表出は制御されている (表示規則 [4]) からである。我々はいつも感情をむき出しにすることはないし、逆に「怒ってみせる」「驚いてみせる」という具合に、そう見えるようにふるまうことも多い [5]。対人コミュニケーションにおける感情の模倣は、舞台やドラマで役者が演技する感情と違って、決して特殊なものではなく、むしろコミュニケーションにおいて本質的な役割を果たす要素だとみなす方が自然であろう。

2.2 ノンバーバル

1950年代、Birdwhistell や Hall などによって始められたノンバーバル・コミュニケーション (nonverbal communication) 論は、「音声伝えるもの」と密接な関係がある。

ノンバーバル・コミュニケーションにおける構成要素には動作・目・対人的空間などがあり [6]、音声もちろんその1つである。コミュニケーションに関係する言語以外のあらゆる要素はノンバーバルであるから、表1の図式における言

Table 1 音声伝える情報 — 藤崎による分類

	文字転記	意識的 制御	範疇的	程度の差
言語的情報 単語 統語構造 など	○	○	○	×
パラ言語的情報 意図 (断定/疑問/勧誘/反論…) 態度 (丁寧/ぞんざい…) など	×	○	○	○
非言語的情報 話者 (性別, 年齢, 個人…) 身体的状態 心理的状态 (感情…) など	×	×	○/×	○/×

*What does speech convey?
by MORI, Hiroki (Utsunomiya University)

¹「感情とは何か」の議論については [3] を参照のこと。

語的情報以外の情報はノンバーバル情報の一種だと言える。

ところで、ノンバーバル情報を日本語にすると「非言語情報」になってしまうため、藤崎の三分法における非言語(的)情報と用語の上で区別しがたい。前者はパラ言語情報の上位概念であり、後者はパラ言語情報とは区別されるものである。本稿ではカタカナのままとする。

2.3 パラ言語

ギリシア語由来の para- という接頭辞は「側にいる (beside)」というほどの意である。(cf. paralegal, paramedic) つまり、語源から言えば、paralanguage は言語そのものではないが言語に従属するもの、ということになる。この意味では、しばしば使われる「周辺言語」[6] という訳語はあまり適当とは言えず、どちらかと言えば「言語周辺」の方が近い。

元々 Trager [7] が使い始めた paralanguage という用語は、かつては音声のノンバーバルな側面が言語学の視点からは周縁的と見なされていたことを反映している [8]。今日では、ノンバーバルな側面は決して周縁的でなく、それこそがインタラクション場面の多くで決定的な要素であることが広く認識されるようになったが、paralanguage という用語はそのまま残った。代表的なノンバーバル・コミュニケーション論の1つである Birdwhistell の動作学 (kinesics) とのアナロジーから、声以外の、表情やジェスチャーといったノンバーバル・コミュニケーションの側面まで含めて paralanguage と呼ぶことはかなり以前からあったようであるが、ノンバーバル・コミュニケーションの構成要素を論じた Vargas [6] は、その構成要素の1つである paralanguage を、人間の声からもたらされる刺激に限定している。Vargas や藤崎、Crystal などに習い、少なくとも日本音響学会では、パラ言語という用語は口から出る音に関連したものに限って使用するのが適当だと考える。

表1では感情は非言語的情報に分類されているが、一般的に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 [9-11]。パラ言語が非言語的情報である個人性を伝えるとされることは稀だが、Interspeech 2010 Paralinguistic Challenge における同定対象は話者の年齢・性別・感情であった。

3 概念の整理

「音声伝えるもの」として、表1では意図、態度、感情、話者の情報などが挙げられている。このほかの重要な概念としては、発話様式、口調、フォーカス、実時間対話支援 [12] などが挙げられる。

これらの概念の中には、感情に代表されるように定義が難しいものもある。また、実際に個別の研究によって、同じ用語でもそれが指す概念がかなり異なる場合がある。ここでは、いくつかの切り口からそれらの概念を整理することで、混乱を解きほぐすためのヒントを提供する。

3.1 メッセージ&心理状態 vs 情報

音声コミュニケーションには、話者と聴き手が関わっている。本稿では、話者が聴き手に伝えようとするものをメッセージ、発話を通して聴き手や観察者に伝わるものを情報と呼ぶことにする。

発話行為論 [14] では、何かを話すことは何らかの行為と見なされる。このコミュニケーション観に従えば、話者のメッセージは何らかの意図²を完遂するための手段である。多くの場合、言語的メッセージ(命題、法、モダリティなど)がその意図を実現する主要な手段であるが、同時に話者の意図はノンバーバルな行動にも反映される。例えば疑問上昇などがそれであり、これら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。

発話行為の背景にある意図とは独立した要因として、感情や態度に関連した話者の心理状態もまた言語的/パラ言語的メッセージの選択に関与する。「うるさい! 静かにしろ!」という発話からは、話者の怒りが推測できる。怒っていないなら、別の語を選んだだろうからである。同時に、話者の心理状態は、自律神経系の作用により音声器官に影響する。例えば怒りのように心理的に覚醒した状態は、喉頭筋などの機能亢進といった生理的反応を導き、声の特徴の変化として表出する。このような感情表出は自動的プロセスであり、メッセージとしての性質を持たない。

これと反対に、話者の心理状態とは独立に、話者の意図を実現するために、言語的あるいはパラ言語的メッセージによって感情や態度を伝えようとする場合もある。これが2.1で述べた模倣で

²表1における“意図”と混同しないこと。

ある。

パラ言語的メッセージと、意図とか態度とか感情そのものを混同すべきではない。それらは(言語的メッセージによっても)パラ言語的メッセージによっても表現されるが、パラ言語的メッセージは意図とか態度とか感情そのものではなく、それらの反映に過ぎない。

音声信号は、言語的メッセージ、パラ言語的メッセージ、話者の心理状態の全てが関与して生成される。ここで話者の身体的特性も影響することは言うまでもない。一方、聴き手は発話から言語的情報、パラ言語的/心理状態情報、話者情報を同時に知覚する。パラ言語的情報と心理状態情報を区別することは、話者がそれを意志によって選択したものか否かを判断することであり、一般に困難である。

以上の関係を図1にまとめておく。藤崎の図式と異なる主な点は、話者側と聴き手側を区別し、メッセージと情報が必ずしも一致しないことを明示したこと、パラ言語的メッセージには話者の意志による感情や態度の表出も含まれること、心理状態が言語的/パラ言語的メッセージの選択に影響を与えることを明示したことである。

3.2 形式 vs 機能

Tragerがparalanguageとして論じたのは主として音声言語を構成するノンバーバルな手がかり(例えば話者性、声質、発声の種類)であり、それが伝えるものについては論じられていなかった。すなわち、Birdwhistellの動作学と同様、それは意味論を扱ったものではなかった。これに対して、藤崎の言うパラ言語的情報は、それが伝えるものを論じている。前者のような形式(form)と、後者のような機能(function)とを区別することは有益であると考えられる。

例えばToBI [13]ではパラ言語の構成要素であるイントネーションの形式を記述するが、パラ言語の機能であるパラ言語的情報の伝達に関しては記述しない。筆者は、“paralanguage”(パラ言語)を専ら形式の意味で使い、パラ言語に関わる機能に言及する際には“paralinguistic”(パラ言語的)を使うことを提案する。

3.3 「伝えようとしたもの」vs「伝わったもの」

藤崎の分類におけるパラ言語的情報は、話者が意識的に選択したものであるという点で言語情報と共通している。「あった?」「あった。」だ

けでもコミュニケーションは成立しているが、最初の話者は、その発話意図が「あったかどうかを尋ねる」であることを示すために、句末を上昇調で発話する。このパラ言語的メッセージは、話者が発話意図の違いに応じて選択したもののよう

に思える。しかし、実際の対人コミュニケーションでは、発話意図を客観的に同定することは困難である。音声インタラクションへの談話行為タグの付与作業は発話意図の同定を含むが、実際の会話では、会話構造は聞き手の解釈によって事後的に形成されることも多い[15]。すなわち、インタラクションは話し手だけが作るものではなく、聞き手が欠かせないということである。このことは、発話の言語的・パラ言語的特徴の違いを発話の原因に求めることを無意味にする。そこで、発話の原因よりも、その効果を記述することが妥当性を帯びる。「伝えようとしたもの(メッセージ)」ではなく「伝わったもの(情報)」を記述しようとする態度である。

3.4 演技 vs 自発

「音声が伝えるもの」に対する言語学的アプローチでは、しばしば実験条件の統制のため、指定された文を読み上げた発話が研究の対象となる。読み上げ音声と自発音声の性質が異なることは改めて書くまでもないが、感情などに関連した研究を統制下で実施する場合は、それらの表出は演技とならざるを得ない。同じ「感情」を研究対象にしていると言っても、読み上げ音声と自発音声では、扱っている要因が全く異なる。

統制された感情音声の研究の多くは、表情研究と類似したパラダイムに則っている。すなわち、話者が「怒り」「喜び」「悲しみ」などの基本感情の下で(しばしば同一の)文を朗読する。適切な感情表出のために、声優などのプロフェッショナルに演技させる、または基本感情の喚起手続き(例えばVelten法)を使用するなどの方法が取られる。図1で言えば、前者はパラ言語的メッセージだけが、後者は主に心理状態が音声に反映される。いずれにせよ、実験の再現性が重要である。

自発音声を対象にする場合は、パラ言語情報のアノテーションが問題となる。顧客の怒り[16]など、興味の対象が限定されている場合には比較的簡単だが、基本6感情の比較と同じパラダイムを持ちこむのは困難である。次元による感情記述は今や標準的な方法の1つとなっているが、

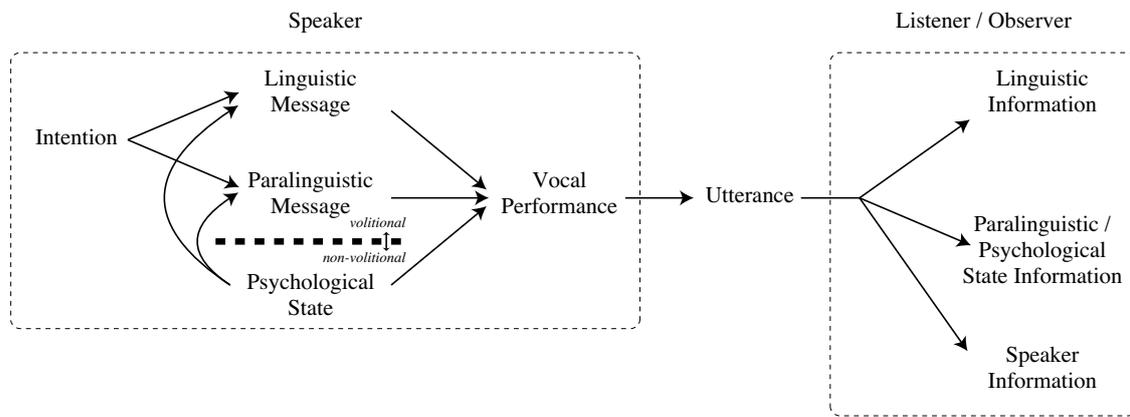


Fig. 1 メッセージ、心理状態、情報

我が国では今だにマイナーである。発話意図のアノテーションはさらに困難である [15]。

4 おわりに

本稿では、「音声伝えるもの」に対する考え方を整理し、今後音声研究が取り組むべき課題を浮き彫りにすることを目指した。用語やそれが指す概念の不統一の問題を指摘し、話者と聴き手の不一致・非対称性を、「メッセージ&心理状態 vs 情報」の図式で説明した。

以下に、今後取り組むべき未解決課題を順不同で列挙する。

- (1) パラ言語のコード化。ToBI にはインタラクションの視点が欠けており、時間構造の記述が不十分。
- (2) パラ言語的メッセージの形態論。(そもそも範疇的であるという仮定が誤っているかもしれない)
- (3) パラ言語的メッセージの生成メカニズム。その普遍性と文化依存性。
- (4) パラ言語的/心理状態情報の記述。(永遠の課題)
- (5) 笑い、咳払い、ため息、舌打ち、相槌、フィラーなどのノンバーバル行動の取り扱い。
- (6) 発話行為の背景にある意図の定式化。特に、タスクに依存しない(自由会話を含む)記述の枠組。ここができないと、パラ言語との対応を研究することが不可能。
- (7) 感情や態度の表出における演技について。特に、日常のコミュニケーションにおける演技をも包含した理論的枠組。
- (8) 自発音声を対象にしたパラ言語情報研究の一般化。タスク/記述法が異なると、個別研究の中で閉じた知見になってしまうのが現状。アノテーション共通化のための理論的検討。

本稿が、「音声伝えるもの」の研究の機運が再び高まる契機となることを期待してやまない。

謝辞

本発表の内容は、国立国語研究所共同研究プロジェクト「パラ言語・非言語情報の研究における基本概念の体系化」(2010-2013年度)における討論の成果である。いつも熱心にご討論いただく共同研究プロジェクトメンバーの諸氏に感謝申し上げます。

参考文献

- [1] 前川, 音講論 (秋), 247-250, 2002.
- [2] H. Fujisaki, in *Computing Prosody*, 27-42, 1996.
- [3] 中村, 音講論 (秋), 1-2-11, 2012.
- [4] Ekman et al., *Semiotica* **1**, 49-98, 1969.
- [5] Banse et al., *J. Pers. Soc. Psychol* **70**, 614-636, 1996.
- [6] M. F. Vargas, 石丸訳, 非言語コミュニケーション, 1987.
- [7] G. L. Trager, *Studies in Linguistics* **13**, 1-12, 1958.
- [8] D. Crystal, *Proc. Univ. Newcastle-upon Tyne Philosophical Soc.* **1**, 93-108, 1966.
- [9] D. Crystal, in *The Body as a Medium of Expression*, 162-174, 1975.
- [10] D. R. Ladd, *Intonational Phonology*, 1996.
- [11] <http://en.wikipedia.org/wiki/Paralanguage>
- [12] Oohashi et al., *Proc. Speech Prosody 2010*, 2010.
- [13] Silverman et al., *Proc. ICSLP '92*, 867-870, 1992.
- [14] J. L. Austin, *How to Do Things with Words*, 1962.
- [15] 高梨, 音講論 (秋), 1-2-12, 2012.
- [16] 野本他, 音講論 (春), 237-238, 2011.