



令和5年度研究助成 【サウンド技術振興部門】より

発話運動のリズム生成における聴覚フィードバックの役割の解明

北陸先端科学技術大学院大学

先端科学技術研究科人間情報学研究領域

特任助教

上江洲安史

1. はじめに

私たちが他者とコミュニケーションを取る上で音声は不可欠な存在であり、音声を生成するために私たちは発話を行う。発話中、言い間違いやイントネーション・リズムの乱れなど、発話に何らかの誤りを生じることがしばしば起こる。それにもかかわらず、私たちはこの発話の誤りを即座に評価し修正することができる。これは発話の際に生じる様々なフィードバック情報に基づいて、実時間での発話の誤りの評価と修正を行っているからであり、特に発話によって生じる舌・顎・喉頭などの音声器官の運動の様子を体性感覚で、発話音声を自身の耳すなわち聴覚でそれぞれモニタリングすることで、これらのフィードバック情報に基づく安定した発話が実現できていると考えられている。このように「聞く」と「話す」は密接に関係しており、ヒトの聴覚-発話運動制御のメカニズム解明は、私たちの音声コミュニケーションの根幹の更なる理解と、新たな音声コミュニケーション技術の基礎知見の獲得に繋がることが期待される。

私たちが何かを話す際、発話運動にはある種のリズムを生じ、音声を介して聞き手に伝わる。これを発話リズムと呼ぶ。発話リズムは言語によって異なることが知られており、日

本語の発話リズムは世界でも少数のモーラ拍リズムである。また同じ言語を母語とする話者でも、発話リズムには個人ごとの特徴が表れる。誰かと音声コミュニケーションを行う状況の思い浮かべたとき、発話リズムの良し悪しが発話内容の理解だけでなく話し手の印象にも影響することは容易に想像できるだろう。このように、発話リズムは音声コミュニケーションの質を大きく左右する重要な要素だと言える。本研究では「聞く」と「話す」の関係の観点から、発話音声の聴覚フィードバックが発話リズムの生成に対してどのような役割を果たしているのかを解明することが目的である。

2. 研究内容

聴覚-発話運動制御メカニズムの解明を目的とした実験パラダイムでは、発話音声に対して外的に何らかの微小な変化（摂動）を与えたものを聴覚フィードバックし、摂動に対する発話の変化を調べる。有名な実験パラダイムの1つに遅延聴覚フィードバック（DAF: Delayed auditory feedback）が挙げられる¹⁾。DAFでは、話者が発話した音声に僅かな時間遅延を付与して、ヘッドホンなどで話者の聴覚にフィードバックする。DAF下で発話すると、話者は発話のタイミングから僅かにずれた音声を知覚することになる。このとき、連発や伸発といっ

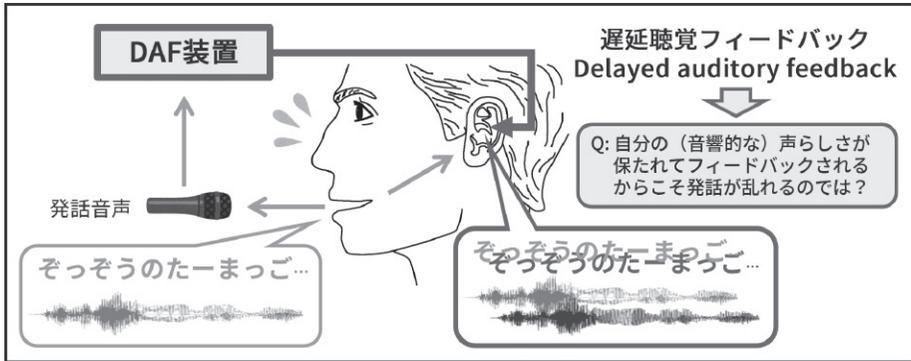


図1 遅延聴覚フィードバック (DAF) と発話の乱れ

た吃音話者に見られる発話の乱れを生じやすくなり、特に遅延量が約200msのDAF下で顕著に生じることがよく知られている(図1)。

DAFを用いた研究の多くでは、フィードバック音声に対して遅延以外を操作しないため、話者自身の音響的な個性がフィードバック音声に十分に含まれていると考えられる。これまでの研究から、発話時に自分の声を聴くと聴覚野の脳活動が抑制されることや、最近では発話時に聴く声の「自分らしさ」が発話行為の主体感に関係していることが明らかになってきた。これらの知見を踏まえると、話者自身の音響的な個性を十分に含んだ音声が遅れて聴覚フィードバックされるからこそ、発話の乱れを引き起こしている可能性がある。本研究ではこの点に着目し、聴覚フィードバック音声に含まれる音響的な個性が発話リズムに重要であるという仮説を立て、音響的な個性を操作したDAF音声が発話リズムの生成にどのように影響するのかを調査する。

これまで、聴覚フィードバックの発話リズムへの影響を調べた数少ない研究では、フィードバック音声の遅延が発話リズムに影響すると言及するのみにとどまっている。さらに、発話リズムは言語による影響も強く受ける。そのため、日本語母語話者を対象に、音声の音響的な個人

性を操作したDAF発話実験を通じて、発話リズムの生成における聴覚フィードバックの役割を多角的に検討する必要があると考える。

3. 現在の成果

現在までに、音響的な個性(以降、話者性)を実時間で変換するシステムを組み込んだDAFを用いて、フィードバック音声の話者性の違いがDAF下での発話に与える影響についての基礎検討を行った(図2)²⁾。

実験は防音室で実施し、日本語母語話者の成人男性9名が参加した。参加者には、ヘッドセットマイクとヘッドホンを装着した状態で、ディスプレイに表示された文章を読み上げるように指示した。発話音声は時間遅延の付与と、ボイスエフェクターによる話者性の操作が行われ、実時間で話者にフィードバックされた。このときの発話音声とフィードバック音声をPCで収録した。実験条件として、遅延量を“30ms”と“200ms”、話者性を“変換なし”と“変換あり”に設定した。ここで“変換あり”条件では、ボイスエフェクターに搭載されたfoとformantの各ツマミを9目盛りを設定した。DAFのない環境でも読み上げづらい文章(例えば「クロレナならコレステロールと別れられるか」など)を対象に、どの参加者も問題なく読み上げられ

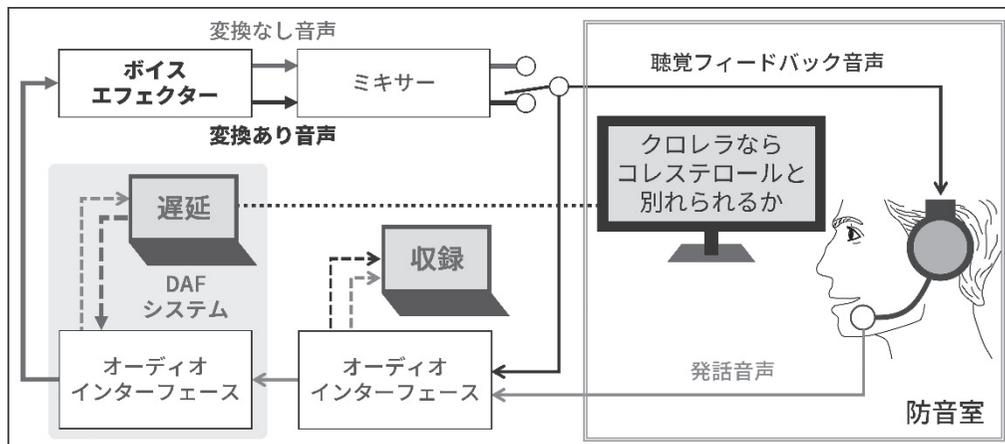


図2 本研究における実験系の概略図の例²⁾

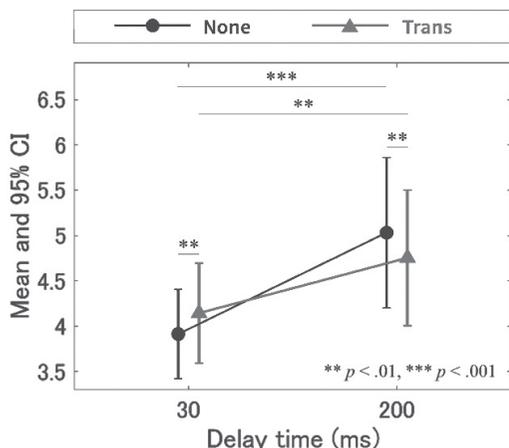


図3 遅延量と話者性の各条件におけるDAF下での発話音声の発話時間長²⁾

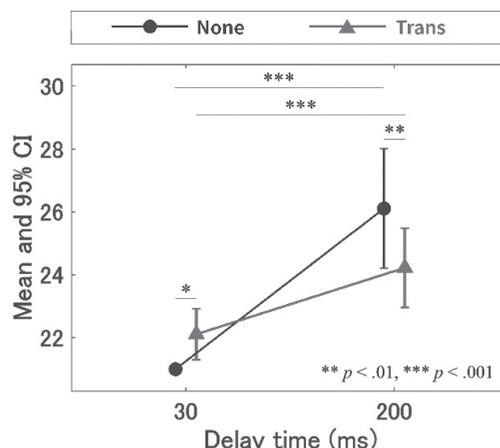


図4 遅延量と話者性の各条件におけるDAF下での発話音声のモーラ数²⁾

ることを練習で確認した後、1文章につき4条件（遅延量2条件×話者性2条件）をランダム順で読み上げる課題を実施した。

収録音声から、発話時間長とモーラ数を抽出し、2×2の反復測定分散分析（RM-ANOVA）を行った（図3および図4）。発話時間長におけるRM-ANOVAの結果から、遅延量×話者性の交互作用、“変換なし”における遅延量、“変換あり”における遅延量、“30ms”における話

者性、“200ms”における話者性の各単純主効果がそれぞれ有意であった。すなわち（1）話者性の変換の有無によらず、DAFの遅延量が増加すると発話時間長も長くなる（2）遅延量“30ms”では、話者性“変換あり”のときに発話時間長がより長くなる（3）遅延量“200ms”では、話者性“変換なし”のときに発話時間長がより長くなることが分かった。（2）と（3）に見られるように、フィードバック音声に含ま

れる話者性の違いが発話に与える影響は、フィードバック音声の遅延量によって異なることが示された。これらの結果はモーラ数においても同様であった。

以上より、典型的な遅延量と話者性の操作を行ったDAF発話実験によって、発話中に聴覚フィードバックされる音声に含まれる話者性が、聴覚-発話運動制御に重要な役割を担うことが示唆された。今後は、本実験で実施した話者性の操作が、音声の音響的特徴をどのように変化させたかを分析するとともに、DAF下における話者性と発話リズムの関係を明らかにすることが課題である。

4. 今後の展望

本研究の完成時には、音声コミュニケーションにとってより良い発話リズムを生成するためには、どのように発話音声を聴覚にフィードバックすればよいのか、その方略を確立できる可能性がある。上述したように、発話リズムは音声コミュニケーションの質に大きく影響する。したがって本研究の成果は、私たちの「聞く」

と「話す」におけるメカニズムの解明だけでなく、日常生活での音声発話サポートや、他言語の発話リズム・音声コミュニケーション学習への応用、吃音をはじめとした発話障がいメカニズム解明・克服といった、音声コミュニケーションを用いる万人のQOLを向上させるための基礎的知見および技術となることが期待される。

謝辞

本研究に対して研究助成を賜りました一般財団法人カワイサウンド技術・音楽振興財団に心より感謝いたします。

参考文献

- 1) Lee, B. S., "Effects of delayed speech feedback," J. Acoust. Soc. Am., 22(6), 824 - 826, 1950.
- 2) 上江洲, "話者性を操作した遅延聴覚フィードバックが発話に与える影響の検討," 日本音響学会秋季研究会, 1-Q-9, 901 - 902, 2023.