

多値化音声信号を利用した骨導通信の検討

石原 学*・白滝 順**・家入 勝吾**

On the Bone Conducted Communication using Sliced
Speech Waves by some levels

Manabu ISMIHARA*, Jun SHIRATAKI** and Shogo IEIRI**

Abstract

This paper describes the experimental results of reducing the normally very complicated speech wave to simple geometrical form without impairing intelligibility, and also the results of articulation tests of simplified speech. The simplification of the speech wave is performed by slicing the original input wave using comparator(s).

1. ま え が き

音声波は、聞きとることができる言語情報として、人間対人間のコミュニケーションには有効な手段の一つである。しかし、音声はいろいろな雑音環境下で聞きとることが要求されるので、現在までに雑音環境下における研究、たとえば交通騒音下での聞きとりに関する研究¹⁾や単音節音声の明りょう度特性などが研究されている²⁾。また、聞きとりの特徴は年齢によって異なることから、聞きとり環境についても検討が行われている³⁾。

音声伝達妨害は、音声と騒音のエネルギー比という物理的な関係に強く依存しているが、このような高騒音場でも十分な信号対雑音比を有する音声伝達を実現する一手段として、骨導振動をピックアップする骨導伝送の適用が考えられている⁴⁻⁶⁾。

また、人間の音声には、多量の情報が含まれているために、その音声情報を考える手がかりとして、零交差波^{7,8)}が研究されてきた。零交差波とは、音声波 $f(t)$ の振幅情報が $f(t) \geq 0$ のとき振幅 A 、 $f(t) < 0$ の時、振幅 $-A$ (または 0) となる方形波に成形した波をいう。この場合においても、音声情報が残っていることが知られており、日本語零交差波の特徴が重永⁹⁾に

よって示されている。また、零交差波を利用した音声認識装置や通信方式⁹⁾などの研究があり、その実用性や有効性が示されている。

しかし、実際に零交差波を聞きとる場合には、ひずみや微小なノイズなどにより、単音節明りょう度はよくない。このことは、零交差波が持つ特徴であり、特に子音成分の変形が原因であると考えられている⁹⁾。零交差波に関する研究では、田宮ら¹⁰⁾が定レベル交差波として最適スライスレベルを求める実験を行っている。しかし、これは音声波の中心(以下“零レベル電圧”と呼ぶ)から正側へスライドしたものであった。しかし、スライドをさせることで、正負によって音声情報が異なることを指摘し、それが欠点として述べられていた。また、スライス情報を取り出すときに、微分波形をとりそれにより時間位置を確保し、さらにパルス波形を作成するという手順で行っており、回路手順に多くを必要としていた。

本論文では、コンパレータを用いた回路によって、波形成形を行い、さらにスライスレベルをいくつか変化させることにより、音声波形のスライスレベルと明りょう度との関係について論じている。この波形成形された音声波を多値化音声信号(3値化, 5値化音声信号)と呼ぶ。さらに高雑音環境下においては、通常のマイクを利用しての通信方式は、環境雑音の混入により、著しい明りょう度の低下をまねく。そのため、明りょう度の改善のために骨導を利用する方式で、高雑

1992年9月25日受理

* 電気工学科非常勤講師

** 電気工学科

音環境下でも利用できる通信方式が考えられている。しかし、骨導音声は、人体によるフィルタ効果によって周波数の高域成分での減衰が大きく、明りょう度の低下が指摘されている。そのために、明りょう度改善を目的として、多値化音声の有効性について検討がされている¹³⁾。本論では、この方式での可能性についてさらに検討を行った。

その結果、(1) 3 値化スライスレベルが -40.5 dB のとき、明りょう度 63% と最大の値を得た。(2) 付加した一対のスライスレベルが $-14.5 \sim -11.0$ dB のとき 68~73% の最大明りょう度を得た。(3) 振幅の絶対値の最大のものについて、この振幅を変えて調べたが特別な変化が見られなかった。(4) 3 値化音声信号の文章理解度は $-45.0 \sim -34.5$ dB で 80% 以上得られた。(5) 骨導音声に多値化への波形処理を施すと、骨導音よりも聞き易くなるのが、主観評価の結果明らかになった。これらの結果より、本実験での被験者は、聴覚正常でかつ受聴に関して特別な訓練を受けていない者を被験者としている。よって、実用上の有用な資料を提示していると考えられる。

さらに、波形処理回路はコンパレータ回路で作成しているために、回路の簡単化および小型化ができると考えられる。

2. 多値化音声信号の明りょう度特性¹²⁾

2.1 零レベル電圧の除法について

2.1.1 実験

回路の構成を Fig. 1 に示す。

音声を零交差波へと波形成形したとき、微小な零レベル付近のノイズ等も信号として処理され、明りょう度に影響をあたえる可能性がある。田宮らは¹⁰⁾ 定レベル交差波として最適スライスレベルの実験を行ったが、これは零レベル電圧より正側へスライドさせた実験であった。そこで本節では、零交差波に波形成形を行うとき、零レベル電圧付近の信号を除去することとした。(以下“3 値化”または“3Level”と記述する)。

Fig. 1 に示すように、回路は Filter (B.P.F. 300~3,500 Hz) の後段に並列回路構成したコンパレータによって正・負両側をそれぞれスライスし、音声波の中心部分である零レベル電圧を除去する。次に加算回路によって、正・負側を合成して 3 値化信号とする。その出力波形を Fig. 2 に示す。

実験に用いた音声資料としては¹¹⁾、NHK アナウン

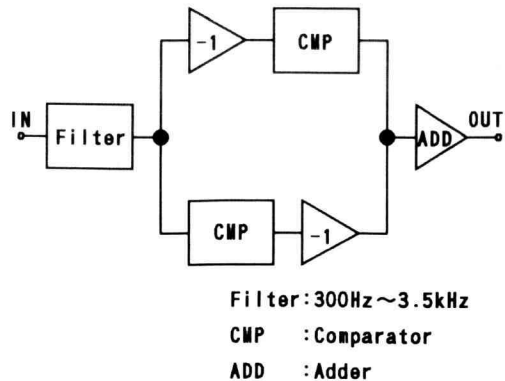


Fig. 1. Block diagram of 3 level.

サーによる 100 単音節の発声テープを使用し、被験者は 21 歳~35 歳までの聴覚正常な男性 5 名である。明りょう度の測定では、ランダムに並べ変えられた 100 単音節を被験者が聞きとった。被験者は聞きとった単音節を解答表に記入する。1 単音節を正解すると 1% とし、100 単音節全部正解で 100% とした。

被験者は、波形成形された信号をオーディオアンプ (テクニクス社, SU-CO1, SE-CO1) を通してヘッドホン (ビクター社, HP660) にて聞きとる。

そのときのオーディオアンプの出力レベルは、被験者の各個人が一番聞きとりやすい値としている。また、室内は静かな 3 m 四方の室を利用した。

スライスレベルは、零レベル電圧より -45.0 , -40.5 , -37.5 , -34.5 , -31.0 , -29.5 および

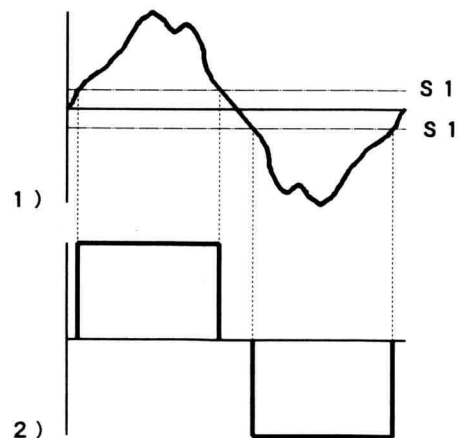


Fig. 2. Schematic illustration of output.

-28.5 dB の 7 通りに変化させた。このとき、音声資料テープに入力されているキャリブレーション信号のピーク値を 0 dB と定義している。

また、零レベル電圧とは、音声波の中心部分を意味し 0 dB とは区別している。

2.1.2 結果および考察

実験結果を Fig. 3 に示す。

スライスレベルが -40.5 dB のとき、63% と明りょう度が高くなる。このことは、-40.5 dB の位置がノイズ成分などを除去し、音声成分を保存している値と考えられる。また、スライスレベルが -28.5 dB のとき、もっとも低い明りょう度の値を示した。

最も明りょう度の良い -40.5 dB の子音別明りょう度でも、/d/, /ny/, /hy/, /my/, /ry/, /gy/ などの各子音の明りょう度は平均値より 30%~40% 低い、これは、3 値化する時点ですでに子音成分が崩れていると考えられる。また、-40.5 dB と -28.5 dB の子音別明りょう度を比べると、/s/, /z/, /ky/, /sh/, /dy/ などの明りょう度が顕著に低下している。このことから、スライスレベルにより影響を受けていることがわかる。この傾向が 3 値化の全体的な特徴であることが実験から認められた。

2.2 スライスレベル可変による実験

2.2.1 スライスレベル可変について

(1) 実験

回路の構成を Fig. 4 に示す。2.1.1 により得られた最適明りょう度 -40.5 dB をスライスレベル S1 とし

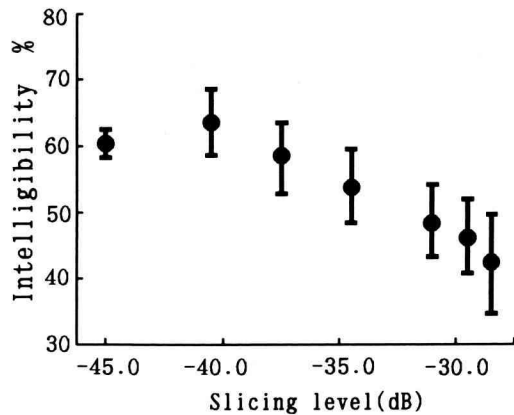


Fig. 3. The articulation of the 3 level-sliced speech wave of various slicing levels.

て設定し、もう一組のコンパレータの比較電圧 S2 を可変させることにより、明りょう度を測定する。このときの出力電圧の関係を Fig. 5 に示す (以下“5 値化”または“5Level”と記述する)。S2 の値は、-28.5, -20.5, -14.5, -11.0, -8.46 および -6.52 dB の 6 通りである。音声資料・被験者などの条件は 3 値化の実験 (零レベル電圧除去) と同一とする。

また、このときの S2 の 0 dB の設定も 3 値化の設定と同一である。

(2) 結果および考察

実験結果を Fig. 6 に示す。S2 の可変範囲が、-14.5 ~ -11.0 dB のとき、明りょう度が 72.8~73.0% と最も高い値を示した。しかし、S2 の値を -28.5 ~ -11.0 dB の範囲にとっても明りょう度は 68~73% 程度と、5% 程度の差である。そして、-11.0 dB を超える電圧から明りょう度は悪化する。すなわち、S2 のスライスレベルについては電圧がある値 (しきい値) より大きくすることに問題があると考えられる。また、子音ごとに明りょう度の高低があると思われるが、-14.5 ~ -11.0 dB 程度のスライスレベルが平均して子音の特徴を保存していると考えられる。また、特に影響を受ける子音は、/s/, /ry/, /gy/ などであると思われる。さらに、スライスレベルが -8.46 dB のときには /ry/ の明りょう度は 0% となる。これらの結果から、子音の明りょう度はスライスレベルに大きく依存していると考えられる。

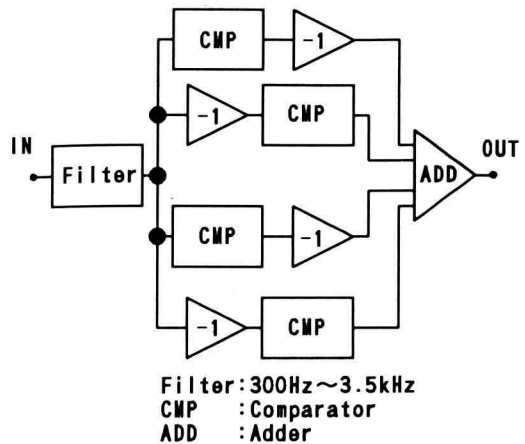


Fig. 4. Block diagram of 5 level.

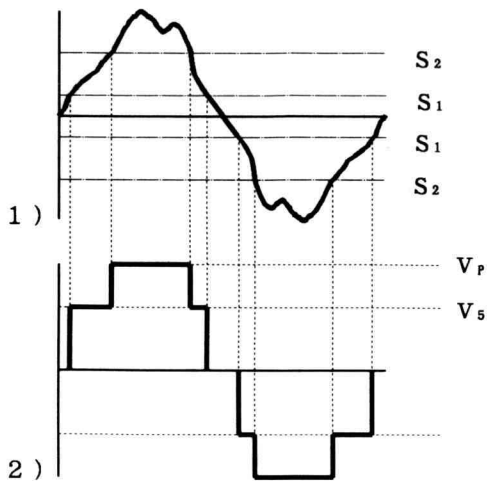


Fig. 5. Schematic illustration of output
 1) Speech wave
 2) 5 level wave forms

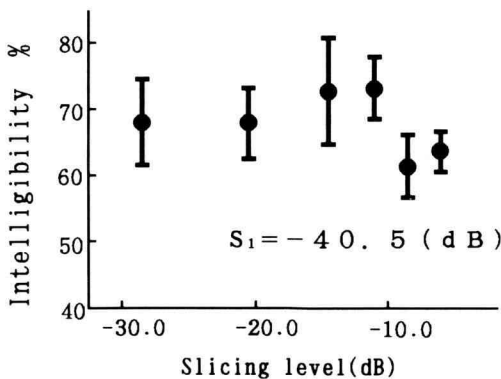


Fig. 6. The articulation of the 5 level-sliced speech wave of various slicing levels.

2.2.2 出力電圧比の可変について

(1) 実験

2.2.1 (1) の回路構成で、スライスレベル S1 を -40.5 dB, S2 を -11.0 dB に設定し、出力電圧の振幅比を可変させる。振幅比は、Fig. 5 に示すように出力波形のピーク電圧を V_p と V_5 との比で定義している。この振幅比を可変させて、明りょう度の測定を行った。その他の実験条件は、2.2.1 (1) と同一である。

(2) 結果および考察

実験結果を Fig. 7 に示す。

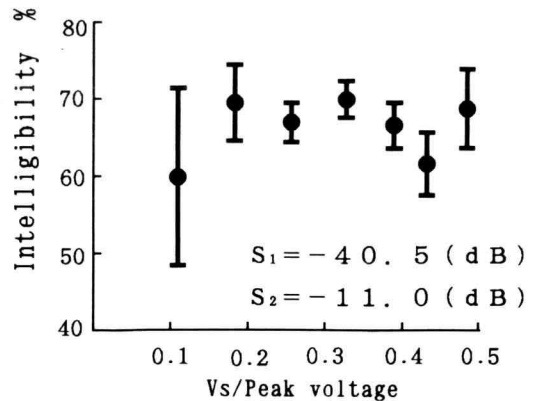


Fig. 7. The articulation score of combinations ratio of each comparator output is varied.

V_p を一定として ($5[V]$) 振幅電圧比 V_5 を変化させた場合には、スライスレベル S2 を変化させたときの実験のように特別な傾向が見られなかった。Fig. 7 から、出力振幅比 $0.185 \sim 0.390$ の範囲では、明りょう度が $66 \sim 70\%$ 程度であり比較的变化が少ない。この実験結果から、比較的明りょう度の変化の少ない、出力振幅比 $0.185 \sim 0.390$ の範囲を利用することで安定した明りょう度を得ることができると思われる。

3. 文章了解度について

3.1 実験

基礎的な研究として、すでに単音節音声の明りょう度は求めた。本節では、3 値化された音声信号に対して、文章了解度を明らかにする。

実験に用いた音声資料としては、NHK アナウンサーによる NHK ニュースを DAT に録音して利用した。被験者は 21 歳 \sim 35 歳までの聴覚正常な男性 5 名である。

了解度の測定では、NHK ニュースをいかに正確に聞きとることが出来たかで評価した。聞きとりの評価として、被験者は聞きとった文章を解答表に記入する。文章中で 1 単音節ごとに正解すると、+1 の得点をあたえ、間違えると 0 点をあたえることとし、文章が全部正解で 100% とした。以下に、計算式を記述する。

$$\text{文章了解度} = \frac{\text{正解の単音節数}}{\text{全文章の単音節数}} \times 100(\%)$$

かい はつか / しぜんほごかで / ぎろんを / おこなわ
 れている / とちぎけんなど / よんけん にまたがる /
 わたらせ / ゆうすいちについて / きょうからしゃか
 いとうの / こっかいぎいんだんの / げんちちょうさ
 が / はじまりました / げんちをおとずれたのは /
 しゃかいとう / かんきょうぶかいの / いわたれすき
 お / だいぎしなど / しゅうさんりょういんの / こっ
 かいぎいん / じゅうにんです / わたらせゆうすいち
 は / あしおどうざんの / こうどくじけんで / ぎせい
 になった / とちぎけんの / きゅうやなかむらが / む
 らぐるみで / いてんしたあとちで / あしのおいしげ
 る / ひろさきさんぜんさんびゃく / へくたーの / ゆ
 うすいちです

Fig. 8. Japanese Text.

被験者は、波形形成された信号をオーディオアンプ (テクニクス社, SU-C01, SE-C01) を通してヘッドホン (ビクター社, HP660) にて聞きとる。

そのときのオーディオアンプの出力レベルは、被験者の各個人が一番聞きとりやすい値としている。また、室内は静かな 3m 四方の室を利用した。

実験に用いた音声資料としては、NHK アナウンサー (男性) による NHK ニュースを DAT に録音して利用した。ニュースの内容は天気予報や実験日の近くで何度も報道されるような頻度の高いものはずし、ローカルニュース的なものを利用した。ニュースの長さは、時間的には 3分~5分程度とし、含まれる単音節数は 350 単音節~550 単音節程度で、意味の途切れない部分で区切った。また、上記の設定のニュースを一文節から三文節程度ごとに 5 秒程度の時間をあけ、書き写す時間を設定した。また、聞き直しは認めず、すべて一回しか聞くことはできない。被験者は 21 歳から 35 歳までの聴覚正常な男性 5 名である。このときの文章例を Fig. 8 に示す。図中の “/” は、書き取りの時間に設えた、一時休止部分である。

了解度の測定では、NHK ニュースをいかに正確に聞きとることが出来たかで評価した。聞きとりの評価

として、被験者は聞きとった文章を解答表に記入する。文章中で一単音節ごとに正解すると +1 点を与え、間違えると 0 点とし、文章が全部正解で 100% とした。また、聞きとれないものも 0 点とした。

スライスレベルは、零レベル電圧より -45.0, -40.5, -37.5, -34.5, -31.0, -29.5 および -28.5 dB の 7 通りに変化させた。このとき、音声資料テープ (DAT) に入力されているキャリブレート信号のピーク値は DAT の LED のレベルメータで 0 dB であり、NHK ニュースを 5 秒間再生したときのピーク値も 0 dB である。

3.2 結果および考察

実験結果を Fig. 9 に示す。

図に示すように、3 値化音声信号の文章了解度は単音節明りょう度の場合よりも最大で約 30% 程度向上する。文章としての了解度を意味理解として考えて、本節では 80% 以上得られれば、意味理解が可能と考えられる。そうすると、文章了解度の実用基準は -45.0 ~ -34.5 dB の間の値を利用すればよいことがわかる。

次に、5 値化音声信号の文章了解度を考えると、単音

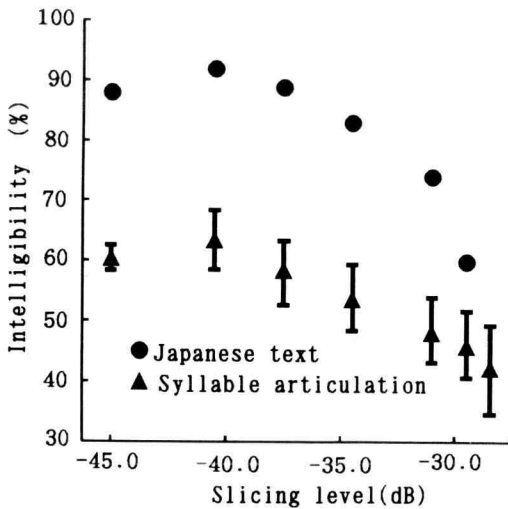


Fig. 9. The articulation of the 3 level-sliced speech wave of various slicing levels.

節明りょう度は3値化音声信号より高く、平均しても70%程度得られている。このことから、5値化音声信号の文章理解度は意味理解を考えた場合の実用基準である80%を十分に得られるものと考えられる。聞き易さの点について述べると、被験者が聞きとるときには、3値化の場合よりもさらに音声信号の自然性が向上し、聞きとり易くなっていく。

4. 骨導通信への利用について

本実験では発声に伴う骨導振動をピックアップする。そのときの、骨導振動は鼻の骨の部分で最も明りょう度が良いと指摘されているために、本実験でも、鼻の部分を使用した¹³⁾。また、この部分は法医学的にも個人差が少なく、表面の皮膚も薄いところである。このときの、骨導音は鼻の部分に超小型圧電型加速度ピックアップ(小野測器, NP-601)を皮膚に押しつけて、加速度ピックアップ用センサンプ(小野測器, PS-601)を介してDAT(SONY, DTC-1000ES)に録音した。次に、DATに録音された骨導音声信号を3値化音声信号および5値化音声信号へ波形成形を行った。北守らの研究¹³⁾では単母音についての分析を行い、母音/i/のFFT分析から明りょう度の低下の原因として第2ホルマントが不明瞭であったと指摘している。

実際に利用する場合は文章であるから、本研究では

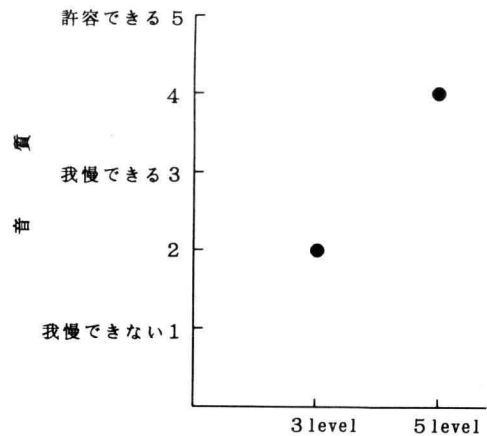


Fig. 10. Quality of sound.

3値化の文章理解度と同じ条件として、骨導の文章理解度について検討する。特に、ここでは3値化骨導音声と5値化骨導音声を受聴したときの聞き易さについて実験を行った。そのときの、聞き易さについての主観評価をFig. 10に示す。

まず、骨導音声のみの場合は、こもった音に聞こえる。これは、骨や皮膚を通しているためにフィルタのような効果があるためと考えられる。次に、3値化骨導音声信号においては、ノイズが発生するものの、骨導だけのこもった音よりも聞き易くなる。さらに、5値化音声信号においては、3値化音声信号よりノイズ音が気にならないで、受聴することができる。その結果、自然性が増加するような効果が生じて被験者にとって聞き易い音となっている。

騒音下での骨導伝送の有効性については、すでに述べられているが、本論文での波形処理を行った多値化音声信号を採用することにより、骨導のみの場合よりも聞き易くなるという効果が得られる。これは、骨導音を通信に利用する場合の有効な方法であると考えている。さらに、現在、この文章理解度特性について、明らかにすべく研究を進めている。

5. む す び

音声波の波形成形を行い、音声の明りょう度について測定を行った。また、骨導音を波形処理して文章を受聴した結果、次のことが明らかになった。

スライスレベルが明りょう度に対して大きく影響す

る。一般に音声信号に含まれる子音成分は振幅が小さく、子音成分のレベルとスライスレベルとのレベル差が影響を与えていると考えられる。

零レベル付近の信号の除去は、田宮らの実験において-30~-40 dBであった。しかし本実験では、零レベル付近のノイズを除去し、良好な明りょう度が得られたレベルはスライスレベルが-45.0~-40.5 dBであった。この相違は、音声波の中心部分を零レベルとして、正・負の両側ともに除去している本実験に対して、田宮らは正側へ偏らせた実験であること、また、0 dBの定義の相違によると考えられる。

子音別にみると、3値化、5値化の順に高い明りょう度が得られる。しかし、/ny/, /hy/, /my/, /gy/については明りょう度が多少上昇するものの、50%以下である。これは、3値化の段階で子音成分の振幅が小さく、この子音成分が切り捨てられたために、5値化においても明りょう度が上昇しないものと考えられる。

被験者によると、5値化の音声は3値化の音声より聞きとりやすいとの意見がある。これは、5値化することにより音声信号の自然性が向上したためによると考えられる。

3値化音声の文章理解度を求めた結果、スライスレベルが-45.0~-34.5 dBの間で、理解度80%以上が得られた。本論文では、意味理解の基準として80%を採用し、このスライスレベル範囲が実用になるものと考えられる。

本実験では、明りょう度の測定を通して音声を聞きとる場合の明りょう度とスライスレベルとの関係を明らかにした。

さらに、骨導音声に3値化および5値化音声信号の処理を行った結果、受聴の主観評価実験から、聞き取り易くなることが明りょう度の向上につながったものと考えられる。

今後、骨導音声よりも骨導音を波形処理した多値化骨導音声信号が聞き易い原因について物理的な意味付けを行いたいと考えている。また、多値化骨導音声信

号の文章理解度について明らかにするつもりである。

参 考 文 献

- 1) 曾根・植松・金指・二村：“ランダム変動騒音下の音声明瞭度の予測（変動騒音の言語聴取妨害に関する研究その2）”，音響学誌，35，58（昭和54-2）
- 2) 石原・白滝・家入：“雑音環境下における単音節明りょう度特性”，電学論C，112，2（平成4-2）
- 3) 鈴木・林：“雑音下における無意味語の若年者および高齢者による聴取について”，人間工学，27，1（平成3-1）
- 4) 柳沢・降旗：“高騒音下の音声通信における振動ピックアップの利用”，音響誌，31，3（昭和50-3）
- 5) 井・渡辺・大山・荘司：“骨導母音分析による上顎骨病変の診断的応用”，ME学誌，16，1（昭和53-2）
- 6) 宇佐川・角・江端・奥田：“騒音下における音声の認識率向上に関する一考察”，信学論A，J70-A，12（昭和62-12）
- 7) J.C.R. Licklider and I. Pollack：“Effects of differentiation, integration, and infinite peak clipping upon the intelligibility of speech”，J. Acoust. Soc. Am., 20, 1（1948）
- 8) 例えば、三浦監修：“新版 音声と聴覚”，信学会（昭和57）
- 9) 重永：“日本語音声の零交叉波時系列 I. 測定方法および母音，II. 無声子音”，音響誌，17，2（昭和36-2）
- 10) 田宮・平松：“定レベル交叉波の性質とその応用”，音響誌，14，2（昭和33-2）
- 11) 明瞭度試験用テープおよび明瞭度試験法の基準，日本音響コンサルタント協会（昭和54）
- 12) 石原・白滝：“音声信号のスライスレベルと明瞭度の関係について”，人間工学，23，4（昭和62-8）
- 13) 北守・滝沢：“明りょう度試験による骨導音声の分析”，信学論A，J72-A，11（平成1-11）