

脳からみた聞くと話すの共通性

～音声変換技術と脳機能計測による人間の音声コミュニケーションの仕組みの解明～

Speech production and perception share common brain pathways

Investigation of the mechanisms of speech communication by speech conversion and brain imaging



人間情報研究部

廣谷 定男 Sadao Hiroya

プロフィール

NTT コミュニケーション科学基礎研究所 人間情報研究部 主任研究員(特別研究員)。2001年東京工業大学大学院総合理工学研究科修士課程修了。同年日本電信電話株式会社入社。2006年東京工業大学大学院総合理工学研究科博士課程修了。博士(工学)。2007年から2008年ボストン大学客員研究員。音声知覚生体メカニズムの解明および音声信号処理研究に従事。IEEE、日本音響学会、電子情報通信学会、北米神経科学学会、言語神経生物学学会各会員。

聞くことと話すことによる音声コミュニケーションは、日常生活において欠かすことのできない重要なものです。「話す」過程では、脳でイメージした伝えたい意図が調音器官を通じて音声として放射され、「聞く」過程では、音声聴覚器官を経て脳で理解されます。聞くことと話すの目的はそれぞれ異なることから、これまでほとんどの音声研究は聞くことと話すを別々に扱ってきました。

人間の音声脳科学研究においても、「聞く」に関与する脳部位はウェルニッケ野、「話す」はブローカ野というように別々に議論が行われてきました。ところが、近年の脳機能計測の発展に伴い、聞くことと話すに共通した仕組みが脳の中に存在することが分かってきました。この共通した仕組みは、話すときに「聞く」に関与する脳部位の活動、聞くときに「話す」に関与する脳部位の活動を調べることで明らかになりました。つまり、聞くことと話すを別々に扱うのではなく、一緒に扱うことが大切なのです[1]。

話すときに「聞く」に関与する脳部位が活動する

自らが発話した音声を200ミリ秒程度遅らせ、それをヘッドホンで聞きながら発話を行うと、どもるようになる、間延びするなど発話が困難になる現象が起こります。また、音声のタイミングを変換するのではなく、「え」と発話した音声が「あ」と聞こえるように音声の共振周波数をリアルタイムに変換した音声を聞きながら発話を行っても発話に変化が見られます。聞こえた音声に応じて発話が変化するということは、人間は自らの音声をリアルタイムにモニタリングしながら話しているということです。これら変換音声を聞きながら発話を行うときの脳活動を調べると、聞くに関与する脳部位(上側頭回)の活動が増加することが分かりました。つまり、モニタリングで発話に誤りが見つかったときに、脳は発話を正しく修正しようとするのです。この誤りに対する発話の修正量の大きさについて様々な検討が行われていますが、最近私たちは、変換音声の音質が良くなると、修正量が大きくなることを発見しました[2]。

聞くときに「話す」に関与する脳部位が活動する

話すときに「聞く」に関与する脳部位が活動することはモニタリングを考えれば理解しやすいかと思います。しかし反対に、聞くときに「話す」に関与する脳部位が活動するという事は起こりうるのでしょうか。もしそうであれば、音声を聞くと、つられて自らの口が動いてしまうこととなります。ところが、口が動かないまでも、話すに関与する脳部位(運動前野)が音声を聞くときに活動することが明らかになっています。このことは、調音器官の運動をシミュレーションしながら音声を聞いているという可能性を示しています。人間がなぜこのようなことをしながら音声を聞いているのかについてはまだ分かっていませんが、音声が聞き取りにくい場合に運動前野の活動が大きくなることから、音声に足りない情報を補うために調音運動のシミュレーションを行うことで音声理解(what)を助ける、また音声の自然性(how:合成音声らしさ、たどたどしさ)を判断するためでないかと考えられています。

音声リズムを変換する

さて、自然性を変換した音声を聞くときに、話すに関与する脳部位の活動がどのように変化するかは分かっていません。私たちは、日本語母語話者の英語音声のリズムを変換する技術[3]により作成した刺激音声を英語母語話者が聞くときの脳活動計測を行いました(図1)。実験の結果、英語母語話者にとって自然でないリズムの音声を聞くときに、発話に関与する補足運

動野の活動が増加することが分かりました[4]。つまり、聞くにおける発話のシミュレーションは音声の自然性を判断するために行われている可能性があります。

まとめと今後の展望

近年、音声変換技術の発展に伴い、音声変換技術と脳機能計測を組み合わせることで、これまで知られていなかった音声脳情報処理の仕組みが解明されつつあります。音声変換技術は、今後も音声脳科学研究において有効な手段であると考えられます。また、この技術を活用して、ネイティブに聞き取りやすい英語音声に変換することで、日本語母語話者の英語での音声コミュニケーションを支援できると期待されます。

話すときのモニタリングに関わる脳情報処理および聞くときのシミュレーションに関わる脳情報処理を紹介してきましたが、実は、モニタリングもシミュレーションも、聴覚野から上側頭回と角回、縁上回を経由して運動前野に音声を送られるという共通した経路です(図2)。しかも、これらは人間の聞くにおける頑健性や話すにおける安定性に深く関わっていると考えられています。しかし、現在の音声認識および音声合成技術ではこの経路に相当する処理が考慮されていません。もしかすると、これら経路における音声脳情報処理の理解が今後音声認識の性能や音声合成の品質を向上させる鍵となるかもしれません。

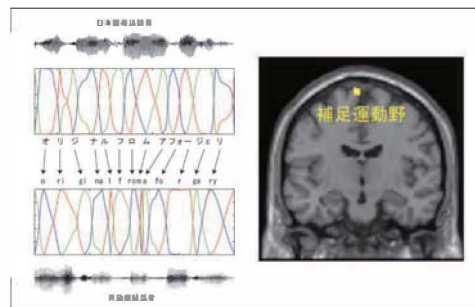


図1:発話リズム変換例(左)。英語母語話者が日本語リズム英語音声を聞いているときの脳活動(右)

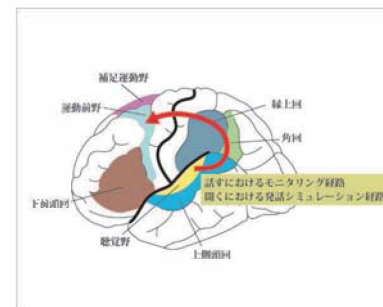


図2:「話す」におけるモニタリング経路と「聞く」における発話シミュレーション経路。「聞く」と「話す」によらず共通した経路

関連文献

- [1] 廣谷定男編著, 第一彦, 辰巳格, 皆川泰代, 持田岳美, 渡辺真澄著, 「聞くことと話すの脳科学」, コロナ社, 2017.
- [2] S. Hiroya, T. Mochida, "Speech sound naturalness alters compensation in response to transformed auditory feedback," in *Proc. J. Acoust. Soc. Am.*, Vol. 140, No. 4, Pt. 2, p. 3228, 2016.
- [3] S. Hiroya, "Non-negative temporal decomposition of speech parameters by multiplicative update rules," *IEEE Trans. Audio, Speech and Lang. Process.*, Vol. 21, No. 10, pp. 2108-2117, 2013.
- [4] S. Hiroya, K. Jasmin, S. Evans, S. Krishnan, M. Ostarek, D. Boebinger, S.K. Scott, "Effects of speaking rhythm naturalness on the neural basis of speech perception," in *Proc. Society for Neuroscience*, 2015.