

高齢者を対象にした定常部抑圧による単音節強調処理の検討

小林 敬[†] 八田ゆかり[†] 安 啓一[†] 程島奈緒[†] 荒井隆行[†] 進藤美津子[‡]

[†] 上智大学理工学部電気電子工学科 [‡] 上智大学言語障害研究センター

〒102-8554 東京都千代田区紀尾井町 7-1

E-mail: [†] kei-koba@ba2.so-net.ne.jp

あらまし 一般的に高齢者は、加齢による聴覚機能の低下が見られることが言われている。その一つの現象として高音域の聴力の低下や継時マスクングの悪化があげられ、音声の子音の聞き取りが難しくなるとされる。荒井ら(2002)の残響環境下における音声の明瞭度改善の研究報告では、エネルギーは比較的大きいが音声知覚にはそれほど重要ではないとされる定常部を抑圧する処理が提案された。本研究ではこの定常部抑圧処理を使って、24 個の CV (子音 母音) 単音節の子音強調を試みたので報告する。50 人の高齢者を対象に明瞭度試験を行った結果、本処理が有効であることが確認できた。特に難聴の高齢者に対して、顕著な効果がみられた。

キーワード 子音強調、聴覚障害、定常部抑圧、高齢者、補聴器

A study of monosyllable enhancement for elderly listeners by steady-state suppression

Kei KOBAYASHI[†] Yukari HATTA[†] Keiichi YASU[†] Nao HODOSHIMA[†] Takayuki ARAI[†]
and Mitsuko SHINDO[‡]

[†] Department of Electrical and Electronics Eng., Sophia University, [‡] Research Center for Communication Disorders,
Sophia University, 7-1 Kioi-cho, Chiyoda-ku, Tokyo, 102-8554 Japan

E-mail: [†] kei-koba@ba2.so-net.ne.jp

Abstract The hearing function of elderly people is generally degraded according to their aging process. As one phenomenon, the drop of hearing level in high frequencies and the abnormal growth of the temporal masking are found. In a previous study, Arai et al.(2002) suggested the steady-state suppression of the speech which improves speech intelligibility in reverberant environments. Steady-state portions are defined as those that have more energy, but are less crucial for speech perception. In this paper, we report the consonant enhancement of 24 consonant-vowel (CV) monosyllables by using the steady-state suppression. And we certified that the technique was effective to improve speech intelligibility according to the results of 50 elderly listeners. Especially, more improvements are detected for hearing-impaired subjects.

Keyword Consonant enhancement, hearing-impaired, steady-state suppression, elderly listener, hearing aid

1. はじめに

一般的に高齢者は、加齢による聴覚機能の低下がみられることが言われている。まず、左右対称に高音域の聴力が低下し、次第に中・低音域の聴力も低下していくが、常に高音漸傾型を示すのが特徴である[1]。語音は、様々な周波数成分の混ざった複雑な構造を持つ複合音であり、特に子音は雑音成分を多く含んだものがある[2]。従ってもし高周波数帯域に対する聴取能力が落ちていけば、子音をうまく聴取できないことが容易に想像できる。加齢と異聴の関係については、設楽

ら[3]の研究があり、平均して 60 歳前後から異聴が多くなるが、これはちょうど 4kHz と 8kHz の聴力が急に低下していく時期と一致していると言う。

以前より、高音域の聴力を補償することを目的とする音声の子音強調処理の研究が行われてきた。Gordon-Salant[4]や Kennedy et al.[5]は、単音節(CV)の振幅のレベル比を変更して子音強調した処理が高齢者や聴覚障害者に対して有効であったと報告している。

また、加齢によって継時マスクング量が増加するという報告もある。Gehr and Sommers[6]は、10ms, 500Hz

の信号に対して 50ms の広帯域マスキングで後向きマスキングを与える実験を若者と高齢者で比較して行った結果、高齢者のほうが多くマスキングの影響を受け、さらにその影響は信号とマスキングが 20ms 離れていてもあったと言う。鈴木・吉住ら[7]は、継時マスキングをレベルの時間変化に対する側抑制的な振る舞いと仮定して極性反転した側抑制機構を用いてレベルを制御する子音強調処理を施す処理を提案し、24~38歳の健聴者に対してノイズで語音明瞭度を下げ擬似難聴にした状態の実験で効果があったと報告した。

音声の知覚において、古井[8]は単音節明瞭度実験の結果から音声の遷移部は非常に重要な役割を果たしていると報告した。また音声の定常部は遷移部と比較するとそれほど重要ではないという報告もある[9]。荒井ら[10]・程島ら[11,12]は、エネルギーは比較的大きいが音声の知覚にはそれほど重要ではない定常部を抑圧する定常部抑圧処理を提案して、残響環境下における音声の明瞭度の改善に効果があったことを確認した。また、荒井ら[13]は、定常部抑圧処理が補聴器の明瞭度の改善に使用できることを示唆した。

本研究の目的は、上記の荒井らの定常部抑圧処理を用いて子音強調した CV が、高齢者の明瞭度の改善に効果があるかどうかを調べることである。

2. 定常部抑圧と子音強調

荒井ら[10]・程島ら[11,12]によって提案されている定常部抑圧処理を用いて、子音強調を行う。まず、16kHz でサンプリングした入力音声を、聴覚特性である臨界帯域を模擬して 1/3 オクターブ幅で帯域分割する。続いて各帯域において、平滑化及びダウンサンプリングした時間包絡の対数をとってから回帰係数を求める。さらに回帰係数の 2 乗平均 (=D) を 10ms ごとに求め、最後にその時間系列を元の標本化周波数に戻す。D は、古井にならない音声のスペクトル遷移を表すパラメータを表す [8]。こうして得られた D が一定以下の場合、入力音声の振幅の 40% に抑圧した振幅を出力とする。抑圧の際、急激に振幅が変化しないように台形補正を加えている。さらに、この定常部抑圧された単音節と原音声の単音節の実効値を揃えた。以上の処理によって得られた CV の例を図 1 に示す。特に子音の直後の母音の定常部が抑圧されている様子が分かる。また子音区間やその後のフォルマント遷移区間が強調されている様子が分かる。なお、母音の語尾(わたり)は続く子音の知覚のために重要な場合があるので、ここでは抑圧せずに残してある。この語尾の振幅は台形補正時の側辺の傾きを変更することによって調整可能である。

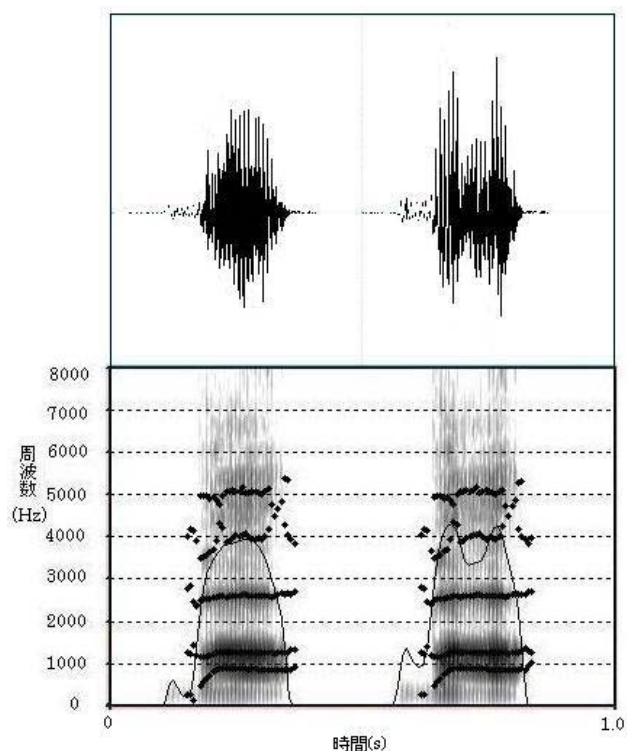


図 1 定常部抑圧処理によって子音強調を施した単音節 /ba/。上図は時間波形、下図はスペクトログラムを示す。左側が原音声、右側が子音強調後の音声である。下図の点はフォルマントを示し、細線はインテンシティを示す。

3. 聴取実験

3.1. 刺激

刺激は日本語の CV で、C として、/p/, /t/, /k/, /b/, /d/, /g/, /s/, /ʃ/, /h/, /tʃ/, /dz/, /dʒ/, /m/, /n/ を、V として、/a/, /i/ を用いた。実験で使用した 24 種類の CV を表 1 に示す。各刺激は、ATR 研究用日本語音声データベース(話者: MAU, 40 才男性)を使用し、同じ母音の CV セットごとに実行値を正規化した。それぞれ原音声(処理なし)と子音強調処理を施した音声(処理あり)の 2 種類を用意した。刺激には残響成分やキャリアセンテンスは

表 1 実験で使用した単音節(CV)の分類

	Voiceless C + Vowel	Voiced C + Vowel
Stop C + Vowel	/pa/ /ta/ /ka/ /pi/ /ki/	/ba/ /da/ /ga/ /bi/ /gi/
Fricative C + Vowel	/sa/ /ʃa/ /ha/ /ʃi/ /hi/	
Affricate C + Vowel	/tʃa/ /tʃi/	/dʒa/ /dʒa/ /dʒi/
Nasal C + Vowel		/ma/ /na/ /mi/ /ni/

含まれない。

3.2. 被験者

被験者は、シルバー人材センターから派遣された東京都千代田区に在住する 56 歳から 90 歳までの日本人の男女 50 名である。全員、痴呆等の症状はなく、また補聴器は未装用である。被験者の平均年齢は 67.9 歳、平均聴力レベルは 21.2 dBHL (3 分法で測定) である。

オーディオグラムにより、健聴者 (グループ A)、高音漸傾型の難聴者 (グループ B)、その他の難聴を呈する難聴者 (グループ C) に分類して、検討を試みた。ここで、健聴者とは 3 分法で 30dBHL より良い聴力がかつ他のグループのような劣化が見られない者とした。グループ C のその他の難聴者とは、低音域の聴力の劣化、特定の周波数のみの劣化、最小可聴値レベルの上昇、左右差あり、など高音漸傾型以外の難聴を呈している者とした。今回は主に高音域の聴力の低下を補うことを目的とする子音強調処理であったため、高音漸傾型とは別のグループに分けた。各グループの人数、平均年齢、平均聴力レベルを表 2 に、オーディオグラム (良聴側) を図 2.1、2.2、2.3 に示す。

表 2 被験者グループ

グループ	人数 (人)	平均年齢 (歳)	平均聴力レベル (dBHL)
A	22	65.4	17.3
B	16	70.4	25.2
C	12	69.0	22.9

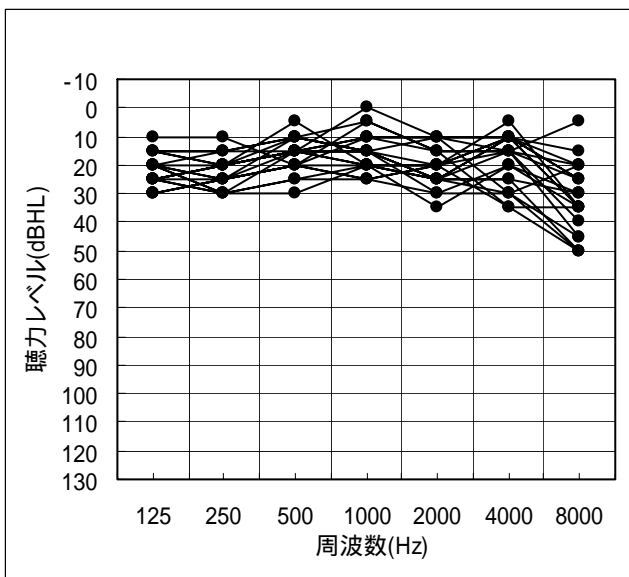


図 2.1 グループ A (健聴者) のオーディオグラム

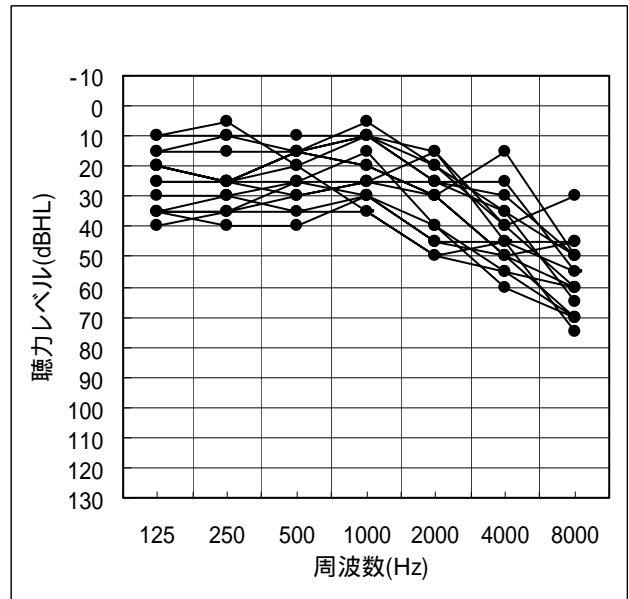


図 2.2 グループ B (高音漸傾型難聴者) のオーディオグラム

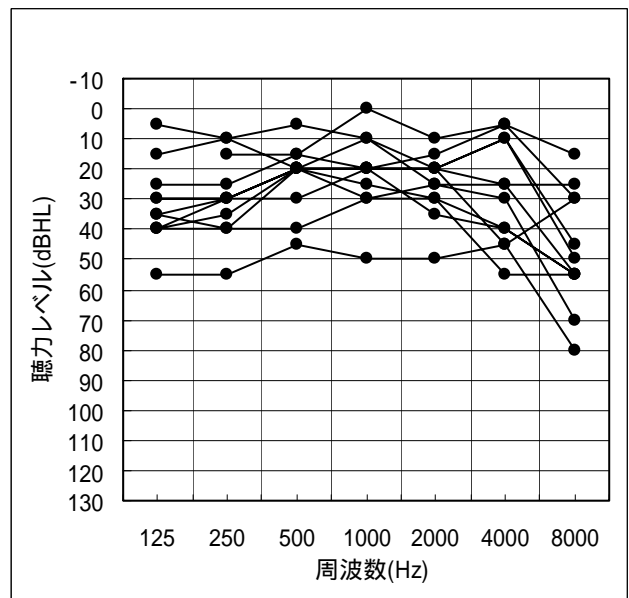


図 2.3 グループ C (その他の難聴者) のオーディオグラム

3.3. 手順

実験の指示は防音室内のコンピュータの画面上で行った。刺激音の呈示はヘッドホン (STAX SR-303) を使い、あらかじめ被験者ごとに適した音圧レベルに調整した。実験の形式に慣れさせるため、本番前に被験者に対して 1~2 回の試行練習を行った。実験の手順は、各試行においてまず刺激音を一度だけ呈示し、呈示終了後回答用紙に聞こえたと思う単音節を記入させた。記入が終わると画面をクリックさせて、次の刺激

が呈示されるようにした。各被験者に対して、計 48 個の刺激（処理なし 24 種類、処理あり 24 種類）をランダムに並べて呈示した。

4. 実験結果

被験者 50 名全体について、処理なしの正解率 86.6% (標準偏差 11.1%)、処理ありの正解率 89.2% (標準偏差 10.2%) であった。*t* 検定の結果、処理なしの正解率と処理ありの正解率の間には 1%水準で有意差が認められた ($p=0.002$)。年齢と聴力レベルの相関係数は 1%水準で有意 (両側) であった。年齢及び聴力レベルと、処理なしの正解率、処理ありの正解率との相関係数もそれぞれ 1%水準で有意 (両側) であった。各グループにおける正解率を表 3 に示す。さらに、各グループで子音の種類ごとに分類した正解率を図 3 に示す。

表 3 実験結果

グループ	処理なし	処理あり	有意確率
A	91.4%	92.6%	0.267
B	79.9%	83.1%	0.054
C	86.1%	91.0%	0.032
B+C	82.6%	86.5%	0.003

5. 考察

グループ A 及びグループ B は、処理なしの正解率と処理ありの正解率の間に有意差が認められなかったが、グループ C において 5%水準で有意差が認められた。グループ B においては後述するように無声摩擦音の処理を今後改善することで、5%水準の有意差が得られると思われる。また、グループ B+C (難聴者全体) では 1%水準で有意差が得られた。

以上の結果から、高齢の難聴者に対して定常部抑圧による子音強調処理の効果が明らかと言える。この処理は、子音部及び母音への遷移部を強調するようにその直後 (10ms 以内) の母音の定常部を原音の 40% の振幅比に抑圧しているため、子音と母音のエネルギー比が変わり、さらに単音節全体の RMS を揃えたことによって、結果的に子音のレベルが原音声 (処理なし) と比べて大きくなるので子音強調の効果がある。従って、特に高音域の聴力が低下している難聴者にとって子音が聞きやすくなると思われる。Gordon-Salant[4]は、子音と母音のエネルギー (RMS) の比が 10dB の条件で子音強調した CV が、高齢者の明瞭度の改善に対して有効であったと示唆しているが、本実験でも同様の効果が得られた。

今回の実験ではグループ C において、もっとも子音

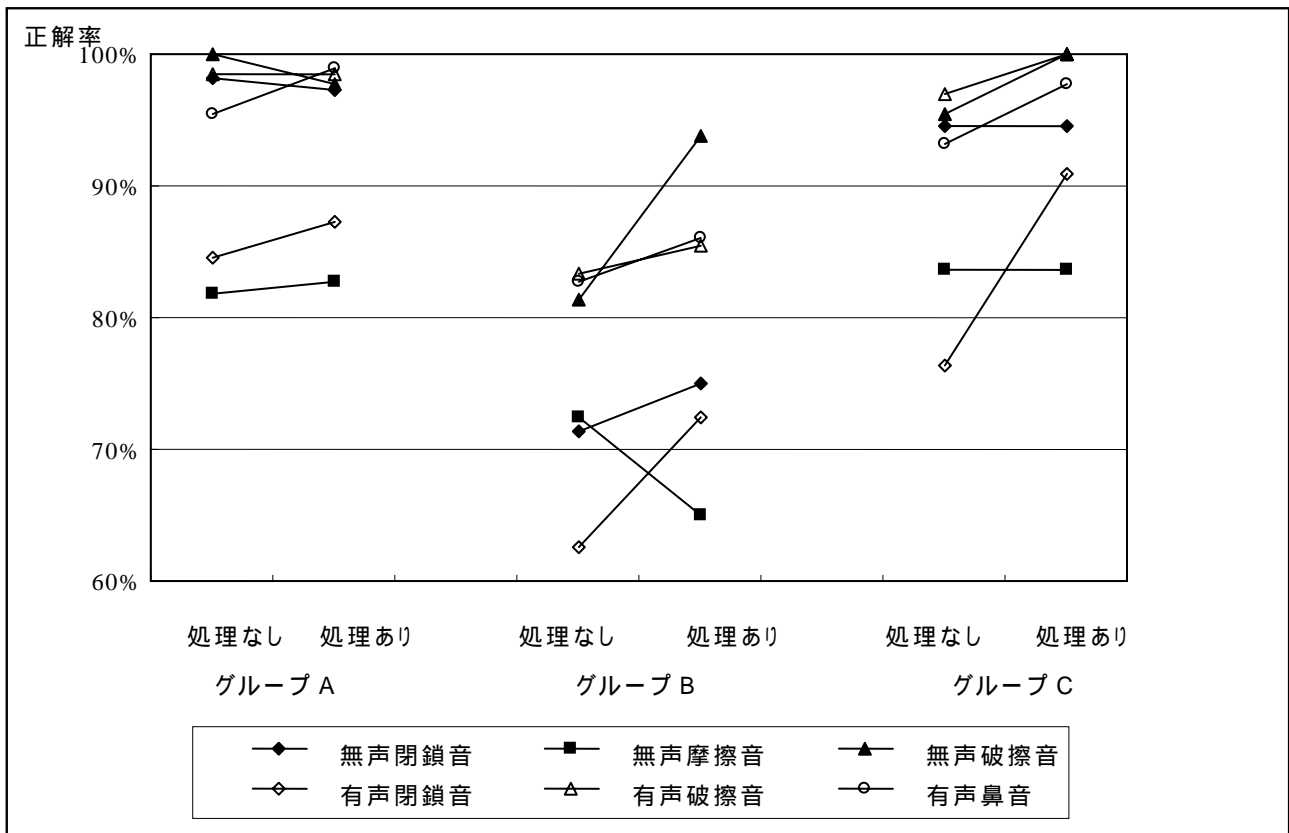


図 3 子音分類ごとの正解率

強調による明瞭度改善の効果がみられた。グループ C は、低音障害型の人が 1 名いる他は全員、高音域の聴力も落ちているので、子音強調の効果が同様にあったと考えられる。またグループ C の処理なしの正解率が全体的にグループ B のそれよりも高かったのは、左右差がある人が多く良耳側で聞いていたり、低音域の聴力レベルが下がっている被験者が多いためにともとも母音による聞き取りにくさの影響を受けにくかったりしていたのではないかと考えられる。また、あらかじめ被験者に適した音圧レベルに調整したため、同様の平均聴力レベルであるグループ B と比較して特に高音域で十分な音圧が確保できていたのではないかと考えられる。

無声摩擦音について、特にグループ B の正解率が下がる結果が見られた。これは、無声摩擦音の子音成分の帯域の一部の時間包絡に定常部分があり、今回これを定常部とみなして抑圧する場合があったためと考えられる。他のグループと比べてグループ B の正解率の低下が顕著だったのは、グループ B が高音域漸傾型のため、抑圧された子音区間が聞きにくかったために異聴しやすかったと考えられる。これは荒井らの処理に工夫を加えることによって今後、正解率の改善が期待できる。また、グループ A における子音分類ごとの正解率を比較すると、無声摩擦音（処理なし）の正解率が低い結果が得られている。特に /ʃi/ の正解率が 40.9% であり、異聴した者は全員が /tʃi/ と回答していた。音節の頭の位置では、破擦音と摩擦音を区別する主な音響的キューは雑音エネルギーの立ち上がり時間と摩擦の持続時間であるとされる[2, 14]。高齢者はこの音響的キューが判別しにくいことが示唆される結果となった。

有声閉鎖音は、どのグループにおいても明瞭度の改善が見られた。特にグループ B, C において改善が顕著であった。さらに詳しく調査したところ、処理なしにおいて /bi/, /da/ をそれぞれ、/ri/, /ra/ に異聴する場面が目立っていた。これは進藤[15]の研究でも、子音を /r/ 音へ異聴する場面が多いことが指摘されている。この /bi/, /da/ の、処理ありの明瞭度はどのグループにおいても改善された。この結果が、有声閉鎖音の明瞭度の改善に結びついている。閉鎖音は音節の頭に現れる閉鎖区間及びその直後のフォルマント遷移区間が極めて短い特徴を持つが[2]、この区間が本処理によって強調された効果があったと考えられる。なお、/gi/ について注釈しておきたい。/gi/ は、実際に呈示していたのは鼻濁音の /ŋi/ であった。このために、健聴者 22 名中 6 名が /ri/ と異聴していたようである。これに関しては処理なしと処理ありで明瞭度に差は見られなかった。

無声破擦音について、グループ B 及びグループ C における明瞭度の改善が顕著に見られた。処理なしでは他の無声音への異聴が目立っていたのが、処理ありでは正確に知覚できるようになっている。破擦音は、閉鎖音のように声道の完全な阻害を伴って生成されかつ摩擦区間を持つが摩擦音の摩擦区間よりも短い傾向にある[2]。閉鎖音同様、この比較的短めの雑音エネルギーを持つ子音区間のレベルが本処理によって強調された効果があったと考えられる。

母音の正解率は、グループ B に属する 2 名を除いて、処理なし処理ありとも全員 100% であった。この 2 名のうち 1 名は処理なしの /ga/, /hi/, /ʃa/ を /bi/, /ma/, /tʃi/ に、もう一名は処理ありの /hi/, /ha/, /bi/, /dʒa/ を /na/, /dʒi/, /ga/, /ka/ に異聴していた。この 2 名のオーゾグラムは、前者は 250Hz から高音域にかけて高音漸傾型を示しており、後者は 1000Hz から高音域にかけてやや急な高音漸傾型を示していた。従って主に第 2 フォルマントの識別が困難であったために母音の異聴を起こしたのではないかと考えている。しかし加齢と語音弁別能力の低下については、コルチ器の感覚細胞とラセン神経節の変形によるところが大きいという報告もあり[1]、純音聴力だけで説明するのは難しい。

本処理の改善が見られたもう一つの理由として、後向性マスキング量の改善効果も考えられよう。Cobbら[16]は、純音と雑音の間隔を 2~30ms に変化させて後向性マスキングを調査した結果、加齢とともに閾値は上昇し、30 歳と 80 歳の間には 30dB の差異が見られた、と報告している。本処理は各帯域の時間包絡のスペクトル遷移を見ており遷移部の直後の母音部（定常部）を検出して抑圧することが出来るので、母音による直前の遷移部及び子音への後向性マスキング量を低減させることが可能である。今回は、設備等の都合で、後向性マスキング量の測定及び正解率との関連の確認は出来なかった。

6. まとめ

本研究では、荒井らによって提案された定常部抑圧手法[10-12]を用い、音声の知覚に比較的重要ではないとされる母音の定常部を抑圧した子音強調処理を施した単音節(CV)による高齢者の明瞭度の改善を調べた。50 名について明瞭度試験を行った結果、本処理が明瞭度の改善に有効であることが確認できた。特に難聴者において、本処理が有効であることが確認できた。子音分類ごとでは、有声閉鎖音と無声破擦音において本処理の効果が顕著であった。また、無声摩擦音に関しては、今回効果が確認できなかったが今後改善できる余地がある。今後は、VCV(母音 子音 母音)の音節においても、同様の実験を行う予定である。

7. 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。また実験にご協力頂いたシルバー人材センター及び被験者の皆様に感謝いたします。

文 献

- [1] 八木昌人, 加我君孝, 中枢性聴覚障害の基礎と臨床, 金原出版, 東京, 2000.
- [2] レイ・D・ケント, チャールズ・リード, 荒井隆行/菅原勉(監訳), 音声の音響分析, 海文堂出版, 東京, 1996.
- [3] 設楽哲也, 耳鼻咽喉科領域における年齢変化, 世紀社, 東京, 1980.
- [4] S. Gordon-Salant, "Recognition of natural and time/intensity altered CVs by young and elderly subjects with normal hearing," *J. Acoust. Soc. Am.*, 80(6), pp. 1599-1607, 1986.
- [5] E. Kennedy, H. Levitt, A. C. Neuman, and M. Weiss, "Consonant-vowel intensity ratios for maximizing consonant recognition by hearing-impaired listeners," *J. Acoust. Soc. Am.*, 103(2), pp. 1098-1114, 1998.
- [6] S. E. Gehr and M. S. Sommers, "Age differences in backward masking," *J. Acoust. Soc. Am.*, 106(5), pp. 2793-2799, 1999.
- [7] 鈴木良二, 吉住嘉之, 目片強司, 山田義則, 田中豊, 河野淳, 船坂宗太郎, "継時マスキングを補償する音声強調方法の検討," *Audiology Japan* 34(5), pp. 335-336, 1991.
- [8] S. Furui, "On the role of spectral transition for speech perception," *J. Acoust. Soc. Am.*, 80(4), pp. 1016-1025, 1986.
- [9] H. Hermansky and N. Morgan, "RASTA processing of speech," *IEEE Trans. Speech and Audio Process.*, 2, pp. 578-589, 1999.
- [10] T. Arai, K. Kinoshita, N. Hodoshima, A. Kusumoto and T. Kitamura, "Effects on suppressing steady-state portions of speech on intelligibility in reverberant environments," *Acoust. Sci. and Tech.*, 23(4), pp. 229-232, 2002.
- [11] N. Hodoshima, T. Inoue, T. Arai, A. Kusumoto and K. Kinoshita, "Suppressing steady-state portions of speech for improving intelligibility in various reverberant environments," *Acoust. Sci. and Tech.*, 25(1), pp. 58-60, 2004.
- [12] N. Hodoshima, T. Arai, T. Inoue, K. Kinoshita and A. Kusumoto, "Improving speech intelligibility by steady-state suppression as pre-processing in small to medium sized halls," *Proc. Eurospeech*, pp. 1365-1368, 2003.
- [13] T. Arai, K. Yasu and N. Hodoshima, "Effective speech processing for various impaired listeners," *Proceedings the 18th International Congress on Acoustics*, 2, pp. 1389-1392, 2004.
- [14] P. Howell, S. Rosen, "Production and perception of rise time in the voiceless affricate/fricative distinction," *J. Acoust. Soc. Am.*, 73, pp. 976-984, 1983.
- [15] 進藤美津子, "中枢聴覚伝導路障害の局在部位と言語音の認知・理解に関する研究," *帝京医学雑誌*, 10(4), 1987.
- [16] F. E. Cobb, G. P. Jacobson, G. W. Newman, L. W. Kretschmer, K. A. Donnelly, "Age-associated degeneration of backward masking task performance: Evidence of declining temporal resolution abilities in normal listeners," *Audiology*, 32, pp. 260-271, 1993.