

音声生成と音声知覚を中心に人間による 音声コミュニケーションに関する研究を考える*

○荒井隆行（上智大・理工）

1 はじめに

「ことばの鎖 (Speech Chain)」[1] に代表されるように、人間による音声を介したコミュニケーションでは音声生成と音声知覚の両者が密接に関係している。この「ことばの鎖」では、まず話者の頭の中で思い描くメッセージなどがあり、語の並びとして文やフレーズが形成される。その際、どのような語が選択されるか、あるいは語の系列がどのようなルールのもとに並んだり変形したりするかなどは、その言語毎に制約が存在する。さらに語の並びは音素の並びとして表現され、その際、音素目録もまた言語に依存すると同時に、音素系列にも言語毎に制約がある(言語学的段階)。次に、その音素列に従い、様々な神経を介して発話に関わる筋肉が動き、発声や調音等が行われる(生理学的段階)。そして、その際に適宜、韻律情報が付加される。そこで作られた音は話者の口から聴取者の耳まで空気中を伝搬する(音響学的段階)。耳に届いた音は外耳・中耳・内耳を経て、聴神経を介して脳へ伝えられる(生理学的段階)。そして脳内において再び言語メッセージとして聴取者に理解される(言語学的段階)。同時に、話者は自分が発した音声をモニタリングし(聴覚フィードバック)、必要に応じて次の発話を調節する。話者と聴取者の関係は随時、入れ替わりながらコミュニケーションは進むことになる。

このように人間による音声コミュニケーションにおいては、言語学、音韻論、音声学、音声科学、聴覚科学、認知心理学、脳神経科学など、様々な分野が関わっていることは言うまでもない。また、人間が生まれてから成長し、老いていく過程においても、言語発達、言語獲得・音声獲得などといった発達の分野や高齢者に関わる加齢の問題なども関係する。他の動物による音声コミュニケーションとの類似点、相違点などを知ることにより、人間による音声コミュニケーションに対する理解が深められることも多い。そして我々が外国語を学ぶとき、第2言語習得の分野にも範囲が及ぶ。いかに外国語を学ぶかは、いかに外国語を教育するかという外国語教育や教授

法などにも通じる。「ことばの鎖」のどこかに何らかの障害が生じると、それはコミュニケーション障害となり、場合によっては言語治療の対象にもなり得る。そして音声言語医学や言語障害学の領域にも関わることとなる。さらには障害者支援や社会福祉も関係し、バリアフリーやユニバーサルデザインにもつながる。音声に関わる犯罪捜査など、法科学分野も重要である。

以上は人間による音声コミュニケーションに関わる代表的な分野を概観しただけだが、そこには学際的領域が広がり、その領域では現在も様々な研究が展開され、その裾野が広がりつつある。その展開が今後も継続的に発展していくためには、音声生成と音声知覚の分野が決して切り離されてはならず、両分野が共に1つの土俵の上で共存し議論を展開させていく必要がある。それがあって初めて議論できる重要なテーマとして、例えば音声知覚における運動理論 (motor theory) があげられる[2,3]。ミラーニューロンの発見後、新しい脳研究が聴覚系と生成系に存在する関係をさらに解き明かすことによって、運動理論が見直される機運が生まれているという点[3]は、とても興味深いことである。

さらに、技術的な発展により、従来できなかった新しい研究が可能となり、音声生成・音声知覚分野においてもそれぞれの機構の解明に加え、その相互作用や両者を総合的に1つのシステムとして理解するなど、これからの研究がより一層期待されている。そして、真理を追究する科学的な側面のみならず、例えば昨今の機械学習による音声機能のモデル化への応用や、言語教育、補聴、医学、テントターテイメントなどより広い分野からの要請があるのも事実である。

このように、人間による音声コミュニケーションにおける音声生成と音声知覚は切り離せない関係にある。次節では、そのことを実例をあげながら考えてみる。

* What to think when researching human speech communication: Speech production and perception, by ARAI, Takayuki (Sophia Univ.).

2 実例

2.1 Lombard 効果

雑音環境下における音声生成・音声知覚の問題を取り扱った例として、Lombard 効果[4-6] があげられる。雑音が多く存在する中で音声を生成すると、静かなときよりも大きな声になる。そればかりか、基本周波数が高くなり、口を大きく開くことで第1フォルマント周波数も上昇するなどの現象が観測される。さらに興味深いことに、Lombard 音声のほうが雑音環境下において明瞭性が高くなることが知られている。そして、残響環境下においても、同様の結果が観測されることも報告されている[7]。いずれにしても、雑音（や残響）が聴覚系から入ることによって生成系に悪影響を与える結果、発話の明瞭度が下がることもあれば、発話の明瞭度を上げるために生成系を制御する場合もある。これら両方向の議論が必要な例であろう。

2.2 日本語における母音の無声化

日本語において、無声子音に挟まれた狭母音が無声化することはよく知られているが、自然発話の場合はその頻度が増す[8,9]。その1つの説明として、話速が上昇すると発話の効率の観点から無声化が促進すると考えられる。一方、摩擦音と閉鎖音に挟まれた母音よりも、摩擦音と摩擦音に挟まれた母音のほうが無声化率が下がるという報告がある[10]。これは、明瞭性確保が優先された結果であると考えられている。発話の効率と明瞭性の両方向の議論が必要な例と言えよう。

2.3 手による音声生成の模倣と発音訓練

Arai (2015) [11] では、手と声道模型を用いて音声生成の模倣を行い「復唱」課題を実施し、その前後で自分の調音器官で実際に音声生成を行った際に、外国語の母音におけるフォルマント周波数が改善するかを調べている。その結果、効果を裏付けるデータが得られているが、ここには音声生成と知覚に加え、手や指の運動や触覚というさらに広い視点が変わってくる例と言えよう。

3 おわりに

本稿では、人間による音声コミュニケーションにおいて音声生成と音声知覚の研究がもともと切り離せるものではないことを述べた。日本音響学会では、2017年3月から生成と知覚の両分野が一緒になって音声コミュニケーションのセッションをスタートさせる。このことが、これらの分野の発展を助けるものと信じている。

謝辞

本稿を執筆するにあたり、協力いただいた皆様から心から感謝申し上げます。内容の一部は日本学術振興会の科学研究費（15K00930）の助成を得た。

参考文献

- [1] P. B. Denes and E. N. Pinson, *The Speech Chain: The Physics and Biology of Spoken Language*, 2nd ed., W. H. Freeman, New York, 1993.
- [2] 箕一彦, “耳は見ている? —音声知覚研究の展望—,” 日本音響学会誌, 72(6), 346-348, 2016.
- [3] J. Ryalls 著, 今富摂子, 荒井隆行, 菅原勉 監訳, 音声知覚の基礎, 海文堂, 2003.
- [4] H. Lane and B. Tranel, “The Lombard sign and the role of hearing in speech,” *J. Speech Hear. Res.*, 14, 677-709, 1971.
- [5] W. Van Summers, D. B. Pisoni, R. H. Bernacki, R. I. Pedlow and M. A. Stokes, “Effects of noise on speech production: Acoustics and perceptual analysis,” *J. Acoust. Soc. Am.*, 84, 917-928, 1988.
- [6] J. C. Junqua, “The Lombard reflex and its role on human listeners and automatic speech recognizers,” *J. Acoust. Soc. Am.*, 93, 510-524, 1993.
- [7] N. Hodoshima, T. Arai and K. Kurisu, “Intelligibility of speech spoken in noise and reverberation,” *Proc. ICA*, 2010.
- [8] T. Arai, N. Warner and S. Greenberg, “Analysis of spontaneous Japanese in a multi-language telephone-speech corpus,” *Acoust. Sci. Tech.*, 28(1), 46-48, 2007.
- [9] K. Maekawa, H. Kikuchi, “Corpus-based analysis of vowel devoicing in spontaneous Japanese: an interim report,” In J. van de Weijer, K. Nanjo, and T. Nishihara, eds., *Voicing in Japanese*, Mouton de Gruyter, pp. 205-228, 2005.
- [10] 河津宏美・前川喜久雄 “子音の調音様式が母音の無声化におよぼす影響—日本語話し言葉コーパスの分析—”, 日本音響学会春季研究発表会講演論文集, 443-444, 2009.
- [11] T. Arai, “Hands-on tool producing front vowels for phonetic education: Aiming for pronunciation training with tactile,” *Proc. of INTERSPEECH*, 1695-1699, 2015.