

研究テーマ	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	報告書作成者	上江洲 安史
研究従事者	上江洲 安史		
研究目的	音声コミュニケーションにおいて、「聞く」と「話す」は密接に関係しており、私たちは自身の発話音声を聴覚フィードバックによってモニタリングすることで、安定した発話制御を実現している。この聴覚フィードバックは発話リズムの生成にも影響を与え、発話内容の理解や話し手の印象に影響を及ぼすと考えられる。 聴覚-発話運動制御メカニズム解明のパラダイムの一つである遅延聴覚フィードバック(Delayed Auditory Feedback: DAF)によって、フィードバック音声に数百 ms の時間遅れが与えられるとき、発話の非流暢性が増加することがよく知られている。発話の非流暢性は発話音声に含まれる音素やモーラといったマイクロな時間的変化の構造(すなわち発話リズム)の乱れに起因すると考えられる。そのため、DAF 下での発話において、発話音声のマイクロな時間的変化の構造を捉えることで、聴覚フィードバックが発話リズムに及ぼす影響を明らかにできる可能性がある。また、従来の DAF を用いた研究を振り返ると、フィードバック音声の話者自身の音響的な個人性(話者性)を十分に含むことで、むしろ発話の非流暢さをもたらしている可能性がある。具体的には、フィードバック音声の話者性が DAF 下での発話にどのように影響するかは十分に理解されていない。このため、DAF 下でのフィードバック音声の話者性を音響的に操作し、この話者性の違いが発話に与える影響を調べる必要がある。 本研究の目的は、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、DAF 下での発話実験を通して解明することである。特に、フィードバック音声の基本周波数(f_0)およびスペクトルの操作を伴う DAF 下において、発話時間長や発話モーラ数、発話リズムに対する影響を詳細に調査する。これにより、フィードバック音声の音響的な話者性が発話全体および発話リズムの生成にどのように影響するかを明らかにする。 本研究は、聴覚フィードバックが発話制御に与える影響をより深く理解するための新たな知見を提供する。これにより、音声コミュニケーションの質を向上させるための応用が期待される。具体的には、発話リハビリテーションや音声合成技術の改善、さらに吃音などの発話障害の治療法の開発に貢献することが考えられる。また、フィードバック音声の操作が発話リズムに及ぼす影響を明確にすることで、話者の個人性に基づいたカスタマイズされたフィードバックシステムの設計に役立つことが期待される。これらの成果は、音声コミュニケーションの質を高めるだけでなく、発話制御に関する基礎的な理解を深めることにも寄与する。		

研究内容	本研究では、主に以下の事項について検証した。 1) DAF 下での日本語読み上げ時の発話リズムへの影響 [1] DAF 下での日本語文章読み上げ実験を通じて、遅延の有無や遅延量の違いが発話リズムにどのように影響するかを検証した。実験は北陸先端科学技術大学院大学内の防音室で実施され、日本語母語話者の成人男女 10 名(平均年齢 24.7±1.85 歳)が参加した。参加者はヘッドホンを装着し、DAF 下で文章を読み上げる課題に取り組んだ。発話音声はマイクで収録され、PC 上の Audapter システム[1]で時間遅延を付与したフィードバック音声を聴取させた。 実験では、フィードバック音声の遅延量として 0 ms、100 ms、200 ms、300 ms の 4 条件を設定し、各 30 モーラの日本語文章 10 種類を読み上げさせた。収録音声に対して Julius 音素セグメンテーションキットおよび Praat を用いた音素およびモーラ単位でのアノテーションを行い、アノテーション結果から区間長を抽出した。発話リズムは音素およびモーラの区間長に基づく発話リズム指標によって定量化し、条件間での差について統計的検定を行った。 2) フィードバック音声の話者性が発話に与える影響 [2, 3, 4] フィードバック音声の基本周波数およびスペクトルを操作し、話者性を保持した音声と音響的に変換された音声の両方を用いた DAF 実験を行った。実験は北陸先端科学技術大学院大学内の防音室で行われ、日本語母語話者の成人男性 12 名(平均年齢 25.0±1.8 歳)が参加した。参加者はヘッドセットマイクとヘッドホンを装着し、DAF 下で文章を読み上げる課題に取り組んだ。発話音声はヘッドセットマイクで収録され、PC 上の Audapter システムで時間遅延を付与したフィードバック音声を聴取させた。 実験条件として、遅延量を 0ms、100ms、200ms に設定し、各遅延量に対してフィードバック音声の基本周波数の操作 3 条件(変化なし・上昇・下降)およびスペクトルの操作 3 条件(変化なし・伸長・圧縮)を組み合わせた 27 条件を設定した。課題には「クロレナならコレステロールと別れられるか」などの 21 モーラからなる2つの日本語文章を用いた。収録された発話音声は MATLAB で解析し、発話時間長およびモーラ数を算出したのち、条件間での差について統計的検定を行った。 [1] 上江洲ら, “遅延聴覚フィードバックが日本語文章読み上げ時の発話リズムに与える影響”, 日本音響学会聴覚研究会, 2024 年 8 月予定. [2] 上江洲ら, “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発表会, 2023 年 9 月. [3] 上江洲ら, “基本周波数とスペクトルを操作した遅延聴覚フィードバック下での発話変化”, 日本音響学会聴覚研究会, 2024 年 2 月. [4] Uezu et al., “Effects of fundamental frequency and spectral manipulations on speech production under delayed auditory feedback”, Proc. 13th International Seminar on Speech Production, May 2024.
-------------	---

研究のポイント	1) この研究のポイントは、DAF 下での日本語文章読み上げ実験を通じて、遅延の有無や遅延量の違いが発話リズムに与える影響を明らかにすることにある。実験は日本語母語話者 10 名を対象に行われ、0 ms、100 ms、200 ms、300 ms の 4 条件で遅延を設定した。収録音声は音素およびモーラ単位でアノテーションされ、発話リズム指標に基づき定量化し、条件間の差異を統計的に検証した。 2) この研究のポイントは、フィードバック音声の基本周波数およびスペクトルを操作し、話者性を保持した音声と音響的に変換された音声の影響を DAF (Delayed Auditory Feedback) 下で検証する点にある。実験は 12 名の日本語母語話者を対象に行い、遅延量と音声変化の組み合わせによる 27 条件を設定し、発話時間長およびモーラ数を解析した。条件間の差異を統計的に検定することで、発話リズムに対する影響を明らかにした。
研究結果	1) 100 ms 以上の遅延を伴う DAF 下では、発話リズム指標は全体的に増加する傾向が見られた。特に子音に関する発話リズム指標が顕著に変化し、DAF による発話運動制御への影響が強く示唆された。発話リズム指標の増加は音素・モーラ区間長の全体的な時間長やばらつきの増加を意味し、これは発話の非流暢性が高まったことを意味している。これより DAF が日本語発話リズムに与える影響は、音素やモーラの時間的構造の乱れに起因することが示された。 2) フィードバック音声の話者性が発話に与える影響は遅延量によって異なることが示された。特に 200 ms 遅延条件下では、話者自身の話者性を保持したフィードバック音声が発話の乱れを顕著に引き起こすことが確認された。また、フィードバック音声に対する操作としては、基本周波数よりもスペクトルを操作した場合で DAF 効果により影響を与えることも分かった。これより、フィードバック音声の音響的な話者性が発話制御において重要な役割を果たすことが示唆された。
今後の課題	共通する課題の一つとして、DAF 下での発話時における音声器官の同時計測が挙げられる。例えば、Electroglottography による声帯振動の計測や、超音波エコーによる舌の調音運動の計測などがある。これらを DAF 下での発話実験時に同時計測することによって、これまで得られた DAF の発話への影響に関する知見を、音声生成機構の実際の挙動と結び付けて議論することが可能となる。 1) DAF と発話リズムの関係については、発話の非流暢性に大きく寄与すると考えられる要素の一つである発話ポーズ部(無音区間)の変化について検証することや、子音・母音の判別が困難なため除外されたもの(促音・撥音・無声化など)についての検証が課題である。 2) これまでに、DAF と話者性が発話に及ぼす影響については確認された。今後は、発話リズム指標を用いた発話音声の分析が課題である。現在、この分析に取り組んでおり、9 月実施予定の日本音響学会で発表予定である。

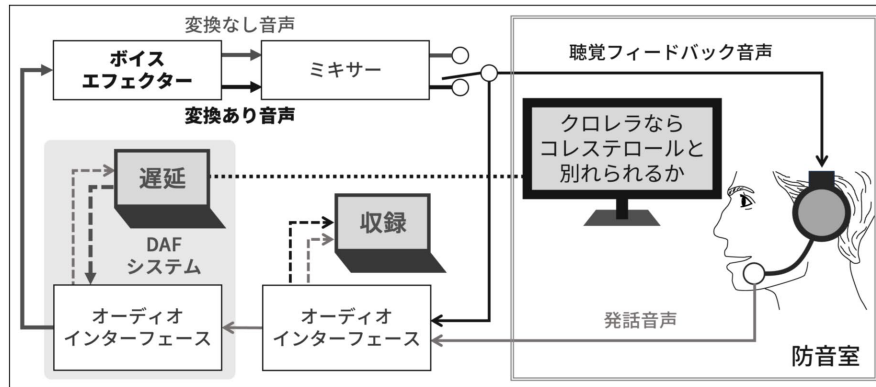


図1. DAF発話実験環境の概略図

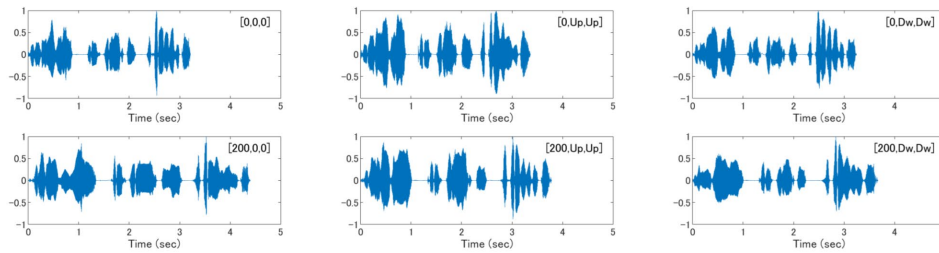


図2. 話者性を操作したDAF下での発話音声の波形

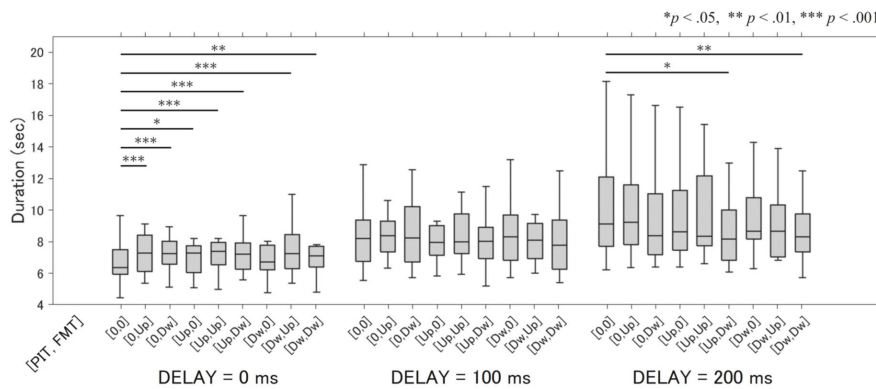


図3. 話者性を操作したDAF下での条件ごとの発話時間長

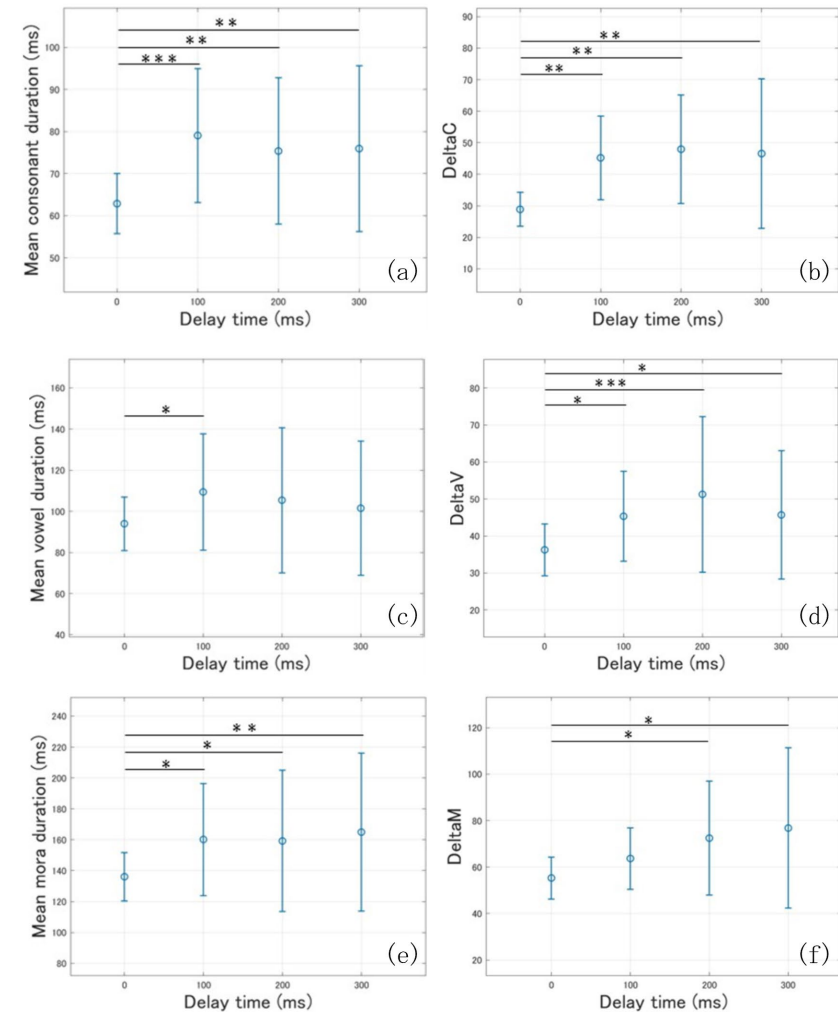


図4. 異なる遅延量における各発話リズム指標の比較:
子音区間の(a)平均と(b)標準偏差
母音区間の(c)平均と(d)標準偏差
モーラ区間の(e)平均と(f)標準偏差