

# 残響環境下における音声明瞭度改善のための前処理 - 定常部抑圧処理の抑圧率と残響の関係 - \*

村上 善昭, 程島 奈緒, 中田 有貴, 林 奈帆子, 宮内 裕介, 荒井 隆行(上智大・理工),  
栗栖 清浩 (TOA)

## 1 はじめに

残響が聞こえを悪化させる原因の1つとして, 先行する音声区間に付加された残響が後続する音声に影響を与える overlap-masking がある<sup>[1-2]</sup>。荒井らは任意の室内空間に適切な前処理装置を設けて, 電気音響的に残響による明瞭度低下を防ぐ手法として定常部抑圧処理を提案した<sup>[3-4]</sup>。これは, 音声知覚に重要な遷移部のエネルギーはそのまま残し, 音声知覚に重要な情報が少ないとされる定常部<sup>[5]</sup>のエネルギーを抑圧する処理である。

定常部抑圧処理にはいくつかのパラメータがあり, 特に定常部を抑圧する度合い(抑圧率)は, 残響による overlap-masking 量と関係がある。そのために, 残響時間が長い室では定常部をより強く抑圧することで overlap-masking 量を減らすことが出来ると考えられる。

定常部抑圧処理の効果を最大限に得るためには, 室ごとに定常部抑圧処理のパラメータを最適な値に調整しなければならない。そこで程島らは抑圧率に着目し, 明瞭度と抑圧率の関係を調べるために聴取実験を行った<sup>[6]</sup>。本論文は先行研究<sup>[6]</sup>よりも長い残響時間をもつインパルス応答を用いて, 残響条件と抑圧率の関係を調べるために聴取実験を行った。

## 2 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<sup>[3-4, 6-7]</sup>で用いた処理と基本的に同じである。本処理は音声のスペクトル遷移を表す  $D$ <sup>[5]</sup>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。 $D$  は基本的には Furui らが行った方法<sup>[5]</sup>と同じであるが, 本論文では帯域分割された信号の時間包絡の対数に対する回帰係数を複数帯域に渡って二乗平均したものを

用いた<sup>[3-4]</sup>。定常部とみなした箇所を元の波形の振幅から, 0, 20, 40, 60, 80, 100%に抑圧した。0%とは定常部の振幅を0にすることであり, 100%は振幅に変化を加えないことである。

## 3 実験

残響条件は次の3種類を用いた。まず, 東大和市大ホールで測定されたインパルス応答に時間包絡をかけて, 人工的に2種類のインパルス応答を作成した<sup>[7]</sup>。これらをIR1, IR2とする。IR1は先行研究<sup>[6]</sup>でも用いられた。また, 東京国際フォーラムで測定されたインパルス応答をIR3として用いた。各インパルス応答の残響時間とD値を表1に示す。本論文で用いた残響時間は, 中心周波数500, 1000, 2000 Hz のオクターブバンドのEDT (Early Decay Time) の平均値とした。

被験者は健聴な日本語母語話者22名(男性14名, 女性8名, 年齢18-23才)であった。原音声は先行研究<sup>[6]</sup>と同じものを用い, 刺激は原音声に定常部抑圧処理を施したもの(処理あり)と処理を施していないもの(処理なし)を用意した。全ての音声信号にインパルス応答を畳み込んで残響が付加された音声刺激を作成した。

実験はコンピュータを用いて, 被験者ごとに防音室内で行った。刺激音は, ヘッドホン(STAX SR-303)を用いてdioticに提示された。音圧レベルは, あらかじめ被験者ごとに聞きやすいレベルに調整した。各試行では刺激音を一度提示し, その後PCの画面上に選択肢

表1. 各インパルス応答の残響時間とD値

	IR1	IR2	IR3
残響時間[s]	1.3	1.7	2.1
D値[%]	62	55	37

\* A Preprocessing Technique for Improving Speech Intelligibility in Reverberant Environments: Relationship between Suppression Rate of Steady-state Portions of Speech and Reverberation, by Yoshiaki Murakami, Nao Hodoshima, Yuki Nakata, Nahoko Hayashi, Yusuke Miyauchi, Takayuki Arai (Sophia Univ.) and Kiyohiro Kurisu (TOA).

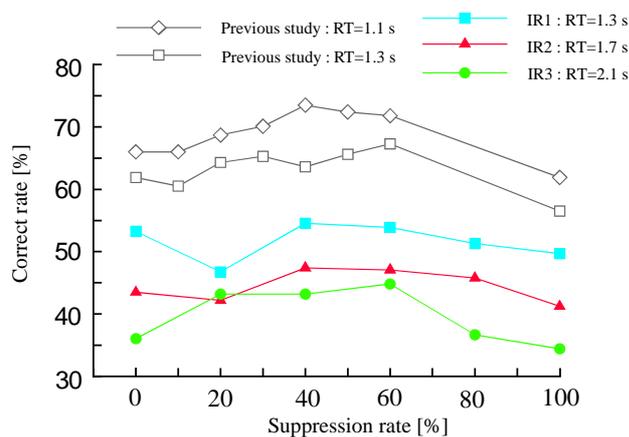


図1. 本実験と先行研究<sup>[6]</sup>の正解率

として実験で使用した14種類の単音節をかなで表示した。被験者には、画面上の選択肢を一つマウスでクリックしてもらい回答させた。各被験者に対して、計252刺激(抑圧率6種類×残響条件3種類×14単音節)をランダムに提示した。

#### 4 結果

残響3種類、処理6種類の計18条件を分析の対象とし、被験者ごとに正解率を求めた。被験者全体の平均値と先行研究<sup>[6]</sup>の結果を図1に示した。

残響・処理条件に対して、繰り返しのある6×3の分散分析を行った結果、残響による主効果が有意( $p < .001$ )であることが示された。よって、残響時間を長くすることにより正解率が減少することが示された。また、処理による主効果は2次回帰曲線で有意であることが示された( $p < .001$ )。下位検定としてSidakの多重比較を行ったところ、IR1とIR2では処理なし音声(抑圧率100%)と処理あり音声の全ての間に有意差は無かった。よって、ある抑圧率で最も定常部抑圧処理の効果があると結論付けられる結果ではなかった。しかし、IR3では処理なし音声の正解率よりも、抑圧率20%と40%の正解率( $p < .006$ )、抑圧率60%の正解率( $p < .012$ )の方が有意に高いことが示された。

#### 5 考察

IR3では抑圧率が20%から60%の範囲のときに正解率が有意に改善し、定常部抑圧処理の効果を確認することができた。しかし、今回用いた3種類の残響条件において、最適な抑圧率を確認することは出来なかったが、抑

圧率が40%から60%の範囲で、処理の効果を最大限に得ることが出来ると推測する。

先行研究<sup>[6]</sup>と同じインパルス応答を用いたIR1は、先行研究<sup>[6]</sup>では処理による効果が見られたが、今回の実験では効果が確認出来なかった。その理由の一つとして、組み合わせた他の残響条件の影響が考えられる。残響時間が長い条件と組み合わせて実験を行うと、正解率が全体的に低下するということが、様々な残響条件を用いた先行研究<sup>[8]</sup>から報告されている。今回の実験では先行研究<sup>[6]</sup>で用いた残響条件よりも残響時間が長いインパルス応答と組み合わせるために、全体的に正解率が下がったと推測できる。

#### 6 おわりに

今後は組み合わせる他の条件により、全体の正解率が変化することに注意を払いながら実験を行わなければならない。また単語や文章を用いて聴取実験を行い、抑圧率と残響条件の関係を調べていきたい。

#### 謝辞

本研究は科学研究費補助金(A-2,16203041)の助成を受けて行った。インパルス応答のデータを提供して頂いた東京大学生産技術研究所(当時)の橋秀樹先生、上野佳奈子さん、横山栄さんに感謝申し上げます。

#### 参考文献

- [1] Nabelek and Pickett, *J. Speech Hear. Res.*, 17, 724-739, 1974.
- [2] Bolt and MacDonald, *J. Acoust. Soc. Am.*, 21(6), 577-580, 1949.
- [3] 荒井他, 日本音響学会秋季研究発表会, 1, 449-450, 2001.
- [4] Arai *et al.*, *Acoust. Sci. Tech.*, 23, 229-232, 2002.
- [5] Furui, *J. Acoust. Soc. Am.*, 80(4), 1016-1025, 1986.
- [6] 程島他, 日本音響学会秋季研究発表会, 607-608, 2005.
- [7] Hodoshima *et al.*, *Forum Acusticum Sevilla*, 2002.
- [8] Hodoshima *et al.*, *J. Acoust. Soc. Am.*, 2006. (to be published)