

騒音下における音声の明瞭度向上に関する基礎的検討 III

大島 静夫

Basic Study on Enhancement of Speech Articulation under Noisy Environment III

Shizuo OHSHIMA

(1996年11月30日受理)

This paper deals with a basic study concerning enhancement of articulation score under noisy environment. A new method adopted here is a partial technique of CLAUDHA (Compensate for Loudness by Analysing the Input signal Digital Hearing Aid) developed for a digital hearing aid by Tohoku University. In order to evaluate this technique, an articulation score test was used. In the experiment, a hearing level of stimuli was adjusted on a level which a subject is comfortable to hear. The stimuli is caught with a headphone in a quiet environment. The results of experiment were as follows. As an example, when a syllable was compensated 20 dB on S/N 10 dB, articulation score has improved 28% (64%-92%) in comparison with a non-processing syllable. Similarly, articulation score has respectively improved 25% (64%-89%) and 17% (64%-81%) in a syllable of 10 dB compensation and 5 dB compensation.

1. はじめに

最近騒音下での公共放送の明瞭度を向上するための、さまざまなシステム[1]~[3]が提案されている。また分野は異なるが、浅野ら[4]の開発したデジタル補聴器は、音声周波数領域に展開し、難聴者の残存聴野に音声スペクトルを収めるように加工し、再度時間領域に展開するという手法を用いており、感音系難聴者に特に有効とされている。感音系難聴者のラウドネス関数は、雑音が入力されたときの健聴者の特性に良く似ており、この手法は、騒音下の公共放送の明瞭度向上にも有効と考えられる。

騒音により音声の明瞭度が低下する原因は、いずれにしろ時間領域もしくは周波数領域における音声情報の欠落と考えられ、前報においては、時間領域における音声情報を補うという立場から、音声の時間領域における騒音によるマスキング分を補償するという方法を検討した。

本報においては、音声を短区間FFTを用いて周波数領域に展開し、周波数領域において第n次高調波成分に対し、騒音によるマスキング分を補償し、再度時間波形に変換する方法を検討した。この音声加工法の評価は、100音節の明瞭度試験で行ってい

る。

2. 音声の加工法について

2.1 ラウドネス補償関数

人間の音の大きさに対する感覚量に関しては古くから調べられており[5]、例えばLochnerらによる純音のレベルとラウドネスの関係などが上げられる。図1に文献の図を示す。図において(a)は雑音がないときのラウドネス特性の理論線、(b)~(e)は生体雑音を含め、最小可聴値を0.01[Sone]との交点で示される値だけ上昇させる雑音を加わったときのラウドネス特性である。ここで述べられていることは、最小可聴値が上昇するような雑音を加えたとき、純音の強さをI、域値付近のラウドネスをS'としたとき $S' = k(I^\alpha - I_0^\alpha)$ でほぼ近似でき、その曲線は純音レベルが上昇するとともに $S = kI^\alpha$ の雑音がない場合の漸近線に急に近づくという補充現象を示すということである。ここでkは係数、 α はほぼ0.3の定数である。

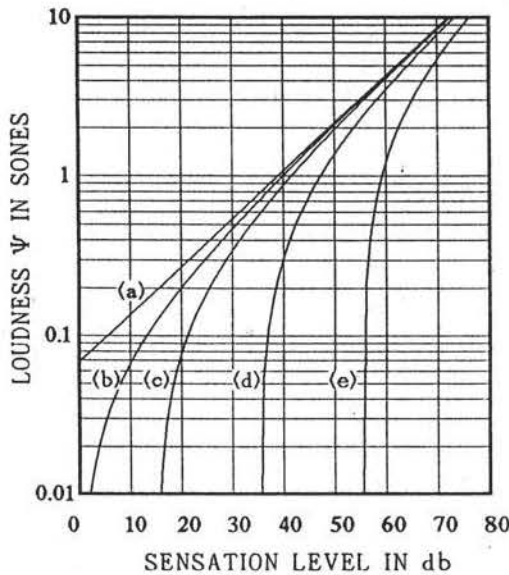


図1 1 [kHz] の純音レベルとラウドネスの関係 [5]

図2はラウドネス補償関数の説明図である。図において対角線は、騒音のないときの入力音に対する感覚量を示す。また a, b は、騒音を加わったときの、最小可聴値および不快レベルの例とする。ここで不快レベルは、人間がうるさく感じ始める音のレベルであり、公共放送においては、ほぼ80[dB]とされている[6]。いま図1の(b)~(e)のようなラウドネス特性が0.1[Sone]以上で図2の破線に対応し、この曲線がほぼ直線近似できるものとする。

図2において p[dB]の音を入力した場合の騒音時の感覚量は c 点となり、対角線との差分 (Gain で示す) が感覚量の減少分となる。この差分を補えば対角線と一致し、騒音がない場合と同じ感覚量を生ずることになる。このことは、もとの対角線に図で示した Gain を加えることと等価であり、同様な作業を a 点で行えば a' 点となり、不快レベルの点を b' とすれば、この間を直線で結んだものがラウドネスを補償する直線となる。これをラウドネス補償関数と呼ぶ[4]。

図において a 点の騒音を加わったときの最小可聴値に対応する音圧を p_a 、b 点の不快レベルに対応する音圧を p_b とする。更にラウドネス曲線の傾きを m 、 p を任意の音圧、 p' を補償後の音圧とし、諸条件を代入すれば

$$m = \frac{p_b}{p_a - p_b}, \quad p' = p^{1-m} 10^{\frac{mp}{20}}$$

で求められる。以降 p_a の値を p_a [dB] 補償と呼ぶ。

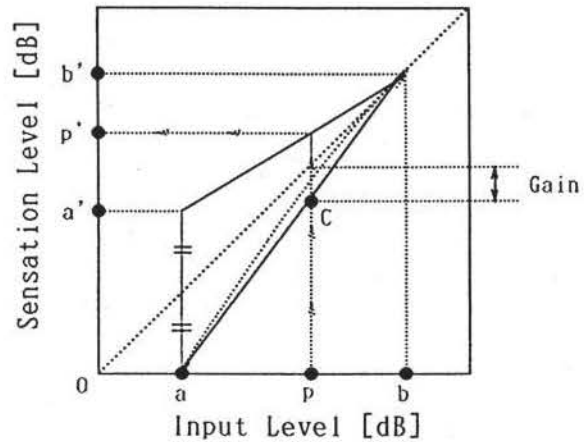


図2 ラウドネス補償関数 [4]

2. 2 音声の加工法

図3(a)に単音節/ki/の時間波形を示す。このような単音節のサンプリング時間60(μS)、128ポイントの短区間FFTを求める。128ポイントを1フレームとした場合の/ki/の子音部としてフレーム8、母音部としてフレーム32のFFTの結果を図3(b), (c)に示す。図において-20 dBに説明のための補助線を引いているが、例えば(b)の子音部では、すでに子音の特徴を示すスペクトルが補助線以下の部分もあり、さらに大きな騒音を加わったときには子音の特徴が失われる可能性が大きい。一方(c)の母音部では、低次のホルムアント周波数のレベルが補助線よりかなり大きく、さらに騒音が上昇しても母音としては判定可能と思われる。事実母音間の異聴は明瞭度がかなり低い場合以外はほとんどあらわれない。このように子音部の判断情報が失われることが異聴の少なくとも一つの原因であると思われる。そこで音声波形をその特徴を損なわないようにする方法として、人間本来の聴覚の特性を用いたラウドネス補償を利用している。

次に問題となるのは聴覚の周波数特性である。Fletcher-Munson 曲線などで良く知られるように周波数によりその最小可聴値が広い範囲で変化する。ところがノイズ下ではその様子が違って来る。文献[7]における白色雑音による純音のマスクングの図を図4に示す。

図4は「実験結果を多少きれいに書き直したものである」との注釈付きであるが、最小可聴値の周波数特性がホワイトノイズのレベルが上昇すると比較的平坦になっていくことがわかる。例えば静寂時には最大約45[dB]の最小可聴値の変動幅が、

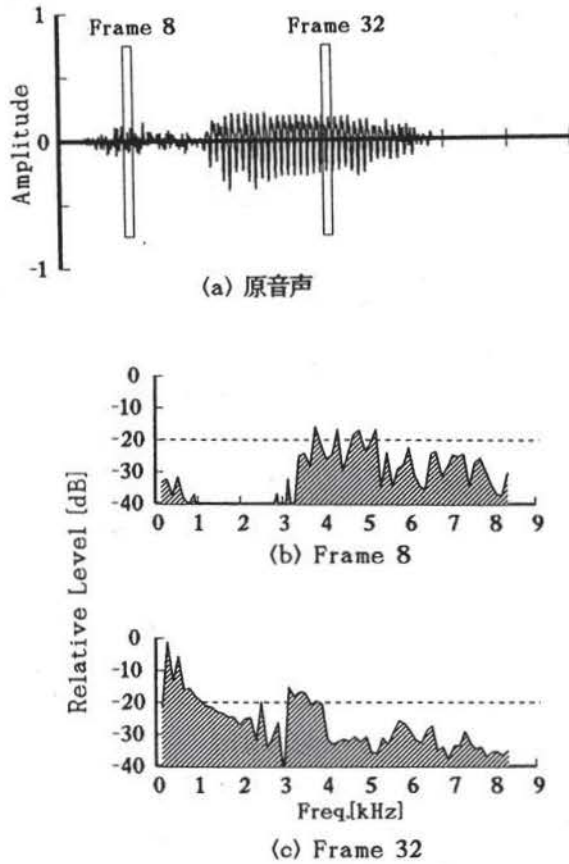


図3 単音節/ki/の原波形および加工波形

20[dB]以上のホワイトノイズ下では、約12[dB]程度になっている。このことより、厳密には p_a は周波数に依存するが、ここでは簡単のために p_a を周波数によらずに一定の値として、スペクトルのマスキング分を補うことにした。騒音の種類によっては p_a に周波数特性をもたせる必要もあると考えられるが、これは今後の課題である。

図5に10 dB 補償した場合の音声波形およびフレーム8, 32の例、図6に同様に20 dB 補償の例を示した。図5, 6からわかるように-20 dBの補助線をスペクトルが超えており、この操作により騒音によりマスキングされた部分が補償されることがわかる。また10[dB], 20[dB]補償したときのスペクトルの形状も原音声のスペクトルと相似になっており、周波数領域での音声判断情報は比較的損なわれていない。

3. 実験について

3.1 実験システムおよび被験者

実験で使用したシステムを図7に示す。図におい

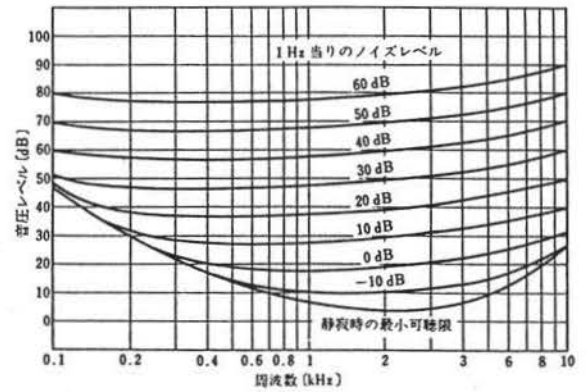


図4 白色雑音による純音のマスキング

てCH1にはノイズ信号、CH2には音声信号を送っている。これらの信号は、D/Aコンバータでアナログ信号に変換し、遮断周波数8 [kHz]のLPFを通して信号を円滑にしている。また、デジタルアッテネータで各信号のレベル調整を行い、音声信号とノイズ信号をミキシングアンプで混合し、ヘッドホン(YAMAHA YHD-3)で再生している。音圧の校正には6 [CC]カップラ、1/2インチコンデンサーマイク(TYPE7023)、精密型騒音計(TYPE6041)およびピストンホーン(いずれもACO社)を用いている。

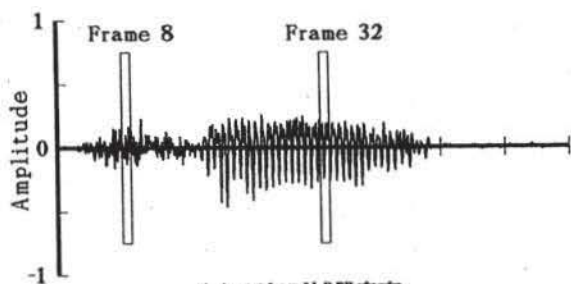
全ての実験において明瞭度試験は100音節で行った。単音節は、量子化量12[bit]、サンプリング周波数16.6[kHz]で録音したものである。単音節の出現順序は20種類の乱数表を用い、提示している。また刺激音は疑似騒音としてのホワイトノイズを提示後150~180[mS]の乱数による調整間隔をあけて音声提示し、ノイズと音声の関係が特定のタイミングにならないように留意している。被験者は19~20歳の男子学生4名で、全員実験の目的は熟知している。

3.2 明瞭度試験の結果

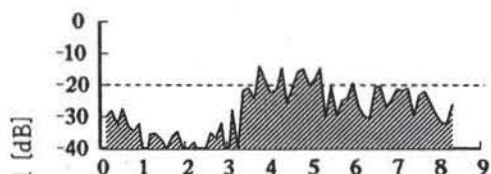
ラウドネス補償法の明瞭度試験の実験総数は、100音節×4点(S/N比5[dB], 10[dB], 15[dB], 20[dB])×5種類(原音声, 5[dB], 10[dB], 15[dB], 20[dB])×4人である。図8に実験結果を示す。

図8よりまず全体的には、補償音声相対的S/N~20[dB]までの範囲で原音声より明瞭度が向上していることがわかる。相対的S/N20~[dB]以上では、原音声の明瞭度の方が高いが、これは補償音声とその音声加工により逆に聞き取りにくい音声となったためと思われる。特に20[dB]補償のとき、相対

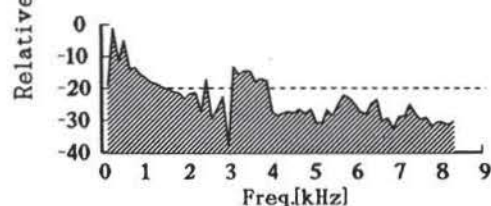
騒音下における音声の明瞭度向上に関する基礎的検討 III



(a) 10[dB]補償音声

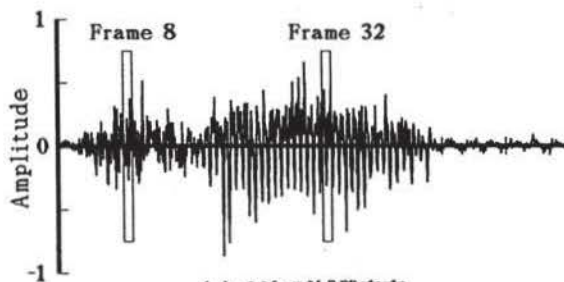


(b) Frame 8

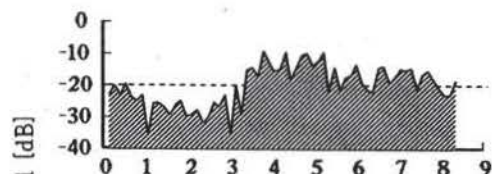


(c) Frame 32

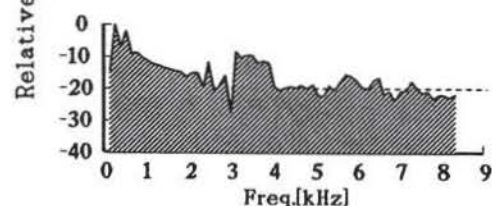
図5 10 [dB] 補償による単音節加工波形例



(a) 20[dB]補償音声



(b) Frame 8



(c) Frame 32

図6 20 [dB] 補償による単音節加工波形例

的S/N20[dB]以上ではS/N比の上昇とともに明瞭度が低下している。実際には相対的S/Nが20[dB]程度であれば、原音声の明瞭度が90[%]を超えていることからわかるように、騒音があまり気にならず補償をほぼ必要としないレベルである。原音声の明瞭度の向上が期待される相対的S/Nが15[dB]までの範囲では、5[dB]補償のとき約10[%]、20[dB]で約15[%]、15[dB]のとき約30~15[%]、20[dB]のとき約40~17[%]程度明瞭度が向上している。個の結果より、ラウドネス補償は騒音下における明瞭度向上に関しても有効な手法であることがわかる。

4. 結 論

図8より騒音レベルが高いときには高い補償を、低いときには低い補償を施し、騒音がないときには原音声そのものを利用すれば、総合的な騒音下での公共放送の明瞭度は向上するといえる。また、騒音のレベルに応じた補償を行うことにより、放送のレベルをあまり大きくせず、より効果的な明瞭度を向上させるシステムが構築可能と思われる。

具体的には、受話者の周囲の騒音レベルをマイク

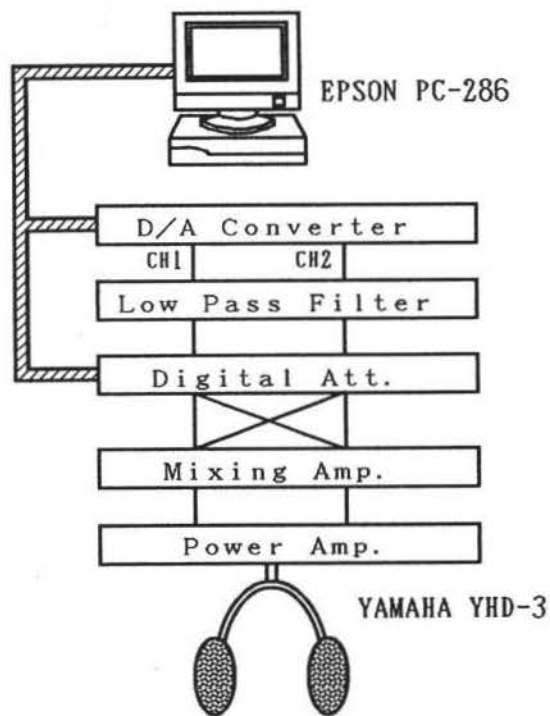


図7 単音節再生システム

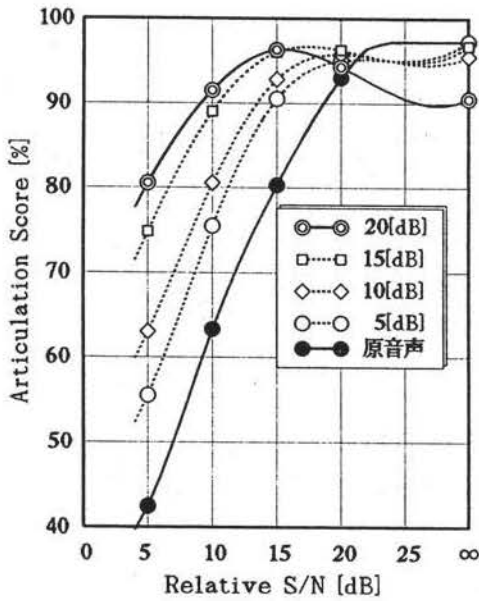


図8 明瞭度試験の結果

で実時間で捕らえ、静かな環境（防音室など）におかれている送話者の音声を、その騒音レベルに応じたラウドネス補償関数で加工し、受話者に送るという方法である。今後の課題としてはDSPの手法などを用いたリアルタイムのシステムの検討、また騒音の種類によっては、ここで一定とした最小可聴値の周波数による違いを考慮に入れたシステム検討なども必要と考えられる。

参考文献

- [1] 安藤善康, 長田総一郎, 野呂雄一, 井 研治, 久野和宏: “子音区間の補強による明瞭度の改善効果について”, 電子情報通信学会信学技報 EA91-7, pp. 25~29, 1991
- [2] 栗栖清浩, 福山和男: “ファジィ, ニューラルネットワークを用いた拡声音制御システム—騒音下における明瞭性改善のための知識獲得—”, 電子情報通信学会信学技報 EA93-82, pp. 1~23, 1994
- [3] 西村竜一, 浅野太, 鈴木陽一, 曾根敏夫: “ウェーブレット変換を用いた音声強調によるSNと明瞭度の改善効果に対する検討”, 電子情報通信学会信学技報 EA95-32, pp. 1~8, 1995
- [4] 浅野 太: “補聴器のためのデジタル信号処理手法に関する研究”, 東北大学審査博士学位論文, pp. 21~26, 1991
- [5] 曾根敏夫, 鈴木陽一: “ラウドネス”, 日本音響学会誌44巻10号, pp. 768~774, 1988
- [6] 城戸勝孝, 野呂雄一, 井 研治, 久野和宏: “騒音下における最適な放送レベルについて”, 電子情報通信学会信学技報 EA91-9, pp. 39~59
- [7] 境久雄, 中山剛: 「聴覚と音響心理」, コロナ社, pp. 111~113, 1978