

テクニカルノート

声量制御のためのホワイトノイズ手法と 音声フィードバック手法

竹川 佳成^{1,a)} 原 史也^{1,b)} 平田 圭二^{1,c)}

受付日 2019年11月8日, 採録日 2020年2月4日

概要: 公共の場で友人との会話が盛り上がった際に、周囲に迷惑をかけるほど声が大きくなってしまいう状況や、プレゼンテーションの際に緊張し声が小さくなってしまいう状況など意識的に声量制御ができない場合がある。この問題を解決するために、周囲の雑音大きい騒音下では、人は不随意的に声が大きくなるというロンバード効果を応用し、2つの声量制御手法を提案する。ホワイトノイズをユーザに聞かせることによって発話時のユーザの声量を増幅させるホワイトノイズ手法、ユーザ自身が聞き取る自身の声量を増幅させることで、発話時のユーザの声量を減衰させる音声フィードバック手法を提案する。ホワイトノイズ手法および音声フィードバック手法それぞれにおいて声量変化の効果を検証しその有用性を確認した。

キーワード: 認知心理学, 声量制御, ホワイトノイズ, 音声フィードバック, 不随意

A White Noise Method and a Voice Feedback Method for Voice Loudness Control

YOSHINARI TAKEGAWA^{1,a)} FUMIYA HARA^{1,b)} KEIJI HIRATA^{1,c)}

Received: November 8, 2019, Accepted: February 4, 2020

Abstract: There are situations in which we cannot immediately control our own voice loudness, even though we ought to. For example, in the case of a person who, having a lively conversation with friends in a public place, raises his voice and bothers the people around him, or the case of a person who becomes nervous in an interview and finds his voice becoming quiet. In order to solve these problems, we propose two voice loudness control methods based on the Lombard Effect by which humans involuntarily speak in a loud voice in a noisy environment. The white noise method changes the white noise that the user hears according to the environment sound, to increase the user's voice loudness when speaking, while the voice feedback method changes the 'own voice loudness' that the user hears according to the environment sound, to decrease the user's voice loudness when speaking. We conducted an assessment to clarify the effectiveness of the proposed methods.

Keywords: cognitive psychology, voice loudness control, white noise, voice feedback, involuntary

1. 背景

近年、ウェアラブルコンピューティング技術を活用し、ユーザの状況を考慮した情報提供手法に注目が集まっている。たとえば、Google Glass や Apple Watch などのウェア

アラブル端末を用いて、目的地までのナビゲーション、買い物時のセール情報、歩数や心拍に基づくヘルスマネジメントなど、ユーザにとって有用な情報を提供することでユーザの行動を支援している。このように、既存の多くの情報提供手法は、ユーザに行動の選択肢を与える情報を提示し、特定の行動を促すものが多い。しかし、公共の場で騒いでいる子どもを静かにさせたい場合や、プレゼンテーションや面接など緊張している状態で萎縮せず明瞭に話したい場合など、ユーザが自分自身の行動をただちに制御できない場合がある。したがって、ウェアラブルシステムにより、ユー

¹ 公立はこだて未来大学
Future University Hakodate, Hakodate, Hokkaido 041-8655,
Japan

a) yoshi@fun.ac.jp
b) hara@fun.ac.jp
c) hiratafun.ac.jp

ザに行動を促すだけでなく、ユーザの行動を制御することは重要である。

そこで本研究では、周囲の雑音が大い騒音下では、人は不随意的に声が大きくなるというロンバード効果 [2] に着目した 2 つの音量制御手法の提案とその検証を目的とする。ホワイトノイズをユーザに聞かせることによって発話時のユーザの音量（以降、発話音量と呼ぶ）を増幅させるホワイトノイズ手法を提案する。また、ユーザ自身が聞き取る自身の音量（以降、聴取音量と呼ぶ）を増幅させることで、ユーザの発話音量を減衰させる音声フィードバック手法を提案する。

本論文の構成は以下のとおりである。2 章では関連研究について述べ、3 章では本研究が提案する 2 つの音量制御手法について説明する。4 章では提案する 2 つの音量制御手法の評価実験について述べ、5 章ではまとめと今後の展望について述べる。

2. 関連研究

人の発声モデルに関する研究は古くから多数行われており、その代表的な例として、Speech Chain モデル [1] がある。耳に聞こえてくる周囲の音などの聴覚情報や相手との距離などの視覚情報をもとに発話をコントロールしている。また、音声フィードバックに特化した発声モデルとしてロンバード効果 [2] がある。たとえば、我々は周囲の雑音が大い環境では、自然と声が大きくなる傾向があるが、騒音やホワイトノイズを聞かせると声が大きくなるだけでなく、基本周波数やフォルマント周波数が高くなるなど、様々な音響的特徴が変化することが知られている。また、Hodoshima ら [5] は、静かな環境・雑音のある環境・残響のある環境において発声の明瞭さを調べている。本研究ではロンバード効果において、ホワイトノイズ以外に、自身の音量をリアルタイムに発話者にフィードバックしたときの効果についても調査している点で既存研究と異なる。

ウェアラブルデバイスを用いて人の行動を制御する研究事例も多数存在する。たとえば、虚偽情報による生体情報制御 [7]、VR (Virtual Reality: バーチャルリアリティ) による視覚提示による食欲制御 [8]、振動モータによる振舞い制御 [9]、[10]、心拍情報の視覚提示による演奏テンポの制御 [11] など様々なアプローチで行動の制御を実現している。本研究に特に関連のある音声制御に関する事例として、聴覚遅延フィードバック (以下: DAF) による発話阻害がある [3]。DAF は人工的に自らの発話を聞かせるタイミングを遅らせると、それ以降の発話に影響を与えるという現象である。DAF を吃音者に適用すると、吃頻度が低減することが示されている。また、DAF の適用後も吃音の軽減が持続することがあり、吃音のリハビリテーションに応用されている [4]、[6]。栗原らは、人工的に自分の発話を遅らせて聞かせる聴覚遅延フィードバックにより、肉

体的な苦痛を与えることなく話し手の発話を阻害するシステム SpeechJammer を提案している [12]。発話の制御という点において本研究と Speech Jammer は関わりが深い。Speech Jammer を適用されたユーザは発話を直接的に阻害されることで発話を止めてしまう。本研究のアプローチは、音量に着目し、発話を阻害することなく発話音量の制御を行うという点で異なる。

3. 音量制御手法

本研究の音量制御手法は、ロンバード効果 [2] を応用したホワイトノイズ手法と音声フィードバック手法の 2 つの手法を提案する。以下、これらの手法について図 1 および図 2 を参照しながら詳細に説明する。

ホワイトノイズ手法 ホワイトノイズ手法は、ユーザの発話音量を増幅させるための手法である。図 1 に示すように、ホワイトノイズにマスクされ、自分自身の声が聞こえにくくなる。これにより、ふだん感じている自分自身の声の聞こえ方（主観的な発話音量）を維持しようとし、発話音量が増幅すると考えられる。雑音に近い音であるホワイトノイズをユーザに与え、ユーザ自身の声を相対的に小さく聞こえさせることで発話音量を大きくさせる。たとえば、就職活動や人前でのプレゼンテーションなどにおいて、緊張しないという人は珍しく、特に面接の重要度が増すにつれて感じる緊張は大きくなる。その後、時間が経過するとともに声が小さくなってしまふ。ホワイトノイズ手法を用いて聴取者である面接官に伝わる音量へと制御することで声が聞こえないという問題を解決できる。

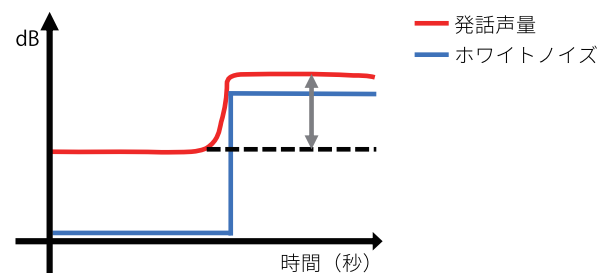


図 1 ホワイトノイズ手法におけるロンバード効果

Fig. 1 Lombard Effect in the white noise method.

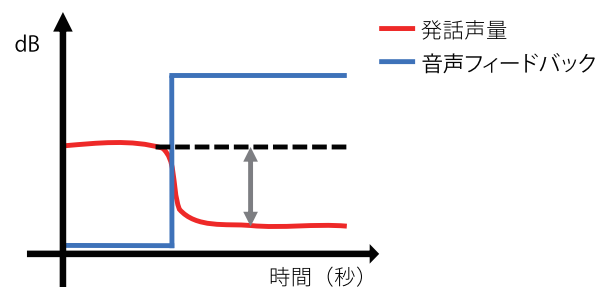


図 2 音声フィードバック手法における逆ロンバード効果

Fig. 2 Reverse Lombard Effect in the voice feedback method.

音声フィードバック手法 音声フィードバック手法は、ユーザの発話声量を減衰させるための手法である。図 2 に示すように、ユーザの発話声量を減衰させるために、ふだん感じている自分自身の声量と比較して現在聞こえる自分自身の声量を増幅させる。ホワイトノイズ手法と同様に、主観的な発話声量を維持しようとし、結果、発話声量が減衰すると考えられる。このように、音声フィードバック手法はロンバード効果 [2] の作用とは逆の作用が働く声量制御手法であり、この作用を本論文では逆ロンバード効果と定義する。たとえば、図書館や機内、車内、病院など、公共の場において大きな声を出すことは好まれない。しかし、話の盛り上がり、酩酊状態などで配慮を忘れ、大きな声で喋ってしまうことがある。本研究の提案する声量制御手法は、不随意的な声量制御であるため、気を配ることができない状態であっても、その場の雰囲気を壊さずに声量を抑えられると期待される。

4. 評価実験

提案する 2 つの声量制御手法の有用性を明らかにするために評価実験を実施する。

4.1 実験目的

実験では、ホワイトノイズ手法ではホワイトノイズ、音声フィードバック手法ではイヤフォンから出力される自分自身の声の大きさ（単位は dB で、以降、増幅量と呼ぶ）をそれぞれ変化させる。このときの声量の変化を検証することで提案するホワイトノイズ手法と音声フィードバック手法の有用性を検証する。

被験者は大学生 7 名である。すべての被験者は本研究の詳細や実験目的を理解しておらず、被験者を募集する際には、文章を音読するというタスクを実施してもらうとだけ伝えた。

4.2 実験システム

評価実験を実施するため、図 3 に示す実験システムを構築した。聴取声量の変化にともなう発話声量の変化を調べることが目的のため、マイクから取得した音声信号を分析する。被験者は防音室に入り、防音室内の 1.7 メートル離れているディスプレイに表示される文章を音読する。マイクで被験者の発話声量を取得し、マイクが生成した音声信号はパソコンとオーディオインタフェースに分配される。パソコンで音声信号を録音し、オーディオインタフェースを通じてホワイトノイズまたは編集済みの音声信号を被験者のヘッドホンへ出力し、被験者はそれを聴取声量として認識する。

ディスプレイは三菱電機社の RDT27WLM-A、パソコンは TOSHIBA 社の dynabook を使用した。マイクはオー

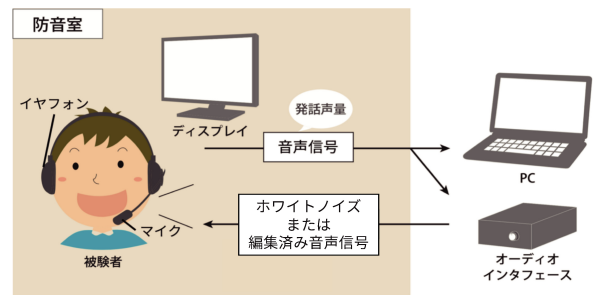


図 3 システム構成

Fig. 3 System structure.

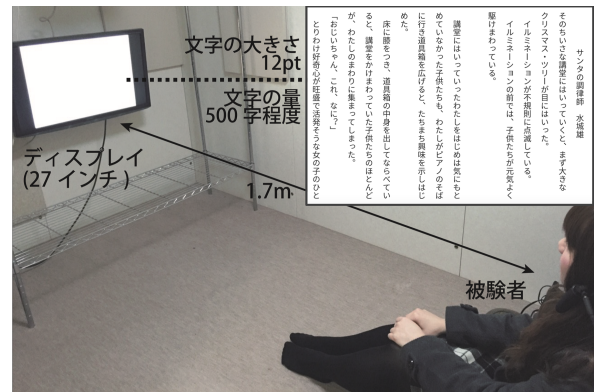


図 4 評価実験の様子

Fig. 4 State of evaluation experiment.

ディオテクニカ社のモノラルマイクロホン AT9913iS、ヘッドフォンは TIMERY 社のヘッドセットマイク HEADSET-A010BK を使用した。オーディオインタフェースはオーディオテクニカ社のマイクロホンアンプ AT-MA2 を使用した。図 4 に実験の様子を示す。

4.3 実験手順

ホワイトノイズ手法ではホワイトノイズ、音声フィードバック手法では自分自身の声の増幅量が変化するなかで、文章を音読するというタスクを行ってもらった。

音読してもらう文章は約 1 分 30 秒で読み切れるニュース記事 [13] である。なお、文章の母音の数などの違いによって発話声量に差が生まれることを防ぐため、同じ文章を利用した。また、十分な発話時間を確保するため、用意した文章を計 6 回繰り返し読んでもらった。6 回中最初の 2 回および最後の 2 回は増幅量を変化させず (0 dB)、3 回目および 4 回目の増幅量を変えた。したがって、増幅量を変えながら音読している時間は約 3 分である。これを 1 試行として、ホワイトノイズ手法と音声フィードバック手法それぞれで 4 試行ごとに行い、計 8 試行を行った。また、各手法で適用する増幅量の試行順番を表 1 および表 2 に示す。各試行の間には十分な休憩時間を確保し、被験者が疲れなように十分配慮した。

さらに、被験者の発話が独り言(他者に伝わる必要のない、声量調整の必然性が薄い発話)にならないように、

表 1 ホワイトノイズ手法の増幅量

Table 1 Amplification in the white noise method.

被験者	試行 1	試行 2	試行 3	試行 4
A	4 dB	16 dB	8 dB	24 dB
B	16 dB	8 dB	24 dB	4 dB
C	8 dB	16 dB	4 dB	24 dB
D	24 dB	8 dB	16 dB	4 dB
E	8 dB	4 dB	16 dB	24 dB
F	4 dB	24 dB	8 dB	16 dB
G	16 dB	24 dB	4 dB	8 dB

表 2 音声フィードバック手法の増幅量

Table 2 Amplification in the voice feedback method.

被験者	試行 1	試行 2	試行 3	試行 4
A	16 dB	8 dB	4 dB	24 dB
B	4 dB	16 dB	24 dB	8 dB
C	8 dB	24 dB	4 dB	16 dB
D	8 dB	24 dB	16 dB	4 dB
E	4 dB	8 dB	24 dB	16 dB
F	24 dB	16 dB	4 dB	8 dB
G	16 dB	4 dB	8 dB	24 dB

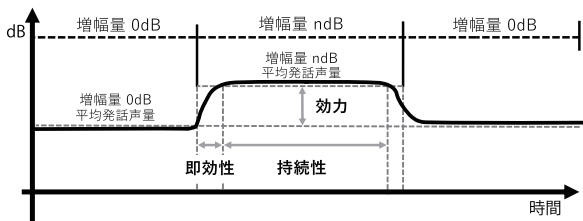


図 5 ホワイトノイズ手法における評価項目のモデル

Fig. 5 Model of evaluation points in the white noise method.

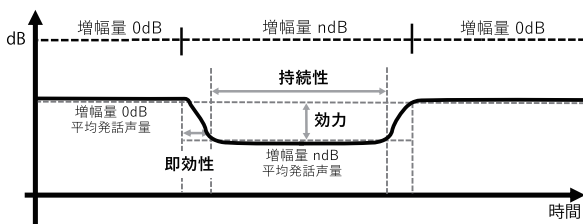


図 6 音声フィードバック手法における評価項目のモデル

Fig. 6 Model of evaluation points in the voice feedback method.

「ディスプレイの位置に人がいると想定して音読するようにしてください」と被験者に指示した。

4.4 評価項目

本実験の評価項目は下記の 4 つである。効力、即効性、持続性は図 5、図 6 を用いて説明する。

汎用性 増幅量適用前から増幅量適用中において、発話声量が変化した人数をカウントすることで評価する。

効力 効力は増幅量適用前の平均発話声量と増幅量適用中の平均発話声量の差分で評価する。

即効性 即効性は増幅量適用中、何秒後に被験者の発話声

量が増加したかを算出することで評価する。

持続性 持続性は増幅量適用中、被験者の発話声量が増加してから何秒間持続するかを算出することで評価する。

4.5 実験結果

マイクから取得した音声のサンプリング周波数は 44.1 kHz で、量子化ビット数は 16 bit とした。また、被験者の音声信号から、振幅の絶対値を取得し、100 区間の移動平均を求めた（以下、振幅絶対値平均と呼ぶ）。振幅絶対値平均は最大 1 から -1 まで変化し、その取得した振幅絶対値平均をデシベルに変換した。その後、評価項目ごとに分析した。各評価項目の結果を図 7 に示す。

汎用性 ホワイトノイズ手法と音声フィードバック手法の 4 パターンの増幅量すべてにおいて、被験者 7 名全員の発話声量は増幅量適用前から増幅量適用中で発話声量が増加した。

効力 ホワイトノイズの増幅量が大きくなっていくと、効力が大きくなることが分かった。すなわち、発話声量が増加していきることが観測された。また、増幅量ごとに Holm の多重分析を適用した結果、4 dB と 24 dB、16 dB と 24 dB に有意差があった ($F(3, 18)$, $p < .05$)。

音声フィードバックの増幅量が大きくなっていくと、効力が大きくなっていくことが分かった。すなわち、発話声量が小さくなっていくことが観測された。また、増幅量ごとに Holm の多重分析を適用した結果、4 dB と 16 dB、4 dB と 24 dB に有意差があった ($F(3, 18)$, $p < .05$)。

即効性 ホワイトノイズの増幅量が大きくなるほど、16 dB 以外、発話声量の変化が短くなり、即効性が強くなった。また、増幅量ごとに Holm の多重分析を適用した結果、4 dB と 24 dB に有意差があった ($F(3, 18)$, $p < .05$)。

聴取声量の増幅量が大きくなるほど、8 dB 以外、発話声量の変化が短くなり、即効性が強くなった。また、増幅量ごとに Holm の多重分析を適用した結果、8 dB と 24 dB に有意差があった ($F(3, 18)$, $p < .05$)。

持続性 ホワイトノイズの増幅量が大きくなるほど、持続性が強くなった。また、増幅量ごとに Holm の多重分析を適用した結果、4 dB と 16 dB および 4 dB と 24 dB に有意差があった ($F(3, 18)$, $p < .05$)。

音声フィードバックの増幅量が大きくなるほど、16 dB 以外、持続性が強くなった。また、増幅量ごとに Holm の多重分析を適用した結果、4 dB と 24 dB に有意差があった ($F(3, 18)$, $p < .05$)。

4.6 考察

効力 ホワイトノイズ手法および音声フィードバック手法

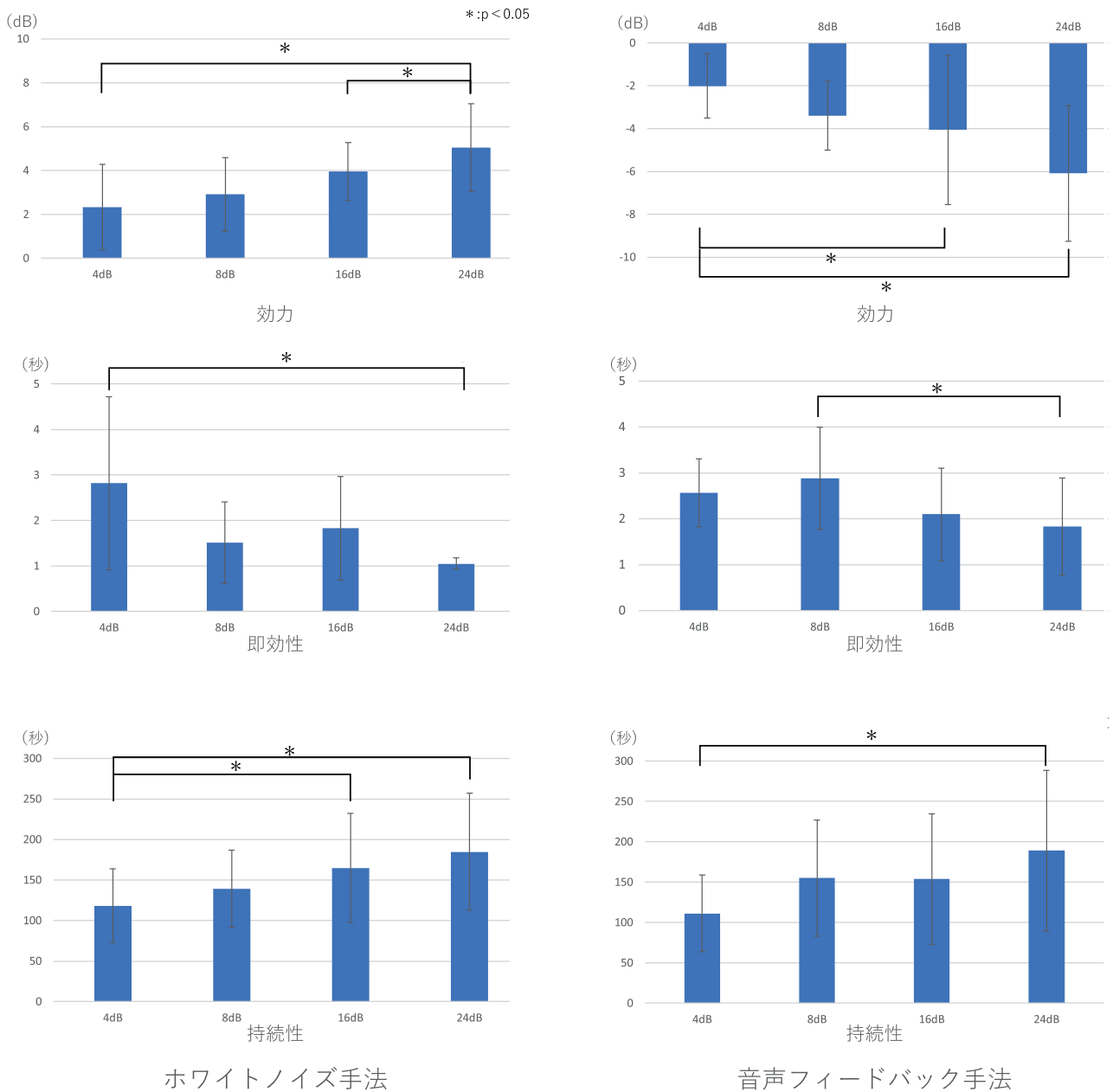


図 7 実験結果
Fig. 7 Results.

いずれも、すべての増幅量で増幅量適用中の効力は増幅量適用前の効力より大きくなることが観測され、増幅量が大きくなるにつれ効力も大きくなる。しかし、本実験の被験者は、実験室内の静かな環境で発話をしているが、周りが騒音環境であったり、ユーザ自身が興奮状態になっている場合にホワイトノイズ手法や音声フィードバック手法が有効かどうかを確かめる必要がある。

即効性 即効性はどちらの手法も強く、被験者は聴取音量の変化に対して素早く順応し、発話できていた。また、ホワイトノイズの方が全体的に即効性が高いという結果となった。ホワイトノイズを与えられた被験者は、ホワイトノイズを雑音と素早く認識し、その雑音に合

わせて発話音量を変化させたと考えられる。一方、音声フィードバックでは被験者自身の声を用いて音量を制御する。したがって、音声フィードバックの増幅量が増幅量適用直後に被験者は自身の声を聞き慣れているために、若干の認識遅れが生じたと考えられる。ホワイトノイズ手法および音声フィードバック手法いずれも、特定の増幅量を適用したときに、数百ミリ秒ではあるものの即効性の鈍化が観測された。有意差がなかったため、個人差が結果に影響したと考えられるが、被験者数を増やすことで、この原因を究明する必要がある。

持続性 持続性はどちらの手法も最低でも約2分ある。その中で一番大きい増幅量である24dBは、どちらの手

法も約3分発話音量を維持できた。発話音量の減衰は、提示されている聴取音量に対する慣れなどが考えられるが、詳細な分析が必要である。

5. まとめ

本研究では、Lombard 効果を応用した音量制御手法を提案した。実験システムを構築し、ホワイトノイズ手法および音声フィードバック手法それぞれにおいて被験者にフィードバックする増幅量を変化させた場合における音量変化の効果について、即効性・効力・持続性・汎用性という4つの評価指標をもとに検証しその有用性を確認した。

今後は、提案手法適用後の後遺症に関する分析、被験者数を増加させ、プレゼンテーションや醜態状態といった被験者の環境や状況を変化させた評価実験の実施や、提案手法を用いたウェアラブルな音声フィードバックシステムの構築などがある。

謝辞 評価法について適切なアドバイスをしてくださった寺井あすか准教授（公立はこだて未来大学）に感謝いたします。本研究は JSPS 科研費 JP16K12560, JP15K00279, JP19H0 4157 の助成を受けたものです。

参考文献

- [1] Denes, B.P. and Pinson, N.E.: The Speech Chain: The Physics and Biology of Spoken Language, W.H. Freeman and Company (1993).
- [2] Lane, H. and Tranel, B.: The Lombard Sign and the Role of Hearing in Speech, *Journal of Speech and Hearing Research*, Vol.14, pp.677–709 (1971).
- [3] Lee, B.S.: Effect of Delayed Speech Feedback, *Journal of the Acoustical Society of America*, Vol.22, pp.824–826 (1950).
- [4] Stuart, A., Kalinowski, J. and Rastatter, M.P.: Effect of Monaural and Binaural Altered Auditory Feedback on Stuttering Frequency, *Journal of the Acoustical Society of America*, Vol.101, No.6 (1997).
- [5] Hodoshima, N., Arai, T. and Kurisu, K.: Intelligibility of Speech Spoken in Noise and Reverberation, *Proc. International Congress on Acoustics*, pp.3632–3635 (2010).
- [6] 遠藤 眞：成人性吃音者の流暢性に対する聴覚遅延フィードバック残効, 日本特殊教育学会特殊教育学研究, Vol.13, No.2, pp.35–42 (1975).
- [7] 中村憲史, 片山拓也, 寺田 努, 塚本昌彦：虚偽情報フィードバックを用いた生体情報の制御手法, 情報処理学会論文誌, Vol.54, No.4, pp.1433–1441 (2013).
- [8] Narumi, T., Ban, Y., Kajinami, T., Tanikawa, T. and Hirose, M.: Augmented Perception of Satiety: Controlling Food Consumption by Changing Apparent Size of Food with Augmented Reality, *Proc. CHI Conference on Human Factors in Computing Systems*, pp.109–118 (2012).
- [9] Rekimoto, J.: Traxion: A Tactile Interaction Device with Virtual Force Sensation, *Proc. Association for Computing Machinery Symposium on User Interface Software and Technology*, pp.427–431 (2013).
- [10] Watanabe, J., Hideyuki, Ando. and Maeda, T.: Shoe-shaped Interface for Inducing a Walking Cycle, *Proc. International Conference on Augmented Tele-Existence*,

pp.30–34 (2005).

- [11] Suzuki, D., Takegawa, Y., Terada, T. and Tsukamoto, M.: A Heart Presentation System for Keeping Music Tempo in Live Performance, *Proc. Institute of Electrical and Electronics Engineers Global Conference on Consumer Electronics*, pp.177–181 (2013).
- [12] 栗原一貴, 塚田浩二: SpeechJammer: 聴覚遅延フィードバックを利用した発話阻害の応用システム, インタラクティブシステムとソフトウェアに関するワークショップ論文集, pp.77–82 (2010).
- [13] 機械に責任を持たせることができるか, 入手先 (<https://webronza.asahi.com/science/articles/2016111600001.html>) (参照 2019-08-18).



竹川 佳成 (正会員)

2007年大阪大学大学院情報科学研究科博士課程修了。同年神戸大学自然科学系先端融合研究環重点研究部助教。2012年公立はこだて未来大学システム情報科学部助教。2014年公立はこだて未来大学システム情報科学部

准教授, 現在に至る。2011年にはMIT Media Lab.にて Assistant Visiting Professor を兼務。博士(情報科学)。教育工学, エンタテインメントコンピューティング, 音楽情報科学の研究に従事。



原 史也

2017年公立はこだて未来大学システム情報科学部情報アーキテクチャ学科卒業。2019年同大学大学院博士前期課程システム情報科学研究科修了。音楽情報科学に関する研究に興味を持つ。



平田 圭二

1987年東京大学大学院工学系研究科情報工学専門課程博士課程修了。工学博士。同年NTT基礎研究所入所。1990～1993年(財)新世代コンピュータ技術開発機構(ICOT)に出向。2011年より公立はこだて未来大学教授。

1993年音楽情報科学研究会初代主査。2005～2007年および2011～2013年本会理事。2010～2015年デジタルプラクティス誌編集委員長。本会フェロー, シニア会員。現在, 音楽情報学に加え, うつ病家族看護者のICT支援, スマートシティの研究に従事。