

後藤崇公, 井上豪, ©大畑典子, 程島奈緒, 荒井隆行 (上智大・理工)

1. はじめに

講演や講義に用いられる室内において, スピーチ(話)の明瞭度は極めて重要となる. 実環境下で音声の拡声が行われるとき, 音声の明瞭度が劣化することがある. その原因の1つとして, かかりすぎる残響によって先行する音素に付加された残響の尾が後続の音素に影響を与える overlap - masking が考えられている [1, 2].

残響環境下において音声明瞭度を改善する手法としては, 主に microphone-array, post-processing, pre-processing の3つが挙げられる. ここで pre-processing は, マイクロフォンとスピーカの間で行われることから, PA (Public Address) の一部に取りこむことで実用化が可能である.

荒井らは, 残響による overlap-masking の影響を軽減するための pre-processing として音声の定常部を抑制する処理を行い, 音声明瞭度の改善を得た [3, 4]. 先行研究 [6, 7] ではコンピュータ上で, いくつかの残響時間を持つインパルス応答を人工的に作成し, 防音室内でヘッドホンを通して聴取実験を行ったところ, 音声明瞭度の改善を得た.

本論文の目的は, 定常部抑圧処理により音声明瞭度の改善が実環境下 (上智大学大講堂) においても実現されるのかどうかを調べることである. 以下では, 聴取実験の結果を示し, 先行研究 [6, 7] の結果と処理の効果について検討する.

2. 聴取実験

2.1 上智大学大講堂の残響時間の測定

聴取実験を行うにあたり, 上智大学 10 号館大講堂 (822 名収容) での残響時間を測定した. 擬似インパルス応答特性の測定結果を基に残響曲線を得た. EDT (early-decay-time) から残響時間を算出した結果, 約

1.2s であることが分かった.

2.2 定常部抑圧処理

本論文で使用した定常部抑圧処理は, 先行研究 [3, 4, 6, 7] で用いた処理と同じものである. 本処理は音声のスペクトル遷移を表す D を計算し, D が一定の閾値より小さい箇所を定常部とした. D は, Furui によって提案されたパラメータ [5] と基本的には同じであり, 帯域分割された信号の時間包絡の対数に対応する回帰係数を複数の帯域にわたり 2 乗平均したものをを用いた. 定常部とみなされた箇所においては, 波形の振幅を 40% に抑圧する処理を施した.

音声情報の損失と overlap-masking 量低下にはトレードオフが存在すると考えられているので, バランスの取れる最適なパラメータ値を探ることが重要である.

2.3 刺激

原音声は, ATR 研究用日本語音声データベースを使用した. これより日本語の単音節 CV (子音-母音) をターゲットとし, キャリアセンテンス「題目としては__といます」に挿入したもの (話者: MAU, 40 歳男性) を作成した. ここで V として /a/ を, C として /p/, /t/, /k/, /b/, /d/, /g/, /s/, /ʃ/, /h/, /tʃ/, /dz/, /dʒ/, /m/, /n/ の 14 種類を用いた. 刺激音としては, 原音声を処理なし, 原音声に定常部抑圧処理を施した刺激を処理ありとし, 計 2 セットを作成した. 被験者全員に対してはランダムに配列した計 56 刺激 (14CV × 処理 2 種類 × 2 回) が同時に示された.

2.4 被験者

被験者は日本語を母語とする健聴者 24 人 (男性 12 人, 女性 12 人 平均年齢 23.5 才) であった. 被験者には, ステージ正面後方ブロックに指定した座席に座ってもらい, 隣接しないよう席の位置を適当にずらした.

**The effect of pre-processing for improving speech intelligibility in the Sophia University lecture hall*

By Takahito Goto, Tsuyoshi Inoue, Noriko Ohata, Nao Hodoshima and Takayuki Arai. (Sophia Univ., Tokyo, Japan)

表 1. 上智大学大講堂での単音節明瞭度試験

	処理なし	処理あり
正解率(%)	62.7	71.8

2.5 手順

各試行において 5 秒間隔で 1 度だけ刺激を提示し、次の刺激が提示される間に 14 単音節が書かれたリスト(かな表示)の中から強制的に 1 つ選択し、回答用紙に書いてもらった。

3. 結果・考察

表 1 に処理条件による正解率の平均値を示す。処理なし、処理ありに対して t 検定を行ったところ、有意に差が示された [$p < 0.001$]。表 1 より、処理を施したほうの正解率が高いことから、定常部抑圧処理を行うことによって明瞭度が改善されたことが示された。

先行研究[6, 7]では、24 種類の単音節を用い、0.4~1.3 s までの残響時間を持つインパルス応答を作成し、防音室内で聴取実験が行われた。原音声は、V として /a/ と /i/ を用いた以外は本稿の実験で用いたものと全て同じものである。結果に対して分散分析を行ったところ、処理による主効果は有意であると示された [$p < 0.001$]。また正解率は全ての残響条件において処理ありの方が高くなり、残響時間 0.8~1.2 s の間では処理による有意差がみられた。

実環境として使用した講堂の残響時間が約 1.2 s であることを踏まえ、先行研究[7]より残響時間が 1.2 s の結果を見ると、処理なしの正解率は 55.1%、処理ありの正解率は 64.2% となり、定常部抑圧処理を施すことによって正解率が向上した。

Moncur and Dirks は monaural 受聴と binaural 受聴について単語理解度試験の正解率を比較し、後者が 14% ほど前者を上回ることを確認した[8]。さらに同じ実験条件で monaural 受聴と diotic 受聴の比較もしているが、差は 2% であった[8]。本実験(dichotic 環境)における処理なしの正解率は先行研究(diadic 環境)[6, 7]の正解率に比べ高かったが、これは[8]の報告と似た傾向である。また、定常部抑圧処理を施した音声においても同様の傾向が確認された。さらに、定常部抑圧処理による処理あり・処理なしそれぞれの正解率の差は変わらず、どちらも処理ありの正解率の方が高かったことから、diadic 環境・dichotic 環境のいずれにお

いても、処理によって同程度の明瞭度の改善が得られたと言える。

以上より、荒井らの提唱する処理が実際の講堂においても音声明瞭度を改善する前処理として有効であるということが確認され、実験室環境の diotic 受聴の場合と同じ刺激を用いて、両耳環境でも効果を得ることが証明できた。

4. おわりに

今回の単音節明瞭度試験によって、実環境下においても定常部抑圧処理による前処理が音声明瞭度劣化を軽減できると初めて確認することができた。今後は、定常部に対する抑圧の度合いを変化させた処理なども用い、明瞭度の改善が得られる残響時間の範囲を調査するとともに、実用化に向けて取り組んでいきたい。

謝辞

本研究を進めるにあたり、講堂を提供して頂いた上智大学関係者のみなさま、聴取実験に参加して頂いたみなさまに心から感謝申し上げます。

参考文献

- [1] R. H. Bolt and A. D. MacDonald, "Theory of speech masking by reverberation," *J. Acoust. Soc. Am.*, 21, pp. 577-580, 1949.
- [2] A. K. Nabelek and L. Robinette, "Influence of precedence effect on word identification by normally hearing and hearing-impaired subjects," *J. Acoust. Soc. Am.*, 63, pp. 187-194, 1978.
- [3] 荒井隆行, 木下慶介, 程島奈緒, 楠本亜希子, 喜田村朋子, "音声の定常部抑圧の残響に対する効果," 日本音響学会秋季研究発表会講演論文集, 1, pp. 449-450, 2001.
- [4] T. Arai, K. Kinoshita, N. Hodoshima, A. Kusumoto and T. Kitamura, "Effects on suppressing steady-state portions of speech on intelligibility in reverberant environments," *Acoustical Science and Technology*, 23, pp. 229-232, 2002.
- [5] S. Furui, "On the role of spectral transition for speech perception," *J. Acoust. Soc. Am.*, 80(4), pp. 1016-1025, 1986.
- [6] 井上豪, 程島奈緒, 荒井隆行, 木下慶介, 楠本亜希子, "定常部抑圧による異なる残響環境下での音声明瞭度の改善," 日本音響学会秋季研究発表会講演論文集, 1, pp. 377-378, 2002.
- [7] N. Hodoshima, T. Inoue, T. Arai and A. Kusumoto, "Suppressing steady-state portions of speech for improving intelligibility in various reverberant environments," *Proc. China-Japan Joint Conference on Acoustics*, pp. 199-202, 2002.
- [8] John P. Moncur and Donald Dirks, "Binaural and monaural speech intelligibility in reverberation," *Journal of Speech and Hearing Research*, 10, pp. 196-195, 1967.