

モーラ毎に定常部抑圧率を変化させた 信号処理による残響下での音声明瞭度改善

☆柳橋歩, 程島奈緒 (東海大・情報通信), 荒井隆行 (上智大・理工)

1 はじめに

空港や駅などの公共空間で、アナウンス音声などに聞き取りづらさを感じる原因の一つは環境内の残響による **overlap-masking**^[1]の影響であり、残響下での音声明瞭度の向上を目的とした信号処理として、定常部抑圧処理^[2-4]が提案されている。この処理は、音声知覚に比較的重要ではないとされるが、エネルギーの大きい定常部を抑圧し、後続の過渡部への **overlap-masking** を軽減し、音声明瞭度を改善するものである。先行研究^[2-5]では、単音節は明瞭度が改善したが、単語了解度試験では効果が得られていないのが現状である。

本研究は、アナウンス音声の明瞭度改善への前段階として、単語了解度改善を目的とした。文章はモーラの組み合わせにより構成されることから、各モーラに適した抑圧を行うことにより効率的な **overlap/self-masking** の軽減効果が得られると考える。そこで、正解率が後続のモーラに比べ低下する先頭モーラ^[6]はエネルギーを残し、後続のモーラは徐々にエネルギーを削減する定常部抑圧処理を提案する。

2 聴取実験

2.1 原音声

刺激の原音声は、関東方言話者1名（年齢24歳、男性）にターゲット語、キャリア文を別々に防音室内で発声させ収録した。振幅を正規化し、キャリア文にターゲット語を挿入し原音声とした。キャリア文は「このような場合には（ターゲット語）することがあります。」を使用し、ターゲット語へのマスクング量の統制を取るため、直前のキャリア文終了部から150msをターゲット開始点とした。ターゲット語は親密度別単語了解度試験用音声データセット2003 (FW03)^[7]より、単語内に含まれる母音が、3種の母音 (/a,i,u/) で構成され、ターゲット語に含まれる母音の合計が各モーラ位置16個ずつとなる48種の単語(親

密度1.0~2.5の4モーラ語)を選定した。

2.2 信号処理

本研究で提案する定常部抑圧処理は、4モーラ語をターゲットとするため、文章を1フレーム4~5モーラで区切り、先頭モーラから徐々に抑圧率を増加させ、全てのフレームに同様の処理を行う。第1モーラの正解率は後続のモーラに比べ低下する^[6]ことから、先頭モーラはエネルギーを残し、後続モーラは定常部を大きく抑圧することによりエネルギーを削減した。抑圧率の変更は、従来法^[2,3]の処理を応用し、定常部判定された波形の振幅を各モーラ位置により異なる抑圧率で抑圧した。本実験で用いた定常部抑圧処理の各モーラ位置の抑圧率をTable1に示す。

2.3 実験手順

原音声に信号処理を4種付加し、インパルス応答2種（残響時間：2.5 s, 3.5 s）を畳み込んで384種の刺激音を作成した。被験者は聴力に問題がなく、日本語を母語とする若年者24名（平均年齢：21.9歳、男性：21名、女性：3名）とし、3名ずつ計8グループにグループ化した。これは、同一単語・条件の提示がないようにグループ内で刺激をランダム化するために行った。被験者には防音室内でパソコン（hash MiniITX ION Fanless Silent）上のインタフェースを操作させ、オーディオインタフェース（TASCAM US-144MKII）を介して、ヘッドホン（STAX SR-303）から刺激を提示し、聴取したターゲット語を回答用紙に記入させた。

Table 1 : 各モーラ位置の抑圧率

条件	抑圧率 [%]				
	第1モーラ	第2モーラ	第3モーラ	第4モーラ	第5モーラ
1	処理なし				
2	従来法(抑圧率:全モーラ40)				
3	80	60	40	20	10
4	50	40	30	20	10

* The preprocessing approach which changes suppression rates of steady-state portions in each mora for improving speech intelligibility in reverberant environments, by YANAGIBASHI Ayumu, HODOSHIMA Nao (Tokai Univ.) and ARAI Takayuki (Sophia Univ.).

3 結果・考察

3.1 結果

統計解析ソフトウェア (IBM SPSS) を用いて、処理 4 条件と残響 2 条件の平均正解率に対して二元配置の分散分析を行い、処理条件間で Sidak による多重比較を行った。また、各被験者グループの平均正解率に対して Wilcoxon の符号付順位検定を行った。

各条件の全体の平均モーラ正解率を Figure1、平均単語正解率を Figure2 に示す。平均モーラ正解率は、処理条件 ($p < 0.05$) と残響条件 ($p < 0.01$) に主効果が見られた。また、平均単語正解率では、残響条件で主効果が見られた ($p < 0.05$) が、処理条件には主効果が見られなかった。

両結果共に処理と残響の条件間には交互作用は見られなかった。また、条件 2 が条件 4 より有意に正解率が高かった ($p < 0.05$) が、その他の条件間では有意差は見られなかった。さらに、各被験者グループ間でも有意な差は見られなかった。

3.2 考察

本研究で提案した定常部抑圧処理は、有意差は見られなかったが、残響時間 2.5 s の条件で条件 3 の平均モーラ・単語正解率が最も高い結果を得ることができた。有意差を得られなかった理由として、各単語の正解率の差が関係していると考えられる。平均正解率が最も高い単語は 91.7%、最も低い単語では 4.2% となり大きな差が見られた。また、条件 4 は全条件の中で最も正解率が低い結果となり、先行研究^[6]と同様に過度な定常部抑圧は正解率を低下させる結果となった。

4 おわりに

本研究は、単語了解度改善を目的とし、定常部抑圧処理による抑圧効果を各モーラ位置で変化させることにより、先頭音節はエネルギーを残し、後続音節はエネルギーを削減した。聴取実験の結果、モーラ正解率、単語正解率ともに本稿で提案した条件 3 の正解率が高くなったが、従来法と比較して有意差は得られなかった。今後は、本実験の結果を踏まえ、モーラ位置ごとの最適な抑圧率の検討を行い、単語了解度の改善を目指したい。

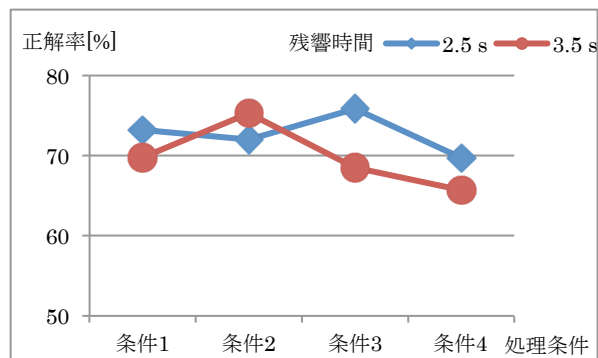


Figure1 : モーラ正解率

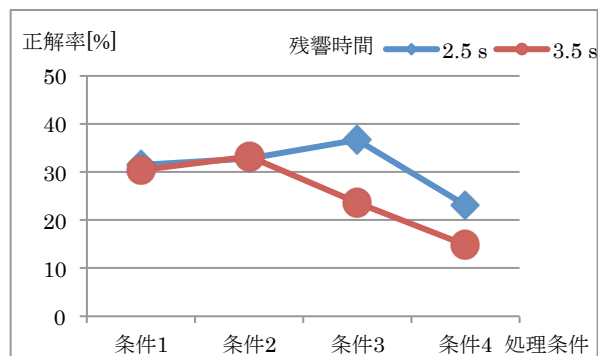


Figure2 : 単語正解率

参考文献

- [1] A. K. Nabelek *et al*, “Reverberant overlap and self-masking in consonant identification,” *J. Acoust. Soc. Am.*, 86, 1259–1265 (1989).
- [2] 荒井隆行 他, “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, 449-450 (2001).
- [3] T. Arai *et al*, “Effects on suppressing steady-state portions of speech on intelligibility in reverberant environment,” *Acoustical Science and Technology*, 23, 229-232 (2002).
- [4] N. Hodoshima *et al*, “Improving syllable identification by a preprocessing method reducing overlap-masking in reverberant environments,” *J. Acoust. Soc. Am.*, 119, 4055–4064 (2006).
- [5] 辻美咲 他, “残響環境における音声明瞭度を目的とした子音強調・母音抑圧による前処理,” 日本音響学会誌, 69(4), 179-183 (2013).
- [6] 坂本修一 他, “単語了解度試験におけるモーラ同定に対する親密度の影響,” 日本音響学会誌, 60(7), 351-357 (2004).
- [7] 天野成昭 他, “新密度別単語了解度試験用音声データセット 2003 (FW03),” 音声資源コンソーシアム (2006).