

残響下での定常部抑圧処理による音声明瞭度改善 -モーラ位置・母音の抑圧率と音声明瞭度の関係-*

☆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は環境内の残響による overlap-masking^[1]の影響により、アナウンス音声に聞き取りづらさを感じる場合がある。この現象は先行する音節の残響が後続の音節に被さることにより発生し、音声明瞭度低下の要因になる。このような残響下での音声明瞭度を向上させるための信号処理として、定常部抑圧処理^[2,3]が提案されている。これは、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することにより後続の過渡部への overlap-masking を軽減する処理である。また、この処理で定常部の振幅を保持する度合い(抑圧率)を40%にし、単音節をターゲット語とした聴取実験では明瞭度改善の効果を得た^[4]が、単語了解度試験では改善を得られていない^[5]。

そこで著者らは、単語の先頭モーラの正解率が後続モーラに比べ低下する^[6]ことに着目し、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た定常部抑圧処理^[7]を提案した。この処理を用いた聴取実験の結果、単語了解度の有意な改善は得られなかったが、残響条件や単語の違いによって単語了解度の改善効果がある適正抑圧率が存在する可能性が示された。

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。先行研究^[7]で得られた各残響条件での適正抑圧率に加え、新たに母音の種類毎に抑圧率を変化させる処理を追加して聴取実験を行った。これは、単語を構成する各母音に対して適した抑圧効果を付加することにより、さらに効率的に overlap-masking を軽減し、単語了解度の改善効果が得られると考えられるためである。これにより、各残響条件の適正抑圧率による有意な改善効果の有無と各母音に対する適正抑圧率の検証を行った。

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キング量の統制を取るため、挿入部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。先行研究^[7]で行った聴取実験では、親密度別単語了解度試験用音声データセット2003(FW03)^[8]より、親密度1.0~2.5の4モーラ語を選定したが、単語の正解率に大きな差があり、結果に影響を与えたことが考えられるため、本実験では4モーラの無意味語をターゲット語として使用した。無意味語の構成は母音3種(/a,i,u)を2種の母音の繰り返し(例:アイアイ、イウイウなど)と同母音の連続(例:アアイウ、イアアウなど)を除外した組み合わせ18種に子音10種(/p,t,k,b,d,g,s,z,m,n/)を単語内で同じものを使用しないようにランダムに組み合わせて144種の単語を作成した。

2.2 信号処理

本実験で用いた定常部抑圧率の各モーラ位置の抑圧率をTable 1に示す。処理番号1は処理を付加せずに原音声を使用し、処理番号2は単音節明瞭度試験^[4]で改善効果の得られた抑圧率40%を使用した。処理番号3・4は先行研究^[6]の結果から得られた各残響条件に対しての適正抑圧率であり、処理番号3は残響時間2.5s、処理番号4は残響時間3.5sでそれぞれ改善効果があるとされたものである。

本研究では新たに母音の種類毎に抑圧率を変化させる処理を提案した。これは、先行研究^[7]で単語正解率やモーラ位置正解率が単語

* The steady-state suppression for improving speech intelligibility in reverberant environment: The relation between speech intelligibility and mora position/ suppression rate of vowel, by YANAGIBASHI Ayumu, HODOSHIMA Nao (Tokai Univ.) and ARAI Takayuki (Sophia Univ.).

Table 1 : Suppression rates (%) in 6 steady-state suppression conditions

Condition	mora					Additional suppression rate
	1	2	3	4	5	
1	Unprocessed					None
2	Conventional processing (all 40)					None
3	80	40	90	30	20	None
4	80	40	40	30	20	None
5	80	40	90	30	20	*1
6	80	40	40	30	20	*2

*1 Suppression rate is decreased by 10% for /a/

*2 Suppression rate is decreased by 20% for /a/ and by 10 % /u/

によって変化することから、各単語を構成する母音にも適正抑圧率が存在し、この値で処理を付加することによって単語理解度の改善効果が得られると考えられるためである。そのため、持続時間が長く、定常部の振幅が大きい母音を含むモーラ位置の抑圧効果を増加させることにより、効率的に overlap-masking を軽減すると考えた。

そこで本実験では、処理番号 3・4 を基に各モーラ位置に母音 2 種 (/a,u/) を含む場合に抑圧効果を増加させる処理条件を 2 種作成した (処理番号 5・6)。これは、先行研究^[7]で母音 (/i/) の持続時間が他の母音 (/a,u/) と比較して短いため、低い抑圧率を設定した場合に十分に抑圧処理が付加できていない可能性があることから、母音 2 種 (/a,u/) に着目した。増加させる抑圧率は、予備実験の結果 (被験者は日本語を母語とする若年者 2 名) から決定した。

2.3 刺激

原音声に処理条件 5 種 (Table 1) を付加し、インパルス応答 2 種 (残響時間 (RT) : 2.5 s、3.5 s) を畳み込んで刺激音を作成した。処理条件 (6 条件) × 残響条件 (2 条件) の計 12 条件である。残響時間は、駅のコンコースと空港のエントランスを想定した。音場シミュレーションソフトウェア (CATT-Acoustic) により残響時間 3.5 s のインパルス応答を作成し、さらに時間包絡の時定数を調整することで 2.5 s の残響時間を持つインパルス応答を作成した。

2.4 被験者

被験者はアンケートにより聞こえに問題がないと回答した、日本語を母語とする若年者 24 名 (平均年齢 : 22 歳、男性 : 19 名、女性 : 5 名) である。

2.5 実験手順

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) から刺激を提示した。条件と単語の組み合わせはカウンタバランスを取り、144 刺激 (各条件 12 音声 × 12 条件) の提示順をランダムにし、各刺激を一度だけ提示した。被験者には、聴取した刺激のターゲット語を回答用紙に記入させた。

3 結果・考察

3.1 単語正解率・モーラ正解率

各条件の平均単語正解率を Figure 1、平均モーラ正解率を Figure 2 に示す。統計解析ソフトウェア (IBM SPSS) を用いて、処理 6 条件と残響 2 条件の平均単語・モーラ正解率に対して二元配置の分散分析を行った。その結果、平均単語・モーラ正解率の両結果共に残響条件 ($p < 0.01$) の主効果は有意であったが、処理条件の主効果は有意ではなかった。さらに、処理条件と残響条件間の交互作用も有意ではなく、処理条件間で Sidak による多重比較も有意差は見られなかった。

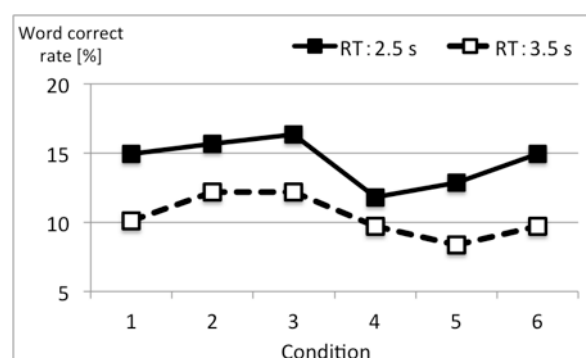


Figure 1 : Word correct rate in each condition

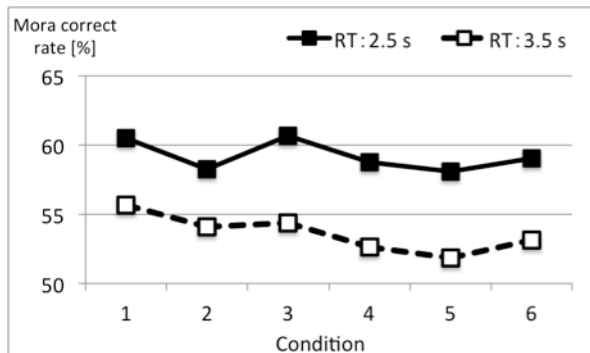


Figure 2 : Mora correct rate in each condition

3.2 モーラ位置正解率

残響時間 2.5 s のモーラ位置正解率を Figure 3、残響時間 3.5 s のモーラ位置正解率を Figure 4 に示す。3.1 章と同様に処理 6 条件と残響 2 条件の各モーラ位置の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。第 1 モーラと第 4 モーラのモーラ位置正解率では残響条件 ($p < 0.01$) の主効果が有意であったが、第 2・3 モーラでは残響条件の主効果は有意でなかった。また、全てのモーラ位置において処理条件の主効果が有意でなく、処理条件と残響条件間の交互作用も有意でなかった。

また、処理条件 6 条件とモーラ位置条件 4 条件の正解率に対して二元配置の分散分析を各残響条件下で行い、モーラ位置条件間で Sidak による多重比較を行った。残響時間 2.5 s では、モーラ位置条件 ($p < 0.01$) の主効果は有意であったが、処理条件の主効果は有意ではなかった。さらに、処理条件とモーラ位置条件間の交互作用も有意ではなかった。また、モーラ位置条件間での Sidak による多重比較では、全てのモーラ位置条件 ($p < 0.05$) が有意となった。残響時間 3.5 s でも、モーラ位置条件 ($p < 0.01$) の主効果は有意であったが、処理条件の主効果は有意ではなく、処理条件

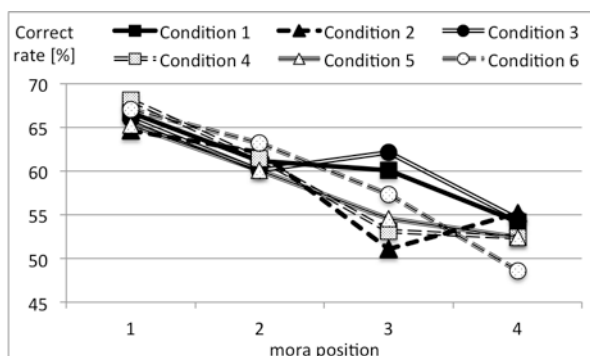


Figure 3 : Mora correct rate at each mora position (RT : 2.5 s)

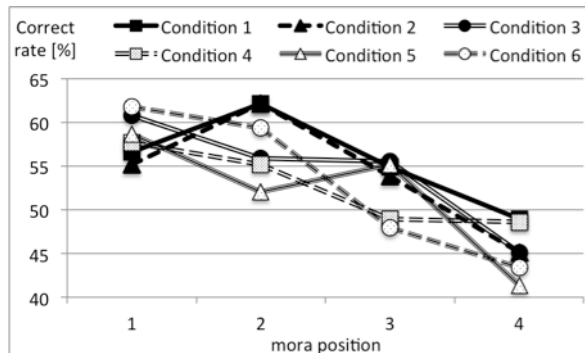


Figure 4 : Mora correct rate at each mora position (RT : 3.5 s)

とモーラ位置条件間の交互作用も有意ではなかった。また、モーラ位置条件間での Sidak による多重比較では、第 1 モーラと第 2・3 モーラ間には有意差が見られなかったが、その他のモーラ位置 ($p < 0.05$) は有意差が見られた。以上の結果より、各モーラ位置に着目した場合でも処理による音声明瞭度改善は得られなかった。

3.3 考察

単語正解率とモーラ正解率 (3.1 章) の結果より、本研究で提案した処理手法による単語了解度の改善は得られなかった。先行研究^[7]で得られた各残響条件での適正抑圧率である処理番号 3・4 (Table 1) は、各適正残響下で平均単語 (Figure 1)・モーラ (Figure 2) 正解率より、残響条件に完全に適応した抑圧率ではないことが示された。この原因として、低親密度の有意語と無意味語とでは使用単語が異なることから、先行研究^[7]で得られた適正抑圧率は、特定の条件下での値であったことが考えられる。しかし、処理番号 3 は残響時間 2.5 s の残響条件下での平均単語 (Figure 1)・モーラ (Figure 2) 正解率の両結果が全ての処理条件間で最も高くなった。これは、第 2 モーラの抑圧率を 40% に設定して overlap-masking を軽減し、第 3 モーラの抑圧率を 90% と高く設定することにより、第 3 モーラの正解率が全処理条件の中で最も高くなったことが影響したと考えられる (Figure 3)。

また、本研究で提案した母音の種類によって抑圧率を増加させる処理番号 5・6 (Table 1) は、平均単語 (Figure 1)・モーラ (Figure 2) 正解率において有意な改善は得られなかった。これは、残響時間 2.5 s (Figure 3) と残響時間 3.5 s (Figure 4) の第 3・4 モーラでの正解率低下が影響を与えている。その原因として先

行研究^[7]の処理を基に抑圧率を増加させているため、後続モーラである第3・4モーラの抑圧率が過度に低くなったことによる音声の劣化^[5]が挙げられる。しかし、残響時間 2.5 s (Figure 3) と残響時間 3.5 s (Figure 4) の第1・2モーラでは、他の処理条件と有意差がないことから、過度な抑圧効果にならない範囲で母音への追加処理を付加することにより、第3・4モーラへの *overlap-masking* を効率的に軽減する可能性がある。

モーラ位置正解率 (3.2 章) より、第1・4モーラは残響の影響により残響時間 3.5 s の正解率が残響時間 2.5 s より有意に低下したが第2・3モーラは残響条件間には有意差が見られなかった。第1・4モーラが残響時間 3.5 s で正解率が低下したのは、先行するキャリア文又は、ターゲット語内の先行する3モーラからの残響の影響をそれぞれ受けたと考えられる。また、第2・3モーラはキャリア文からの残響の影響が少なく、ターゲット語内で先行するモーラも少ないため残響の影響が少なかったのではないかと考えられる^[1]。

また、各正解率に単語の正解率が少なからず影響を与えていると考えられる。本実験ではランダムに組み合わせて作成した無意味語を使用した。正解率が0%の単語が56種あり、先行研究^[7]の結果と比較しても正解率が大きく低下した。これは、本実験で使用した無意味語は先行モーラから後続モーラの予測が困難であり、人間が日常生活で単語を理解するために用いている心的辞書^[6]による予測を使用できないためであると考えられる。

4 おわりに

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。先行研究^[7]で得られた各残響条件での適正抑圧率に加え、新たに母音の種類毎に抑圧率を変化させる処理を用いて聴取実験を行った。

各残響下での適正抑圧率を用いた定常部抑圧処理による単語了解度の改善は得られなかった。しかし、残響時間 2.5 s の適正抑圧率である処理番号3において、他の処理条件では正解率が低下した第3モーラで正解率が最も高くなり、単語了解度改善の可能性を示した。

また、提案手法である母音毎に抑圧率を変化させる処理においても単語了解度の改善を

得られなかった。改善策としては、モーラ位置毎に抑圧率を変化させる処理と組み合わせる場合には、抑圧率があらかじめ低い第4モーラは母音の種類による追加抑圧は付加せずエネルギーを確保し、正解率が他の処理条件と差がない第1・2モーラは追加抑圧することにより、効率的に *overlap-masking* を軽減し、単語了解度の改善に繋がる可能性がある。

今後は、本研究で得られた単語・モーラ・モーラ位置正解率で子音の異聴を検証し、モーラ位置と母音、残響条件だけではなく、子音などの他の要因による定常部抑圧処理の効果の差を検証する。この結果より、定常部抑圧処理による単語了解度の改善を目指したい。

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] 柳橋歩 他, “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,” 日本音響学会秋季研究発表会講演論文集, 357-358 (2013).
- [8] 天野成昭 他, “親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).