

◎宮内裕介, 程島奈緒, 安啓一, 林奈帆子, 井上豪, 荒井隆行 (上智大・理工), △進藤美津子 (上智大・外国語)

1. はじめに

講演ホールのように残響時間の比較的長い広い室内においては音声の明瞭度が低下し、スピーチの内容を理解することが難しくなることがある。その原因の1つとして、先行する音声区間に付加された残響の尾が後続する音声区間に影響を与える *ovrelap-masking* が考えられている[1,2]。

残響環境下における音声の明瞭度の劣化を防ぐ方法としては、音声に残響が付加される前に処理を行う *pre-processing* と、音声に残響が付加された後で処理を行う *post-processing* がある。*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) の一部に取り込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 *pre-processing* として音声の定常部を抑圧する処理を行い、音声明瞭度の改善を得ている[3,4]。程島らは健聴者を対象に残響時間をパラメータとした聴取実験を行った。その結果、定常部抑圧処理により残響時間 0.7s~1.2s で音声明瞭度が有意に改善された[5,6]。

一方、我々の聴覚は加齢に伴い聴力の低下と時間分解能の低下が生じる。高齢者は健聴な若者よりも残響による明瞭度への影響を受けやすいと報告されている[7]。

本論文の目的は、高齢者を対象に聴取実験を行い、定常部抑圧処理の効果を調べることである。以下に実験結果を示すと共に、先行研究[5,6]の結果と比較しながら処理の効果について検討する。

2. 聴取実験

2.1 被験者

被験者は日本語を母語とする 56 歳から 90 歳の計 50 名 (男性 21 名、女性 29 名) であった。

2.2 刺激

原音声は、ATR 研究用日本語音声データベース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」に挿入して作成した。ここで V として /a/ を、C として /p/, /t/, /k/, /b/, /d/, /g/, /s/, /ʃ/, /h/, /dz/, /dʒ/, /tʃ/, /m/, /n/ の 14 種類を用いた。

実験に用いたインパルス応答は上智大学 10 号館講堂で測定したもの (残響時間 1.3s) と東京都東大和市大ホールのインパルス応答を人工的に残響時間 1.0s にしたものである。

原音声に上記のインパルス応答を畳み込むことによって擬似的な残響環境を実現し、刺激音を作成した。実験には原音声に残響を畳み込んだ刺激セット (処理なし) と原音声に対して定常部抑圧処理を行った後に残響を畳み込んだ刺激セット (処理あり) の 2 種類を使用した。

2.3 定常部抑圧処理

本論文で用いた定常部抑圧処理は先行研究[3-6]で用いたものと同様のものである。音声の定常部に含まれるエネルギーは比較的大きいが音声の情報量は少ない。一方音声の遷移部は音声知覚に関して重要な役割を果たしていることが分かっている。そこで音声の定常部をあらかじめ抑圧することにより、残響による *overlap-masking* の影響を軽減するというのが定常部抑圧処理[3,4]である。なお、本実験では先行研究[3-6]と同様に、定常部を抑圧する割合は 40% にした。

2.4 実験手順

実験は防音室内のコンピュータを用いて行った。刺激音はヘッドフォン (STAX SR-303) を用いて提示し、

*Pre-processing for improving speech intelligibility in reverberant environments: The effect of steady-state suppression on elderly people.

By Yusuke Miyauchi, Nao Hodoshima, Keiichi Yasu, Nahoko Hayashi, Tsuyoshi Inoue, Takayuki Arai and Mitsuko Shindo (Sophia Univ., Tokyo, Japan)

表 1. 実験結果の正解率 (%)

	東大和ホール		上智大学講堂	
	処理なし	処理あり	処理なし	処理あり
本実験	34.4	45.8	37.3	44.9
先行研究[6,7]	58.1	69.5	56.1	62.2

刺激音のレベルは各被験者に合わせて調整した。実験中は使用した 14 種類の CV を選択肢として示し、刺激音を一度だけ提示した後に被験者に選択肢の中から一つ選択させ書き取らせた。以上の試行を計 56 刺激 (14 単音節×残響 2 種類×処理あり/なし) に対して行った。なお、刺激はランダムに提示し、回答時間は制限せず被験者のペースで実験を進めた。

2. 結果・考察

表 1 に各残響環境における処理ありと処理なしの正解率の平均値と、比較のため先行研究[5,6]における正解率の平均値を示す。表 1 における先行研究の正解率は健聴者を被験者として本実験と同様の刺激 14 種類、同様のインパルス応答、同様の受聴条件 (防音室内でヘッドフォンを使用) を用いた時のものである。

高齢者を対象とした本実験の正解率は健聴者を対象とした先行研究の正解率と比較すると、両残響条件において処理なし、処理あり共に正解率は低くなる結果となった。Fitzgibbons and Gordon-Salant の報告[7]より、加齢により聴力が衰え時間分解能が低下し、残響の影響を受けやすくなることから本実験の結果は予測されるものであった。

本実験の各残響条件において *t* 検定を行ったところ、先行研究同様に、処理ありの正解率が処理なしの正解率に比べて有意に高いことが示された ($p < 0.01$) ことから、定常部抑圧処理を行うことによって明瞭度が健聴者に加え高齢者に対しても改善されることが示された。

健聴者よりも残響の影響を受けやすいとされる高齢者の処理による音声明瞭度の改善の度合いは健聴者のそれよりも大きかった。これは定常部抑圧処理により、健聴者に比べて大きく劣化していた音声明瞭度が改善された結果と思われる。

4. おわりに

本論文では荒井らにより提案された定常部抑圧処理

[3,4]の高齢者に対する効果について検討した。今回の実験では先行研究[5,6]において健聴者に対して処理が有効であった残響条件において、高齢者に対しても処理による効果が確認でき、改善の度合いもより大きいものであった。

今後は実環境 (ホールや講堂など) における聴取実験を行い、本研究や先行研究[5,6]の結果を受聴環境、被験者に対して比較し、定常部抑圧処理の効果を検討していきたい。

謝辞

本研究は日本学術振興会科学研究費補助金(A-2, 16203041)の助成を受け行ったものである。本研究を進めるにあたりインパルス応答のデータを提供していただいた東京大学生産技術研究所橋秀樹先生、上野佳奈子さん、横山栄さん、実験に協力していただいた皆様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] N. Hodoshima, T. Goto, N. Ohata, T. Inoue and T. Arai, "The effect of pre-processing approach for improving speech intelligibility in a hall: Comparison between diotic and dichotic listening conditions," Acoustical Science and Technology, vol. 26, No. 2, 2005.
- [6] 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.
- [7] P. J. Fitzgibbons and S. Gordon-Salant, "Profile of auditory temporal processing in older listeners," J. Speech, Language, and Hearing Research, 42, pp. 300-311, 1999.