

母音知覚におけるホルマントピーク制御の効果と 振幅圧縮処理への応用

板橋貴史(久留米高専専攻科) 坂田聡・上田裕市(熊本大) 渡邊亮(熊本県立技術短大)
池田隆(久留米高専)

1. はじめに

音声には音源や声道の共振特性をはじめとして様々な情報が含まれる。聴き手は音声から言葉はもちろん性別、年齢などの個人性や感情も知ることができる。その中で特に母音の知覚に関連の深いパラメータを探ることは、情報伝達の質を確保するために重要なことである。とくに聴取能力の低下により音声の聴き取りに支障のある場合の補償では、効果的な信号処理方式を実現することが望まれる。本稿ではホルマントの共振のピークレベル制御に着目し、母音を知覚するため合成音声の F2 ピークレベルを変化して健聴者により聴取実験を行い検討する。これらの結果をもとに補聴器における振幅圧縮処理で圧縮係数の算定に参考にできるような知見を得ることを目的とする。

2. 合成音声の生成

ホルマントピークレベルの制御を容易にするため、インパルス音源を用いて並列に配置した共振回路を駆動する図 1 の方式を用いる。

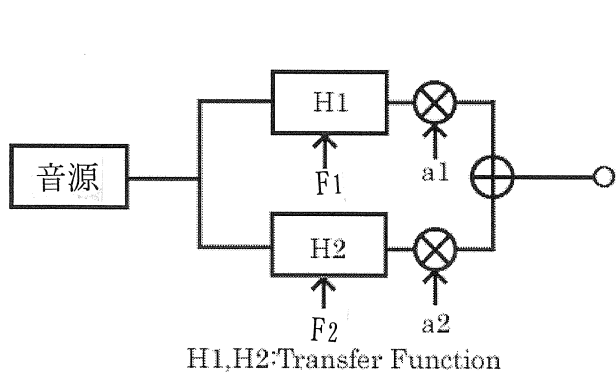


図 1. 合成音声の生成

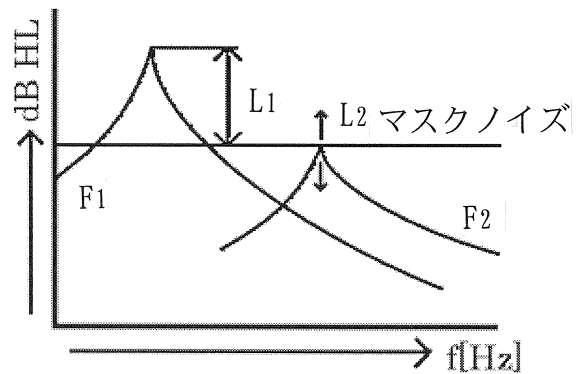


図 2. ホルマントピークレベル

インパルスを第 1, 第 2 ホルマント共振回路に入力し、ピークレベルの係数を乗じた後加算し出力している。

ホルマントのピークレベルは、図 2 のようにマスクノイズの平均スペクトルレベルを基準にそれぞれ L_1, L_2 と定義する。以後の実験では $L_1=15$ [dB] とし、第 1 ホルマントピークレベルは、常に可聴領域内に保つ。F2 のピークレベルは、第 2 ホルマント成分との関連を調べるために変化させる。

3. 予備実験での検討

可聴領域内に存在する音声のホルマント情報と母音判別との影響を探るため、典型的な 5 母音の聴取を 2 人の健聴者で行った。ピッチ周波数 $F_0=129$ [Hz], 時間長 $T_d=500$ [ms], サンプル周波数 $F_s=12$ [kHz], 合成母音のホルマント周波数は $/V/(F_1$ [Hz], F_2 [Hz]) で示すと, $/a/(800, 1400)$, $/i/(200, 2400)$, $/u/(300, 1300)$, $/e/(600, 2200)$, $/o/(500, 800)$ である。 L_2 は, -42 [dB] ~ $+12$ [dB] の範囲で 6 [dB] おきに各母音で 10 段階設定

*The effect of controlling formants peak levels and its application to amplitude compression for vowel identification.

し、50 個の試料をランダムリストにより提示、休憩を挟み 10 回聴取した。マスクングノイズの時間長は 1.5 [s] で、聴力レベル (HL) で平坦特性になるようホワイトノイズにインパルス応答を畳み込んで作成し母音試料はノイズの中央位置に加算後、提示した。被験者は聴力検査室内に着席し、予め MCL に音量を調整後、100 個の提示を 1 セッションとし、PC から自動的にヘッドホンの片耳に提示された合成音声を 5 母音の何れかに同定し記録した。聴取結果を要約すると、/a/ では、ノイズの有無によらず高い正答率を示し、他の母音試料では、マスクノイズレベルと第 2 ホルマントピークレベルが接近する $L_2=+6\sim 0$ [dB] 付近から、正答率の低下傾向が認められ、 $L_2=0$ [dB] を超えるレベル範囲でも正答率が漸傾している事が分かった [1]。

4. 母音カテゴリを用いた評価実験

4.1 刺激

提示する合成音声のホルマント周波数を図 3 に示すように F_1 - F_2 平面で格子状にとり、聴取の変化についての検討を行う。 $200 \leq F_1 \leq 800$ [Hz] に 100 [Hz] ステップ、 $600 \leq F_2 \leq 2600$ [Hz] に 200 [Hz] ステップで 60 個の組を定め母音カテゴリの変化と L_2 の関連を調べる。 L_2 は -12 [dB] \sim $+15$ [dB] まで変化させ、予備実験の半分の 3 [dB] 刻みで変化させた。音源は予備実験と同様インパルスとした ($F_0=129$ [Hz], $T_d=500$ [ms], $F_s=12$ [kHz])。

4.2 実験方法

1 セッションで、600 試料をランダム提示する。5 セッションを 2 人の健聴者 (成人男性) で実施した。被験者は、聴力検査室内に着席し、ヘッドホンの片耳に MCL で PC から提示後 5 母音に同定し入力する。

4.3 実験結果

図 5, 図 6 に聴取結果のカテゴリ図を示す。

図 5 はマスクノイズ無しで、 F_2 のホルマントピークレベルが $L_2=6, -3$ [dB] の例である。両者を比較すると /e/-/a/, /u/-/o/ の境界域に F_2 方向の広がりが増える。 /u/ では L_2 の減少にともない、安定に聴取する領域が減っている。ノイズを重畳した場合のカテゴリ図 6 では、 L_2 が減少していくにつれて F_2 方向への空白領域が増え、 F_1 情報による判定傾向が顕著であり、/i/, /u/ と /e/, /o/ と /a/ の 3 カテゴリ化が進んでいるのが分かる。

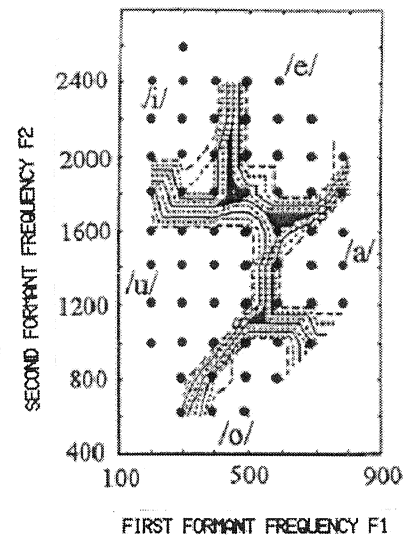
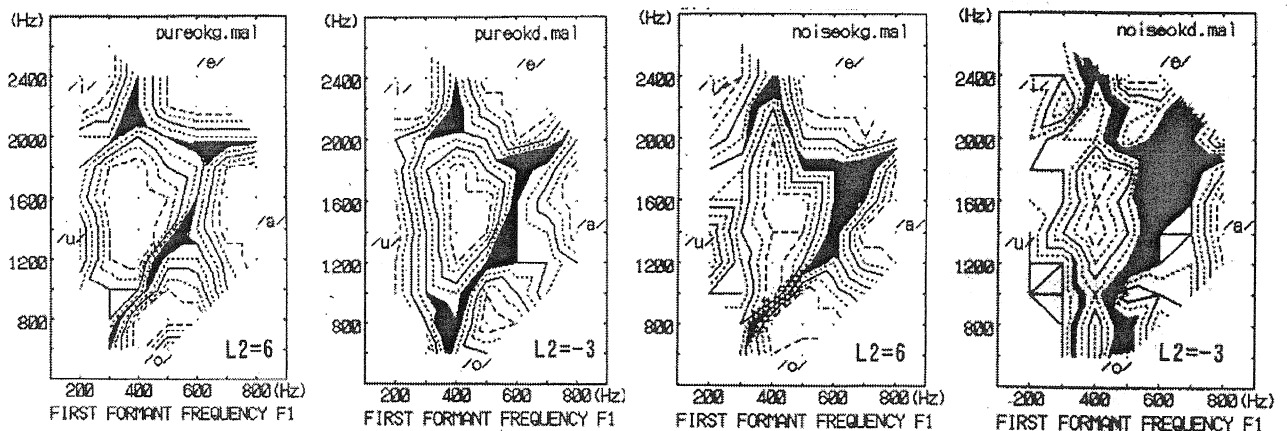


図 3. 試料の F_1 - F_2 配置



マスクノイズ無
図 5 カテゴリの変化

マスクノイズ有
図 6. カテゴリの変化

4.4 考察

カテゴリ類似度

各母音のカテゴリを考察するために、次のカテゴリ類似度を用いる[2]. 5母音においてF2ピークレベル(L2=15[dB])を基準カテゴリとし、それに対する各L2のカテゴリ類似度を

$$\delta_i = \frac{(P_i \cdot P'_i)}{(|P_i| \cdot |P'_i|)}$$

$$P_i = (p/a/, p/i/, p/u/, p/e/, p/o/)$$

$$\bar{\delta} = \frac{\sum_{i=1}^{60} \delta_i}{60}$$

$$P_i = \text{基準カテゴリ}, P'_i = \text{比較カテゴリ}$$

求めた. 図8に、マスクノイズ無しでの

カテゴリ類似度変化を示す. 類似度は各母音でL2=15[dB]の提示音声60個のうち各母音へ70[%]以上同定された分(/a/: 10個, /i/:5個, /u/:16個, /e/:7個, /o/: 5個)と, 全提示音声60個の平均の延べ6通りについて示している.

図9に、マスクノイズ無しのL2=15[dB]の結果を基準カテゴリとしてマスクノイズを重畳した時の各ピークレベルの聴取結果で求めたカテゴリ類似度変化を示す.

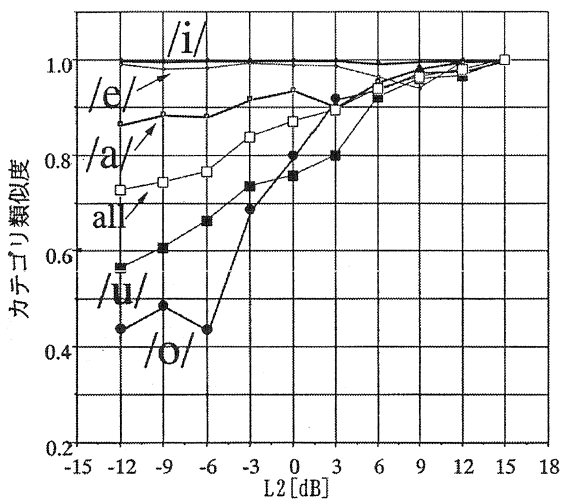


図8. カテゴリ類似度変化

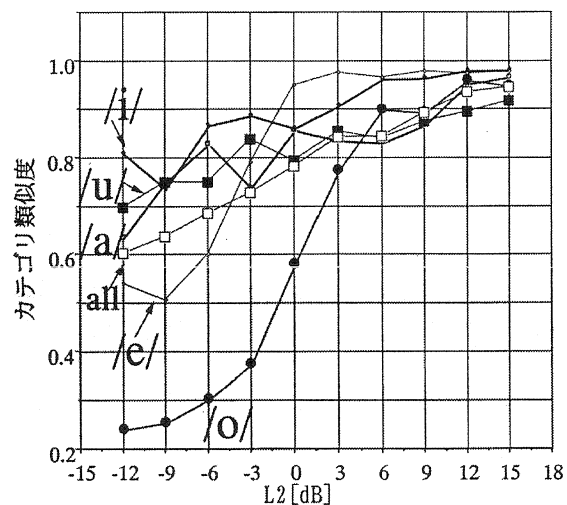


図9. カテゴリ類似度変化

図8では、L2の低下に関わらず/e/, /i/では9.4以上、また/a/では8.7以上と高い値を維持している. これに対し、/u/, /o/はL2の低下に伴い類似度も減少している. またマスクノイズを重畳した場合の図9では、/e/がL2=0~-6[dB], /o/がL2=6~-3[dB]で急激なカテゴリ類似度の減少傾向を示す. 図8, 9よりマスクノイズ環境下では、/o/はマスクノイズの情報に存在する領域から類似度の低下が始まっているのに対し、/e/ではほぼL2=0[dB]付近から低下を生じている. このことから、ホルマントピークレベルを制御して聴き取りを改善する際にはF1-F2との関連を考慮する必要があると考えられる[3].

マスクノイズを重畳しない場合と同等の効果を与えるL2の検討

マスクノイズを重畳した時の聴き取りが、どの程度のL2で確保できるかを調べる. 各F1-F2の格子点毎に、ノイズを重畳しないL2=15[dB]の聴取と同等の聴き取りができるマスクノイズ重畳時の同じ点でのL2を実験結果から抽出する. 同じ格子点で満足する複数のL2がある時は最小のL2をとる. 図10に、この条件で抽出したレベルをF1-F2格子点ごとに表示する. 次に求めた各格子点でのL2での条件での聴取結果に従いマスクノイズ重

畳でのカテゴリ図を描きなおすと図 11 のようになる。

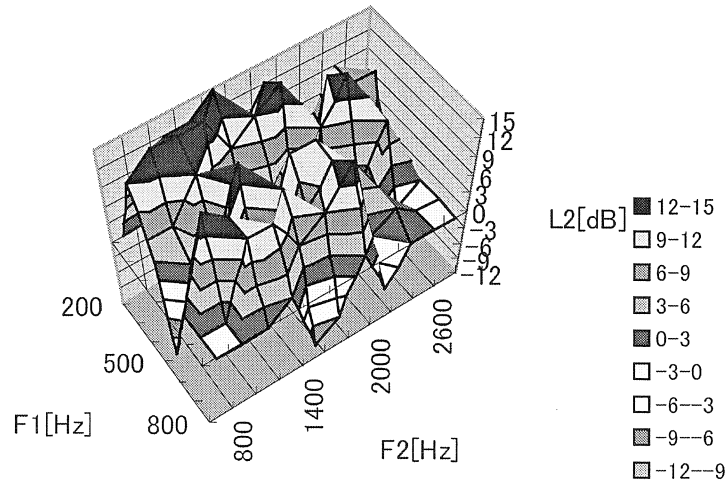
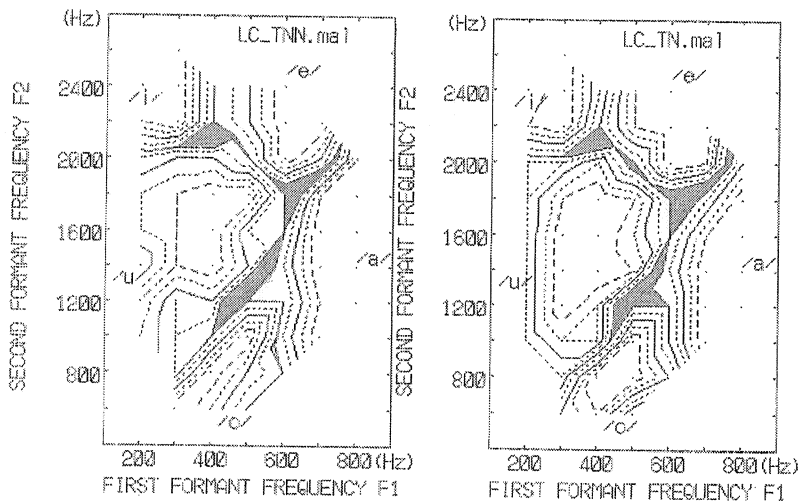


図 10. F1-F2 平面での L2 の最小値抽出



ノイズ無, L2=15 [dB] ノイズ有, データ抽出
 図 11. L2 が最小時の抽出結果のカテゴリ分布

図 11 で左がマスクノイズ無のカテゴリ分布で右がマスクノイズ重畳時の聴取実験結果から抽出した値を用いたカテゴリの状態である。マスクノイズ無し L2=15 [dB] のカテゴリ分布に近い状態まで回復しているのが分かる。

5. まとめ

F1, F2 を含む合成音声を用いた健聴者による聴取実験で母音知覚とホルマントピーク制御について検討した。健聴者の聴力特性で平坦となるマスクノイズを用い第 1 ホルマントのピークを可聴領域に維持した状態で、第 2 ホルマントのピークレベルを変化しカテゴリ類似度、カテゴリ図により考察した。その結果 F1-F2 平面で F2 ピークと母音カテゴリは、同平面の位置に応じ、聴取カテゴリを保つのに必要なピークレベルが異なる傾向が観察された。

今後、これらの傾向を更に検討して、振幅圧縮型補聴器での圧縮係数算定に反映させる方法を考察し実験と検証を行うことが必要である。

参考文献

- [1] 池田, 板橋, 上田, 坂田, 渡邊, ” 模擬難聴状態でのホルマントピークレベル比と母音知覚領域の関係”, 日本音響学会 2004 年講演論文集, pp. 483-484, 2004 年 9 月
- [2] Y.Ueda and A.Watanabe, ” Computer-assisted system for digital hearing aid and its application to the fitting procedure” Proc.H3-3, 15th ICA, China, Sep., 1992
- [3] 池田, 小埜, 坂本, 上田, 坂田, 渡邊, ” 母音同定に必要な第 1, 第 2 ホルマントピークレベル比と可聴領域の関係”, 電子情報通信学会技術報告会資料, SP2004-162 (2005-03), pp. 49-54