

残響環境下における音声明瞭度改善のための2種類の定常部処理の評価*

◎中田有貴, 村上善昭, 林奈帆子, 宮内裕介, 程島奈緒, 荒井隆行 (上智大・理工), 栗栖清浩 (TOA)

1 はじめに

残響時間が長いホールでは音声の明瞭性が低下し, 聞き取りが難しくなることがある。その原因の1つに, 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 overlap-masking が挙げられる[1]。

荒井ら[2,3]は, 残響による overlap-masking の影響を軽減するための前処理として音声の定常部を抑圧する処理(定常部抑圧処理)を行い明瞭度改善を示唆する結果を得た。そして林ら[4,5]は電氣的に残響時間を変えることのできる XEBEC ホール(神戸)にて残響時間 2.6~3.3 s で健聴者を対象に聴取実験を行った。その結果, 処理の効果は見られなかったが, これは残響時間が長すぎることによる床効果が起こったためだと思われる。一方, 音声にある長さの零系列を挿入する処理(定常部零挿入処理)を Arai[6]が提案し, 同じホールにおいて音声明瞭度の改善を得ている。

本論文の目的は, 2種類の定常部処理の効果を擬似残響環境下(diotic 受聴)において調べることである。今回は比較的短い残響時間も加え, diotic 受聴とホールでの dichotic 受聴(先行研究[4,5,6])を比較した。

2 聴取実験

2.1 被験者

日本語を母語とする 20~23 歳の健聴者 26 名(男性 14 名, 女性 12 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。V として/a/, 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 XEBEC ホールで測定したもの(残響時間 0.7~3.3 s)であり, 刺激音は, 原音声に上記のインパルス応答を畳み込むことで作成した。実験には原音声に残響を畳み込んだ刺激(処理なし)と原音声に定常部処理を施した後, 残響を畳み込んだ刺激(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[2-5,8,9]で用いたものと同じである。本処理はエネルギーは大きい, 音声知覚に必要な情報量が遷移部に比べて少ない定常部を予め抑圧することにより, overlap-masking の影響を軽減することができる[2,3]。

2.4 定常部零挿入処理

本論文で用いた定常部零挿入処理は先行研究[6]で用いたものと同様のものである。前節と同じ方法で音声の定常部を検出し, ある長さ(T_z)の零系列を挿入することで音節と音節を時間的に離す。これにより overlap-masking の影響を軽減する[6]。本実験では先行研究[6]と同様に T_z を 50 ms, 100 ms とした。

2.5 実験手順

実験は防音室内でコンピュータを用いて行った。刺激音はヘッドホン(STAX SR-303)を用いて提示し, 刺激音のレベルは各被験者が聞きやすいレベルに合わせて調節した。実験中は刺激音を一度だけ提示した後に 14 種類の CV を選択肢として示し, 被験者に選択肢の中から一つ選択させた。刺激はランダムに提示し, 以上の試行を 9 セット計 252 刺激に対して実験を行った。

3 結果・考察

表 1 に本研究(diotic 受聴)での各残響条件における処理あり/なしの平均正解率と, 比較のため XEBEC ホール(dichotic 受聴)で行った

*Evaluation of two steady-state processing methods for improving speech intelligibility in reverberant environments by NAKATA, Yuki, MURAKAMI, Yoshiaki, HAYASHI, Nahoko, MIYAUCHI, Yusuke, HODOSHIMA, Nao, ARAI, Takayuki (Sophia Univ., Japan) and KURISU, Kiyohiro (TOA Corp., Japan).

表 1. 実験結果の平均正解率(%)

| 残響時間 (s) | 受聴 | 処理 | 定常部処理 | | | | |
|----------|----|----|-------|------|-------|------|---|
| | | | 抑圧 | 零挿入 | | | |
| | | | | 50ms | 100ms | | |
| 0.7 | ○ | なし | 86.3 | * | / | | |
| | | あり | 92.6 | | | | |
| 0.9 | ○ | なし | 55.8 | * | | | |
| | | あり | 66.2 | | | | |
| 1.1 | ○ | なし | 36.5 | * | | | |
| | | あり | 41.5 | | | | |
| 2.9 | ○ | なし | 29.1 | 31.9 | * | 36.3 | * |
| | | あり | 31.6 | 36.8 | | 44.2 | |
| | ● | なし | 45.6 | 44.2 | * | 49.3 | * |
| | | あり | 42.6 | 50.7 | | 57.8 | |
| 3.3 | ○ | なし | 24.2 | 23.6 | | 31.9 | * |
| | | あり | 23.1 | 23.1 | | 36.8 | |
| | ● | なし | 34.1 | 37.6 | * | 37.6 | * |
| | | あり | 32.5 | 42.6 | | 45.2 | |

- : 本研究 (diotic 受聴)
- : 先行研究 (dichotic 受聴)
- * : 有意差あり ($p < 0.05$)

先行研究[4,5,6]の平均正解率を示す。

残響時間 2.9~3.3 s で先行研究 (dichotic 受聴) [4,5]における正解率と比較すると本研究 (diotic 受聴) の正解率は 10 % 前後下がる結果となった。これは[7]の dichotic 受聴に比べて diotic 受聴では音声明瞭度が低下するという報告から予測されるものであり, diotic 受聴と dichotic 受聴を比較した先行研究[8]と同様の傾向を確認できた。

また, 残響時間が短くなるにつれて処理あり/なし共に正解率が上がるという傾向も確認できた。定常部抑圧処理は diotic 環境において残響時間 0.5~1.2 s で効果があるという報告[9]と同様に, 残響時間 0.7~1.1 s で処理ありの正解率が処理なしの正解率を有意に上回り, 明瞭度の改善が得られた。

一方, 残響時間 2.9 s, 3.3 s での定常部抑圧処理の結果については処理あり/処理なしの正解率の差に有意差はみられず, 明瞭度の改善はみられないことが確認された。これは床効果が起こってしまったためだと思われる。

残響時間 2.9 s, 3.3 s での定常部零挿入処理の結果については, 残響時間 3.3 s, $T_2 = 50$ ms のみ有意差がなかったが, 他はすべて処理ありが処理なしの正解率を有意に上回った。このことから比較的長い残響時間においては定

常部零挿入処理は音声明瞭度を改善させる効果があることが確認された。

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と定常部零挿入処理の効果について検討した。今回はヘッドホンを用いた防音室環境での diotic 受聴による実験を行い, 比較的短い残響時間(0.7~1.1 s)では定常部抑圧処理の効果が有意に得られた。また比較的長い残響時間 (2.9 s, 3.3 s) では先行研究と同様に定常部抑圧処理による音声明瞭度の改善は得られなかったが, 定常部零挿入処理による音声明瞭度の改善は一部を除いて確認することができた。また, diotic 受聴では実環境よりも正解率が 10 % 前後下がることも確認できた。

本研究により, diotic 受聴と dichotic 受聴における明瞭度に相関があることが確認された。これより, diotic 受聴で行った先行研究[9]で定常部抑圧処理の効果を得ることができた残響時間 0.5~1.2 s で実験を行えば実環境でも処理の効果が出ることが予測される。また, 定常部零挿入処理については処理による改善が得られる残響条件の範囲を検討したい。

謝辞

本研究は日本学術振興会の科学研究費補助金(A-2, 16203041)の助成を受けて行ったものである。

参考文献

- [1] Nabelek and Robinette, J. Acoust. Soc. Am., 63, pp. 187-194, 1978.
- [2] 荒井隆行 他, 日本音響学会秋季研究発表会講演論文集, 1, pp. 449-450, 2001.
- [3] Arai et al., Acoust. Sci. Tech., 23, pp. 229-232, 2002.
- [4] 林奈帆子 他, 日本音響学会春季研究発表会講演論文集, 1, pp. 537-538, 2005.
- [5] Hayashi et al., Proc. Interspeech, 2005.
- [6] Arai, Acoust. Sci. Tech., 26, 2005.
- [7] Helfer, J. Speech Hear. Res., 37, pp. 429-438, 1994.
- [8] Hodoshima et al., Acoust. Sci. Tech., 26(2), pp. 212-214, 2005
- [9] Hodoshima et al., Acoust. Sci. Tech., 25, pp. 58-60, 2004