

言語学 2011 1726 号

残響および多重音声による
通話品質劣化の研究

井 研 治

目 次

第1章	まえがき	1
第2章	音声通信と妨害雑音	3
2.1	はしがき	3
2.2	音声通信と妨害雑音	3
2.3	妨害の評価法	5
2.4	まとめ	6
第3章	通話品質に及ぼす残響の主観的効果	7
3.1	はしがき	7
3.2	残響付加方式	7
3.3	人工残響音のプレファレンス	10
3.3.1	プレファレンス測定法	10
3.3.2	信号対雑音比の規準化	12
3.3.3	残響時間の規準化	14
3.4	残響時間によるプレファレンスの表現	15
3.5	まとめ	18
第4章	残響の主観的評価量と物理量の対応	19
4.1	はしがき	19
4.2	残響特性の評価	19
4.3	プレファレンスとインパルスレスポンスとの関係	20
4.4	多重エコーを伴う音声のプレファレンス	22
4.5	まとめ	24
第5章	残響の明りょう度に及ぼす影響	26
5.1	はしがき	26
5.2	音声の時間的性質	26
5.3	擬似音声による検知限測定	29
5.4	人工残響付加による明りょう度の低下	32
5.5	まとめ	35

第6章	多重音声による通話妨害	37
6.1	はしがき	37
6.2	多重音声の性質	37
6.3	多重音声の騒音レベル分布	43
6.4	多重音声による明りょう度低下	46
6.5	まとめ	50
6.6	付 録	50
第7章	あ と が き	52

謝 辞

参 考 文 献

第1章 ま え が き

原始社会において人類が集団生活を営みはじめてから、たがいに通信を交わす方法がいくつか考えだされてきたが、とりわけ音声通信は人類が発見した通信手段の中で他のどんな方式よりも便利な方法であり、言語という符号体系が習慣として確立するとともに音声伝達できる情報量は極めて多くなった。しかし人間の発する音声の大きさには限度があって遠距離通話が困難であることや音声現象が瞬時的であるためたちまち消散するという2つの欠点を持ち続けていたが、1876年 Bellが電話機を、またその翌年 Edisonが機械的録音方式を発明したことによって音声通信の利用は飛躍的に拡大し、20世紀の科学技術の急速な進歩とあいまって音声有していた空間的・時間的な制約はほとんど克服されつつある。

音響に関する学問は古くから主として音波の伝わる速さや楽器の共鳴についての研究として続けられてきたが19世紀終りに Helmholtz や Reighleigh などによって音波の基本的な性質が明らかにされ科学的な発展を遂げた。そして音響学は音波の反射、吸収などの現象から建築学と結びついて研究が進められ、1895年 Sabine は室内の残響現象に注目して残響時間を定義し、以後の室内音響学の基礎をつくった。

一方、通信とはより少い損失でより多くの情報を相手に伝達することが究極の課題であり、この目的の遂行のため通信工学の多くの分野では詳しい解析がなされてきた。しかし音声による通信では物理的な取り扱いが困難な受話者の聴覚特性まで含めが議論する必要がある。上に述べた残響が聴覚に及ぼす影響は近年 Haas⁽¹⁾により単一エコーの研究として発表され、また Lochner^(2,3)も聴覚に及ぼす反射音の効果について詳しい実験を行っている。またわが国でも柳沢ら^(4~6)によって反射音が与える心理的影響についての報告がなされている。これらの研究は残響音を両耳による受聴としてとらえ拡散音場内で主観評価することによって考察されており、Meyer⁽⁷⁾は室内における拡散性を表現する尺度を提案している。また Schroeder ら⁽⁸⁾は音楽の観賞という立場からコンサートホールの音響効果と心理的印象との関連性を追求しており、最近の研究は情緒的な効果を反映させる物理因子を抽出する方向に発展する様相を見せている。

他方、19世紀にはじまる電話機の普及とともに電話伝送に関する基礎的指針を明確にする必要が生じ、Bell が設立した Bell 研究所において1920年代から音声通信に関する研究が始められた。そこでは Fletcher によって精力的な活動がなされ、1950年代には伝送線路の物理特性と聴覚を介した通話の性能を対応づけることが可能となった。⁽⁹⁾

わが国では音声による通信システムの1つである電話が近年飛躍的に整備されつつあるが将来の電話機利用の一形態としての拡声電話機^(10,11)についてもその研究、開発が進められている⁽¹²⁾。拡声電話機の利用は諸外国では古くから試みられているが^(10,11)、この場合通話の性能に影響する主な妨害要因として残響による妨害及び外来騒音の影響が考えられる。特に前者の妨害は送話諒及び受話側が4線になること、ならびに通常の室内が示す残響特性によっていわゆる送話エコー、受話エコー

を生じ残響感を増加せしめ、その結果音声の品質が低下する。

本論文はこのような拡声電話機等の音声通信においてその品質を劣化させる要因と考えられる残響の効果及び音声を主体とした騒音についてこれらが通話に与える影響を調査した実験的研究である。そして音声通信のより効率的な利用に際して音声信号と人間の聴覚に着目した基礎資料を得ることを主なねらいとし、前述の要因が音声通信に及ぼす妨害効果を工学的立場から定量的に評価することを試みる。

先ず第2章では通話に及ぼす妨害要因について若干の考察を試み、問題点の整理を行うとともに工学的立場から各種の心理測定法について検討を加える。

第3章では残響の存在が通話に及ぼす主観的效果についてプレファレンスが等価な雑音によって置きかえて評価する手法を導入し、各種の人工的な残響音の等価雑音がレベルをもとに残響時間との対応づけを規準化によって試みる。そして従来発表されている結果と比較検討し規準化の有効性を確認する。

続いて第4章では主観量に対応させた等価雑音レベルを単純化したインパルスレスポンスを用いて表現しエネルギー比との関連性を追求する。また多くのエコーを伴う音声、つまりマルチエコーを伴う音声の示す主観的效果を通話にとってのプレファレンスの面から考察し上記の関連性を実験的に裏付ける。

第5章では残響が通話に及ぼす影響を通信の実用的な尺度である明りょう度を用いて検討する。ここでは日本語の平均的な構造を念頭に置いて会話音声のモデルを考え、音声のマスクングを考慮して明りょう度計算の可能性を示すと同時に残響時間と会話の発話速度の関係についても論じる。

次に第6章では音声による騒音が通話に及ぼす影響を前章の会話モデルを用いて音声の瞬時振巾確率密度によって論じ、統計的な立場からいくつかの推定を試みるとともに妨害音声を騒音として分析し種々の性質を明らかにし通話に及ぼす実際的な影響を明りょう度によって考察する。

第2章 音声通信と妨害雑音

2.1 はし が き

音声通信系における最も基本的な妨害要因の一つは雑音による妨害であるがここでは音声通信システムとしての見地から雑音の発生、混入の可能性について問題点を整理するとともにこれら妨害雑音に起因する通話品質の劣化の評価法と受聴実験に用いる測定装置の特性についてのべる。

2.2 音声通信と妨害雑音

言語情報を音声によって伝達するいわゆる音声通信は最も効率的かつ便利な通信システムであるが、これは図2.1に示すように送話者の大脳で形成された言語のメッセージを遠距離伝達に適した音響信号つまり音声波形に変換する過程、及びこの音声波形がもとになってつくられる神経系の情報から再び言語メッセージに変換する過程によって代表される。¹³⁾

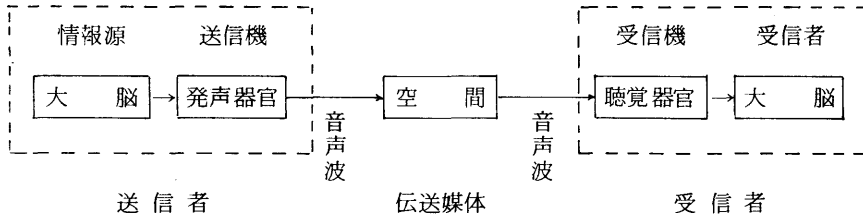


図 2.1 音声による情報伝達の過程

図に示した伝送媒体として最も簡単なものは送話者の発声器官によってつくられた音声波形がそのまま受話者の聴覚器官に到達する場合であり対面通話などはこの例である。

一方音声は直接伝達できる限界を越えて通話を可能にするような通信系では伝送媒体として電気音響変換器を含んだ単なる増巾器や複雑な変調、復調システムが介在することもあり、これらの媒体は電子工学の技術的な進歩によって性能が急速に高められた。

通信系の目的は言うまでもなく情報伝送であるが、伝送途中でいくばくかの雑音または確率的な変動が目的の信号に影響を与えることが特徴であって、そのため送話者の意図する言語メッセージが正確に受話者の大脳に伝達されるとは限らない。このため伝送の品質に対して一定の評価を与え、その性質を把握することが肝要であろう。通信系ではこの目的のために信号対雑音比、いわゆるSN比によって評価することが多い。ここでいう雑音とは伝送媒体の内部あるいは外部から混入する不要あるいは有害な成分を指していることはいうまでもないが、音声通信系においては通信機器

に関する雑音の他に送・受話端における音響伝達経路に混入する外来騒音も含めることができる。^(11,12) この外来騒音はいわばわれわれが居住している生活環境に付随する根源的な雑音であり最も多くの時間を過ごす建築物の中での音であって、その一つは物理的環境に起因する残響音であり、他の一つは社会的環境つまり集団の騒音といえよう。

前者については、われわれの生活する空間の壁面でなされる音の吸収が完全でなく必ず反射を伴うこと、及び音の伝播時間による遅延信号が重量されて発生するひずみが聴覚で感じとられるためである。このような残響現象を表現する物理尺度の一つに残響時間があり、これは定常状態に励振している音源を取り去ったとき音の強さが 60 dB 減衰するまでに要する時間によって定義されている。⁽¹⁴⁾ 残響が聴覚に与える影響については Haas による単一エコーを用いた研究があり、⁽¹⁾ それによれば約 30 ms 以内のエコーは直接音に隠れいされ音の聴覚的な大きさ、すなわちラウドネスを高める性質があると指摘している。反面 50 ms 以上の遅れを伴うエコーは直接音と分離した感じを与えるとのべており同様の報告は Lochner⁽²⁾ によってもなされている。残響は多数の複雑なエコーによって構成されるため単一エコーのように簡単に模擬することが困難であり定量的な評価はなされていないが、Knudsen⁽¹⁵⁾ は残響が受聴の際に及ぼす効果を分析し、残響時間の増加が室内の音圧を上昇させるため音声の受聴が効率よく行われる効果と、個々の音を引き伸ばす残響の作用のため受話者に与える混乱作用の 2 つを考えている。そのため残響時間には適当な範囲が存在し、会話もしくは音楽など用いる信号源の性質によってそれぞれ最適値が提案されていて⁽¹⁴⁾ 一般に会話に対しては短かく、音楽などには長い残響時間になっている。一方、残響時間以外の尺度も提案されており Meyer⁽⁷⁾ は直接音から 50 ms 以内に到達する反射音が一律に聴覚に貢献する立場から D 値についての考えを明らかにし、これは音の明りょうさ (Clearness) を反映する値であると述べている。また柳沢ら⁽⁴⁾ は反射音について心理的な分析を行い、D 値についての関連性を追求したが、通信工学的な立場から定量化の試みは他にあまり例を見ない。残響が付加された音声の品質を物理尺度に変換することができれば聴覚を介した反応を数値化してながめることができるが、そのためには主観を介させた測定方法も詳しく吟味する必要があり、これは次節で扱う。

さてもう一つの外来騒音、つまり社会的環境による集団の騒音についてはどうであろうか？ この問題はわれわれが雑踏などで会話を行うときよく経験するように希望の信号以外の騒音を無意識に抑制している。つまり信号対雑音比の上昇を神経系で実現しておりこの作用はカクテルパーティ効果⁽¹⁶⁾ として知られ、両耳聴機能を含めて議論せねばならないがここでは基本となる片耳受聴に限定して問題への接近を試みる。多くの音声通信系などを集中的にコントロールしている航空管制塔などでは航空機の輻奏時に通話が重なり了解性が悪くなることから近年研究が始められており^(17~21)、音声レベルに着目した考察は MacLean⁽²²⁾ や Plomp⁽²³⁾ によって、また明りょう度に関しては Pollack⁽²⁴⁾ によって実験がなされている。このような音声による騒音は会話音声の重畳されたものであり、発話者数及びそれらと受話者の幾何学的配置によってかなり正確に記述できる。また基本となる会話音声は母音と子音のくりかえしであることからこれら音声の基本特性を念頭において通話に及ぼす影響を考察してゆく。

2.3 妨害の評価法

通話の良さを定量的に表現するものを通話品質といい古くから種々のものが提案され用いられているが、そのうち主なものに明りょう度、平均オピニオン値、弁別限、等プレファレンス等がある。⁽²⁵⁾中でも明りょう度は伝送系の物理条件をもとに計算によって示すことが可能な尺度であり、日本語100音節を含んだ平等率試験音節表によって発話された音節が正しく受話者に伝達された割合によって通話者間の情報伝達効率を代表させている。母音・子音をそれぞれV、Cの記号で表わせれば平等率試験音節表とはVのみの単音が5種類、CVの組み合わせたったものが95種類、合わせて100の音節から成り立っている。一方米国では種々の音節表が使われ、⁽⁹⁾例えば100種類のCVC音節から成り立つものや、CVおよびVCがそれぞれ50種類含むハーバード大学形式のものなどがあるため、海外の測定結果と対比するときは言語の表音形式のちがいがいも含めて注意しなくてはならない。この明りょう度の測定には熟練した試験員が必要であり、練習の効果についても報告されているが、⁽²⁶⁾測定、整理に時間を要するのが欠点である。

一方平均オピニオン値⁽²⁹⁾は通話者が通話の良さについての満足度を主観的な判断によって数値化する方法であり、被験者は用意されているいくつかの評価段階に投票する。この結果から各段階に与えられた評点を重平均値を求める試験方法である。この方法は通話の良さを直接定量化する点では他の試験法に比してすぐれているものの結果に人的要素が混入し再現性に乏しいとされているが、非定常雑音の妨害等で明りょう度によって適切な評価をしがたいときに用いられる。

さて心理測定の一つに恒常法といわれる測定法があり、⁽²⁷⁾これは他の調整法や極限法などの方法に比べて最も正確かつ適用範囲が広いとされている。この方法はまず用意された一組の変化刺激を単独もしくは基準になる刺激とともに呈示し、被験者は各呈示毎に与えられた判断基準にしたがってまたは一どちらかで応答させる。この結果から刺激の大きさについて累積頻度を求めるとこれは精神測定関数と呼ばれる正規分布の累積分布関数になって、この曲線から通常累積分布が0.5となる刺激を閾値としている。恒常法の長所は閾値の決定が正確であり、計算の手続きが洗練されていることその他に、判断が比較的簡単にできるため被験者にとって好ましい方法であるが、反面測定誤差の決定方法が十分満足すべきものでない短所も持ち合わせている。聴覚に関係した各種の測定についても恒常法によるものが多く、検知限などはこの方法によっている。第3章では恒常法によって残響のプレファレンスを定量化する手法を導入し、通話についての好ましさを観点から考察を行う。

さて前節で述べた妨害音声による通話品質の評価は本来両耳受聴をもとに測定を行うことが望ましいが、本論文では基本的な性質を把握するため、片耳受聴に限定して実験を行った。これは妨害が通話側に存在し、音声モノホニック線路を経由して片耳受聴する場合に相当し、また両耳聴機能は片耳受聴が基本となると考えられるためである。⁽²⁸⁾

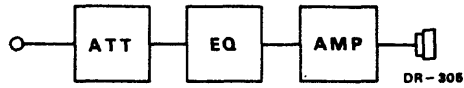


図 2.2 受聴装置

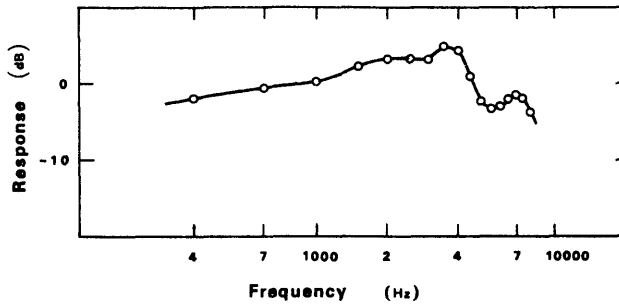


図 2.3 受聴装置のレスポンス

本研究に用いた受話系の構成は図 2.2 に示されており、標準受話器 DR - 305 による片耳受聴が同時に 4 名まで行える。この受話系の総合特性は受話系入力電圧を e_R とし、受話器を耳にあてたときのラウドネス感覚に等しいラウドネスを音場で誘起するような自由音場電圧の値を R_f としたとき $20 \log_{10} R_f / e_R$ で定められ、⁽²⁹⁾ 本研究におけるこの特性を図 2.3 に示す。この総合特性が平坦であることは受話者が音場で空間系を通して受聴する場合の音響効果と同一であることを示し、本研究で用いた装置はこのレスポンス偏差が約 9 dB 以内である。また受聴試験は通常の実験室で外来騒音の少ない時間を選んで実験を行ったが約 35 ~ 45 dB の暗騒音が存在している。

2.4 ま と め

以上の議論で明らかになったように本論文で扱う雑音とはすべて音響伝達経路に混入する音声信号であって、残響妨害とはこの妨害音声は通信の目的である音声信号と高い相関を有する場合といえる。他方通話内容と独立な音声による通話妨害も考えられ、この 2 点について以後考察をすすめる。

また通話品質を表わすいくつかの尺度についても触れ、更に受聴実験に用いる受話系の特性を示した。

第3章 通話品質に及ぼす残響の主観的效果^(30~32)

3.1 はし が き

残響の性質を表現する方法の一つとして残響時間が広く用いられているがその主観的な影響については多くの研究をみない。ここでは残響が付加された音声の品質をプレファレンスを用いて物理的な尺度に変換する方法をもとに、遅延素子と帰還による人工残響音について音声通信、とりわけ電話伝送の立場から品質評価を行い、遅延時間による規準化によってこれら人工残響音の性質を規準化残響時間によって対応づける。また若干の室内残響音についての測定結果から30msの遅延時間で定められる人工残響音がプレファレンスの点で室内における残響音と類似の性質を持つことを述べる。更にプレファレンスの観点から残響時間を信号対雑音比および%ディスタージョンの点から考察する。

3.2 残響付加方式

一般に残響付加方式は鉄板やスプリングなどの機械振動を利用するものと、遅延素子を帰還回路に組み合わせる電子回路方式に大別することができる。前者は電気信号を一たん鉄板やスプリングの機械振動に変換し、その自由振動を再び電氣的にピックアップする方式であるが、この場合に得られる残響音の特性を詳しく表現することは一般に困難である⁽³³⁾。一方遅延と帰還による後者の残響付加はその応答も単純であり、最も簡単な残響現象を模擬したものと考えられ残響の性質も明らかにできる場合が多い。これらの理由によってここでは後者の方法、つまり遅延と帰還による人工残響音を用いて実験をすすめる。

さてここで用いる遅延素子としては磁気録音機における録音及び再生ヘッド間のテープ走行時間差を利用したものや、⁽³⁴⁾ BBD素子を用いるものなどいくつか考えられるが、ここでは遅延時間の変更が容易であることから遅延素子はデジタル計算機とAD、DA変換器を組み合わせたものによって、またその他の部分はアナログ計算機によって構成する。遅延と帰還による残響付加方式のうち最も簡単なものは図3.1に示す回路で表わすことができる。ここで τ は利得が1の遅延素子であり、 g は利得を表わしている($|g| < 1$)。

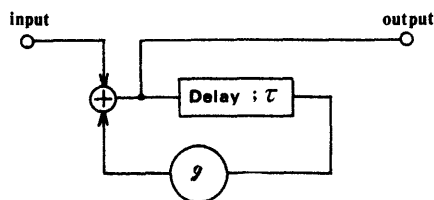


図3.1 簡単な人工残響付加回路

さて残響音の特性を示すパラメータのうち最も一般的なものは残響時間⁽¹⁴⁾であり、これを記号 T_r で表わせば図 3.1 において帰還ループの利得 g 、及び遅延時間 τ によって

$$T_r = \frac{-3 \tau}{\log_{10} g} \quad (3.1)$$

で与えることができ g 及び τ によって T_r を独立に変化させることが可能である。一方この回路の周波数特性を $H_1(\omega)$ とすれば

$$|H_1(\omega)| = \frac{1}{\sqrt{1+g^2 - 2g \cos \omega \tau}}$$

の関係があり、⁽³⁵⁾ 周波数特性が $1/\tau$ の間隔で規則的に変動するいわゆるくし形特性を示す。ここで変動の大きさ d は

$$d = 20 \log_{10} \frac{1+g}{1-g} \quad (\text{dB}) \quad (3.2)$$

となるため図 3.1 の回路では $H_1(\omega)$ を固定しつつ残響時間 T_r を変化することができない。いいかえれば残響時間の変化に伴って周波数特性も変化する。この欠点は人工残響音の主観的な評価を行う場合好ましいことではない。このため Schroeder は前記の欠点を取り除いた残響付加回路を提案している。^(35~37) 図 3.2 はそのブロック図であり、 g を利得とすれば直接音、反射音に対応する信号をそれぞれ $-g$ 、 $1-g^2$ の比で出力に導びくときこの回路が示す周波数特性 $H_2(\omega)$ は

$$|H_2(\omega)| = 1 \quad (3.3)$$

となり定常信号に対しては理想的な残響付加が行われる。そして遅延素子による遅延時間を τ とすればこの場合も残響時間 T_r は式 (3.1) で表現することができる。

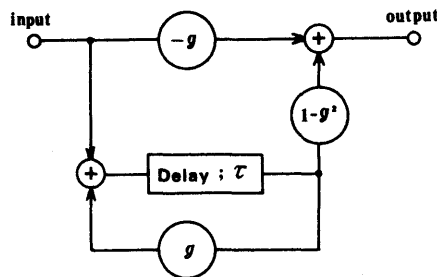


図 3.2 Schroeder の提案した人工残響付加回路

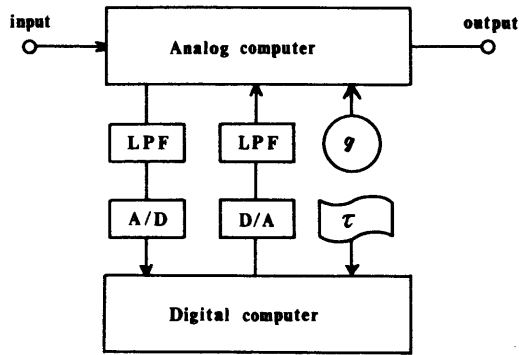


図 3.3 人工残響付加回路の構成

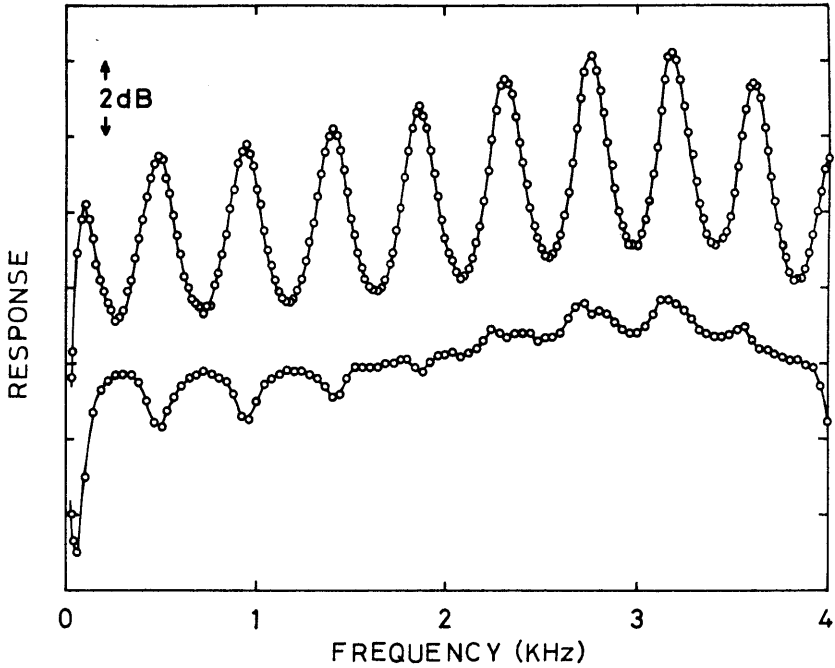


図 3.4 図 3.1 (上段) 及び図 3.2 (下段) による残響付加回路の特性

さてこのような残響付加回路は図 3.3 に示す構成によって実現した。A/D 変換器は $50 \mu\text{s}$ の間隔で 12 ビットのデジタル量に符号化し、あらかじめ紙テープにより読み込まれた遅延時間 τ に相当する時間を経たのち再びアナログ信号にもどされる。一方アナログ計算機では利得 g などの係数を定めており、デジタル的な遅延素子とともに残響付加系を構成している。

図 3.1 および 3.2 に示した残響付加系をこの回路によって実際に構成して周波数特性を測定した

結果を図 3.4 に示す。いずれの場合も $g = 0.3$, $\tau = 2.2 \text{ ms}$ であって上段に示される単純な残響付加系では式 (3.2) より約 5.2 dB の変動が規則的に現われている。また Schroeder による残響付加回路において式 (3.3) が成立するためには回路に含まれる各素子が周波数に無関係な定数となる条件が必要であり、そのうえ回路が正帰還になることから各素子の誤差が拡大されて出力側に影響を与えるため、素子の選択には十分注意する必要がある。⁽³⁸⁾ ここで用いた Schroeder による残響付加は図 3.4 の下段に示したように周波数変動を $\pm 2 \text{ dB}$ 以内に抑えることができた。

3.3 人工残響音のプレファレンス⁽³⁹⁾

3.3.1 プレファレンス測定法

さてこのように Schroeder による残響付加方式によって人工的に残響付加回路を構成したが、ここでは音声通信、特に電話伝送における残響の影響を調査することが目的であるため、防音室内において女性の短文章朗読を録音したものを原音声とし、この音声に種々の残響を付加した。受聴テストは 2 章で述べた恒常法の応用である一対比較法により行い、次の試験音を用いた。

- (A) 原音声に白雑音を混入したもの。
- (B) 原音声に残響付加を行ったもの。

図 3.5 に一対の呈示信号を示す。試験音 (A) については原音声に対する白雑音のレベル、つまり SN 比を 8 段階ランダムに変化し、それぞれを試験音 (B) を組み合わせ、更に順序効果をなくすため (A), (B) を入れ換えたものを含む合計 16 対の試験音を用意し、受聴者に呈示する。この場合試験音 (B) の残響付加はすべて同一のパラメータ (g と τ) によって固定されている。受聴テストにおける受聴者数は 4 名で各 10 回のテストを行った。受聴者はいずれもこれらの受聴テストには習熟している男子学生および男女職員であり、呈示された一対の試験音について

「どちらが通話に好ましいか」

の判断をする。これらの試験音声は受聴装置で 300 ~ 3400 Hz の電話帯域に周波数制限され、約 70 dB で受聴者に与えられる。

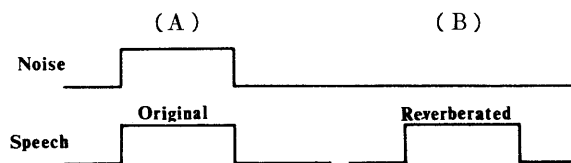


図 3.5 一対比較テストに用いた信号

このような受聴テストによって得られた結果を試験音 (A) の SN 比を横軸に、また残響音 (B) を好むと答えた割合を縦軸にプロットした一例を図 3.6 に示す。ここで試験音 (B) を好む割合が

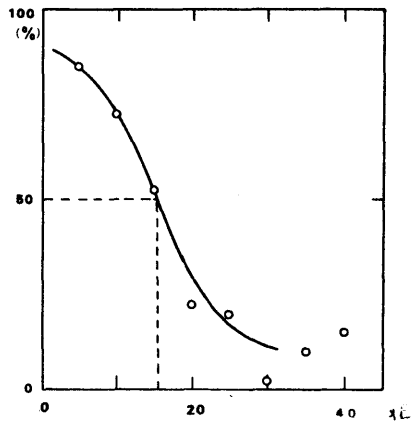


図 3.6 残響音を好む割合を S/N の関係についての一例

50%となるSN比は試験音(A), (B)を等しい割合で選ぶことを示しており, この点で定められる試験音(A)のSN比は試験音(B)の持つ残響特性とプレファレンスが等しいといえる。⁽⁴⁰⁾このようにして残響音をプレファレンスが等しいSN比に置き換え, この値を残響音と等プレファレンスとなるSN比と定めればこの値は残響音の主観的な性質を反映した一つの尺度となる。

一方SN比は明りょう度や各種のオピニオンテストにおいて多く用いられるパラメータであり, このような点から残響音は等プレファレンスとなるSN比を介して他の尺度に換算することも可能であると考えられる。

このように主観的な評価が困難であった残響音をプレファレンステストにより, 容易に扱える物理尺度の一つである信号対雑音比に置きかえることが可能であることから, 次にいくつかの人工残響音について前述の受聴テストを行った。遅延時間 τ は12.5msから200msまで対数的に選り, 利得 g は0.1から0.5までの値を用いたが, τ の値によっては更に多くの g についても実験を繰り返した。

このように一連のプレファレンステストから等プレファレンスとなるSN比を求めることができるが主観的な残響の性質は単一エコーの場合もそうであったように利得 g のみでなく遅延時間 τ も考慮した尺度つまり時間減衰特性によって左右されると思われるため従来より室内音響の尺度として広く用いられている残響時間, すなわち式(3.1)で定められた T_r によって結果を表わしたものが図3.7である。なおこれらの結果はすべて40回の受聴結果をもとにして得られた値であって平均の標準誤差は利得 g が0.2~0.3で最大となり約2.5dBであった。

図3.7から残響時間が増加するにつれて等プレファレンス S/N で定められる品質は低下してゆき, ある一定値に漸近する傾向がうかがえる。そしてこれらの結果は遅延時間 τ によって分布して

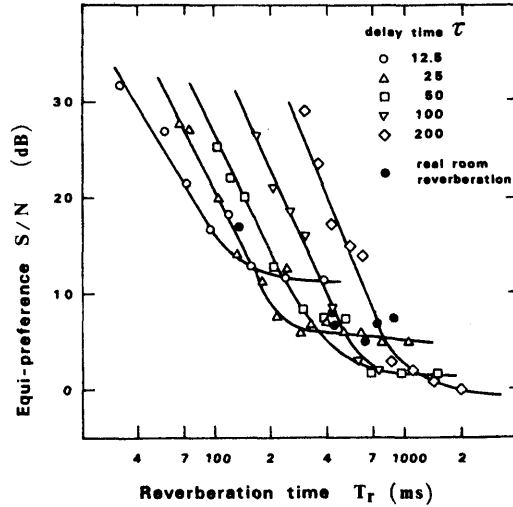


図 3.7 等プレファレンス S/N と残響時間の関係

いるが S/N 比と τ の関係はすべて同様の傾向を示しており、そのため遅延時間 τ と曲線間の差を求めれば一つの曲線で代表させることが可能であると思われる。

一方受聴テストに用いた原音声を通常の室内において残響効果を付加したものについて、人工残響音と同様のプレファレンステストにより S/N 比に変換した結果も図 3.7 に併せて示した。これらの残響時間は 0.5, 0.7, 1, 1.5, 2, 3 および 4 KHz における減衰波形から 60 dB の減衰に要する時間をそれぞれ求めて平均した値を用いた。⁽¹¹⁾ これらの受聴結果を Schroeder 方式による人工残響音による結果と比較した場合、遅延時間にしてはほぼ 25 ~ 50 ms に位置するものと推定される。

さて人工残響音の等プレファレンス S/N が遅延時間 τ によって移動することは次のように考えられる。

すなわち遅延と帰還によってつくられた人工残響音の主観評価において規則的な時間差を持つ信号が聴覚に対して同時に作用することにより、残響時間で定められる以外に遅延時間 τ の効果がプレファレンステストに大きく影響していると思われる。そして図 3.7 において遅延時間 τ が等プレファレンス S/N に与える影響を推定すればそれは

S/N 方向の移動

T_r 方向の移動

の 2 種類の効果に分割して考えることができ、これらの点について検討を加える。

3.3.2 信号対雑音比の規準化

図 3.8 に示したものは τ が 12.5 ms の場合を規準に S/N 方向の差を残響時間が大なる部分よ

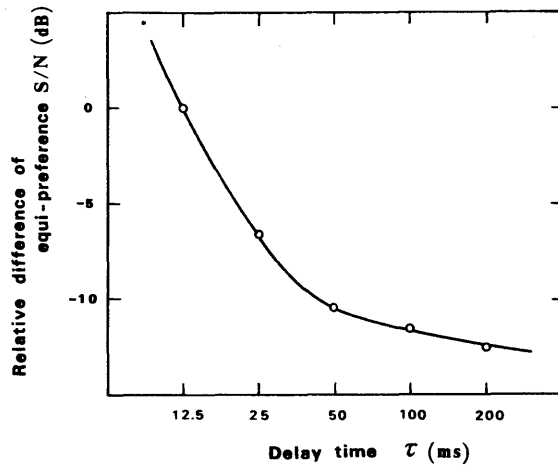


図3.8 $\tau = 12.5$ ms に対して求めた S/N の差

り求めた結果であり、人工残響音が示す品質の差を表現したものといえる。

さて Lochner ら⁽²⁾ によるスピーカを用いた単一エコー妨害を測定した結果では 30 ms 以内の遅延信号は単に 3 dB 法則で加算することができ、聴覚的にはこの場合明りょう度の向上に貢献することが報告されている。一方 Haas による同様の研究⁽¹⁾においても遅延時間が約 30 ms 以内のエコーはラウドネスを増加させ、聴感上こちよい印象を与えると述べており、また 50 ms 以上遅れた音を伴う場合は分離した感じを与えるとともに妨害効果を生みだすことが明らかにされている。これらの報告をもとにここでは Lochner が提唱したように単一エコーを信号成分と雑音成分に分割する考え方を導入し、残響音の品質を SN 比に関連づければ、30 ~ 50 ms 以内の反射音は SN 比を上昇させ、反面遅れの大きいエコーは SN 比の減少につながり品質の劣化をまねくものと考えられる。ここでは単一エコーにおける妨害についての結果を考慮して SN 比の移動を示す図 3.8 の曲線を改め、図 3.9 で示されるように遅延時間にして 30 ms における値を 0 dB に置いた曲線によって規準化を行えば異った遅延時間で定められる人工残響音が潜在的に含んでいる SN 比の増減を明確にすることができ、図 3.7 における縦方向の分布に関して遅延時間が異なることに起因する影響を補正できると考えられる。

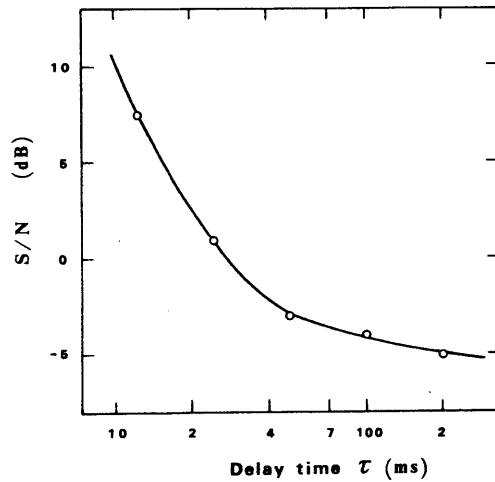


図 3.9 信号対雑音比の規準化曲線

3.3.3 残響時間の規準化

図 3.7 における人工残響音についての残響時間 T_R は式 (3.1) から求めた値であるが、遅延時間の異なる人工残響音について更に残響時間方向に分離する傾向が見られることはここでも規則的な繰り返し間隔 τ によって主観的な判断が単に残響時間のみでなく τ を含む項が式 (3.1) につけ加えられた値に依存することを意味している。ここで用いたような規則的な繰り返しのある人工残響音の場合、数学的に定められる残響時間と主観に作用する残響感の間に差があることから、次式で定められる T_R^* を帰還方式による人工残響音に関しての規準化残響時間として導入する。

$$T_R^* = k(\tau) \cdot T_R \quad (3.4)$$

さてこの場合も図 3.7 に見られる室内残響音のプレファレンスの性質から遅延時間が約 25 ms の場合に相当していること、ならびに信号対雑音比の補正に用いた 30 ms の規準化と考え合わせ、ここでも遅延時間が 30 ms に対して規準化を行い

$$k(30 \text{ ms}) = 1$$

に選ぶ。このとき $k(\tau)$ は図 3.10 に示される値をとり、遅延時間 τ が大きくなるに従って規準化残響時間 T_R^* は残響時間に比べてより小さい値となることを示しており、 $k(\tau)$ の実験式は

$$k(\tau) = 0.046 \tau^{-0.88} \quad (3.5)$$

となる。従って規準化残響時間は式 (3.4)、(3.5) より

$$T_r^* = \frac{-0.14 \tau^{0.12}}{\log_{10} g} \quad (3.6)$$

によって実験的に表わすことができる。

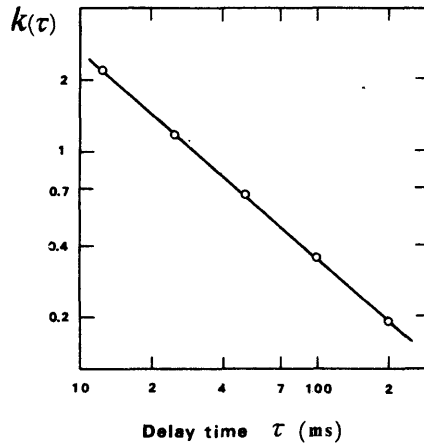


図 3.10 $k(\tau)$ と遅延時間 τ の関係

3.4 残響時間によるプレファレンスの表現

さて式(3.6)および図3.9を用いて図3.7を30msの遅延時間について規準化を施すことにより人工残響音の遅延時間が異なることによって現われた影響を除くことができる。図3.11は人工残響音についてこの規準化を行った結果をプロットしたものであり、室内残響音の示す等プレファレンス S/N とほぼ同様の傾向がみられ、3.3.2, 3.3.3で用いた30msの規準化があながち否定されるとはいえない。この結果より規準化残響時間にして約220msの残響音が一つの境界値と考えられ約7.5dBの等プレファレンス S/N に相当しているがこれを別の資料によって考察する。

図3.12は受聴レベルが70dBにおいてSN比を変化させたとき、A:非常によい、B:よい、C:まあまあ、D:悪い、E:非常に悪い、の5段階のオピニオンについてそれぞれ4, 3, 2, 1, 0の重みをつけて平均したときの結果であり、測定されたオピニオン値とSN比の関係を表わしている。⁽⁴²⁾ このときSN比が約7.5dBの音声ではオピニオン値の平均が約2.4である。一方このオピニオンテストにおいて5段階評価のうち、D, Eを選ぶ割合が10%のとき平均オピニオンが2.4となることから、⁽⁴²⁾ 7.5dBのSN比は10%の受聴者が「悪い」という判断、つまり妨害を感じるものであるといえる。このため図3.11における境界値のSN比は7.5dBで

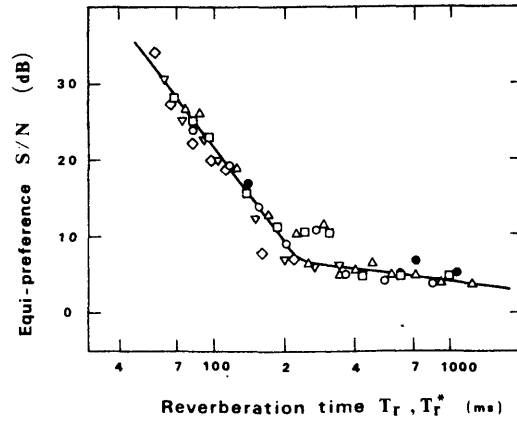


図 3.11 実験結果の規準化表示

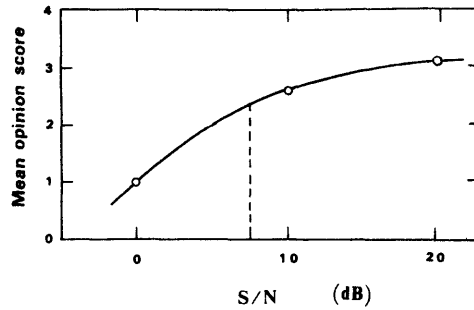


図 3.12 S/Nと平均オピニオンの関係

示されることから、この点で定められた規準化残響時間が約 220 ms の残響音は約 10% の受聴者が妨害を受けたと感じる場合であると推定される。

従って残響音の限界値は規準化残響時間にして約 220 ms、また等プレファレンスとなる SN 比では約 7.5 dB であり、妨害を感じる程度は受聴者の約 10% であると結論されるが、更に Bolt らによる % ディスタバンス⁽⁴³⁾との関係について考察する。

図 3.11 から判断された境界値における規準化残響時間 T_r^* は遅延時間 τ の選び方にかかわらず約 220 ms であるが、単一エコーの場合と同様に各遅延時間に対する利得 g を求め、この関係を図 3.13 に実線で示す。この関係で定められる人工残響音は図 3.12 によって 10% の受聴者が妨害を感じる程度であると述べたが、Bolt らによる単一エコーの 10% ディスタバンスを定めた破線と同様の傾向を示している。このように単一エコー及び人工残響音について 10% の妨害を

示す関係に顕著な差が見られないことは人工残響音の初期遅延信号が主観に大きく作用しているためと考えられ、主観的な残響感が減衰を定めている初期傾斜に強く依存するという報告⁽⁴⁴⁾と矛盾しない。

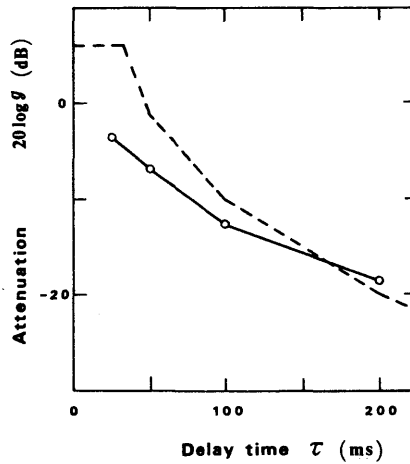


図 3.13 遅延時間と利得 g の関係。
破線は Bolt らによる 10%
ディスタージョンス曲線を表わす。

ここで論じた単一エコー及び人工残響音について 10% の妨害が類似の条件で与えられることは山本の指摘したとおり 10% ディスタージョンスが一つの検知限としての性質を備えていると考えられる⁽⁴⁵⁾。

更に山本⁽⁴⁶⁾の調査によれば放送アナウンススタジオとして許容される残響時間は約 250 ms 以内という結果が報じられており、ここで電話伝送を対象として実験を行った結果、約 220 ms と推定した境界値と非常に近いことから人工残響音を規準化することによって残響音の尺度をより密接に主観と対応させることができると思われる。

このように遅延素子と帰還による人工残響付加音声についてインパルスレスポンスの規則性に起因する聴覚的な変動をとりあげ、規準化を施すことによって人工残響音についての実験結果を通常用いられている残響時間の尺度にあてはめることができ、また実験結果より得られた境界値は単一エコーおよび通常の残響音場において妨害を推定した条件と同様の傾向を認めることができ、ここで用いた規準化の有効性がうかがえる。

3.5 ま と め

以上のことをまとめると次のようになる。

まず従来主観評価が困難であった音声通信系における残響音を等プレファレンスとなるSN比に置き換えることにより尺度化できることを示した。

次にこの方法を用いて Schroeder 方式による人工残響音の品質をプレファレンスの観点から測定した結果、等プレファレンスとなるSN比の曲線は遅延時間にして約30 ms を中心に分布するものと思われ、これは規則的な遅延が聴覚に作用することに起因すると考えた。このため人工残響音の通話における感覚的な評価は残響時間にある係数を乗じた規準化残響時間によって対応づけることができ、また単一エコーと同様に時間差が30 ms 以内の遅延信号群は負の妨害、つまり品質の上昇をもたらす傾向を推定した。

従って遅延時間が30 ms 付近に設定された人工残響音および室内における残響音についてそれぞれ電話伝送系を通過したものはプレファレンスの点で非常に似ており、電話系を対象とした室内残響の効果を単純に模擬する場合の一条件となろう。

またこれら電話伝送系に混入する残響音は残響時間にして約220 ms において一つの境界値があり、この点でプレファレンスを介して求めた等価SN比は約7.5 dB に相当している。そしてこの値は10%の受聴者が妨害を感じる場合であることが裏付けられ、信号間の時間差とレベル差の点では単一エコーと同様の傾向を示すことが確かめられた。

第4章 残響の主観的評価量と物理量の対応^(47,48)

4.1 はし が き

前章では残響が通話に与える影響を「通話にとっての好ましき」の面から追求し、プレファレンスが等しい信号対雑音比の値に置きかえて評価を試みたが本章では更にこの信号対雑音比が室内音響で用いられているD値を変形したもの、すなわち約40～50msでインパルスレスポンスを二分したときの間接者エネルギー比に相当していることを見いだすとともにこれを実験的に検証する。その結果、プレファレンスの観点からあたかも直接音およびすべての間接音が聴覚上有効な信号成分として作用し、同時に直接音から約40～50ms以上の遅れを伴う部分は雑音としての妨害作用をもつことを述べる。

4.2 残響特性の評価^(49,50)

残響の性質を表現する物理量の一つである残響時間は室内音響の尺度として広く用いられているが、ここでは3章で述べた受聴テストの結果と残響時間の関係について更に検討する。

さて前章では遅延素子を含む回路により人工的に合成された規則的な残響音の主観評価は遅延時間によって異った傾向を示した。この原因はわれわれの感覚が日常経験する複雑な残響音に順応しており、ここで用いた人工的な残響付加回路のように単純化した残響音を用いた評価では聴覚上特異な残響としての印象を受けるためと思われる。そしてこれらの品質が本来残響時間と結びつけられるとしたとき、この特異な印象の補正を遅延時間が30msの場合に読みかえることを規準化として導入し評価することを前章で試みた。これらの実験結果を規準化したものは図4.1に示されるようになりに強い直線関係が認められ、更にこの図に併記した若干の室内残響による等プレファレンスSN比についてもほぼ同様の傾向を認めることができる。このため電話伝送系を通過した室内残響音が示すと思われる等プレファレンスSN比を規準化された人工残響音による等プレファレンスSN比で代表させて議論する。

さて図4.1を詳細に観察すれば残響時間が約200msまでは残響時間の増加にともなって等プレファレンスSN比がほぼ直線的に低下してゆくが、残響時間が約300ms以上、1sec程度までは品質の劣化する割合がより小さいと思われる。前章で述べたように約220msを境界値と考えたが、等プレファレンスSN比を P (dB)で表わしたとき残響時間が200ms以下について P と T_R の関係を最小二乗法により求めてみると次のようになる。

$$P = 2.19 / T_R^{0.97} \quad [dB] \quad (4.1)$$

つまり残響時間 T_R が200ms以下では等プレファレンスSN比である P はほぼ残響時間に反比

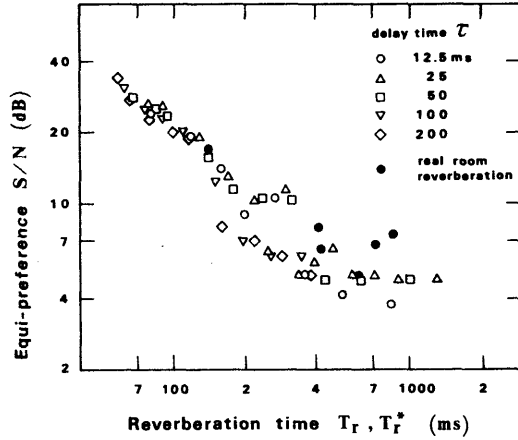


図 4.1 実験結果の規準化表示

例して減少する。ちなみに T_r が 250 ms, および 300 ms 以下の実験結果によっても式 (4.1) の分母はほとんど差がなくそれぞれ $T_r^{1.02}$ および $T_r^{0.97}$ であるため T_r が 300 ms 近くまで式 (4.1) の関係を認めることができる。このように残響による電話音声の品質を表わす尺度として導入した等プレファレンス S/N 比が残響の物理特性の一つである残響時間と反比例することは興味深い。

そのため残響音声の主観を反映した値 P と残響時間 T_r の関係を以後簡単のため次式のように考えて議論を進める。

$$P = 2.2 / T_r \quad [\text{dB}] \quad (4.2)$$

4.3 プレファレンスとインパルスレスポンスとの関係

前節の議論から残響音声の主観を反映した品質 P と残響時間 T_r は比較的単純な関係で結びつけられることから次に残響時間以外の物理尺度との関係について考察する。

通常の室内において時刻 $t = 0$ で音源を停止した場合、室内の音響エネルギーを $I(t)$ で表わせば良く知られているように室内の拡散性が良い場合に $I(t)$ は

$$I(t) = A e^{-\alpha t} \quad (4.3)$$

によって代表させることができる。⁽⁵²⁾ このように室内における残響音が平均的に指数減衰すると考えられる場合残響時間 T_r は $T_r = 13.8 / \alpha$ で与えることができ、この関係を用いて式 (4.2) を書きかえると次のようになる。

$$\begin{aligned}
 P &= 0.159 \alpha \\
 &= 10 \log_{10} e^{0.037\alpha} \quad [\text{dB}]
 \end{aligned} \tag{4.4}$$

つまり等プレファレンス SN 比は音響エネルギーの時間的減衰を示す定数 α による関数 $e^{0.037\alpha}$ を dB 表示したものによって計算できることがわかる。

さて室内の音響指標として用いられるものの一つに Mayer らの提案による D 値 (Definition) があり、⁽⁷⁾ 次式で定義されている。

$$D = \int_0^{50 \text{ ms}} I(t) dt / \int_0^{\infty} I(t) dt \tag{4.5}$$

これは音源停止後 50 ms 以内の音を直接音とみなしたとき、直接音のエネルギー比を示している。したがって音源停止後の室内の音響エネルギー $I(t)$ が式 (4.3) で与えられるように指数減衰し、かつ直接音として考慮する時間を t_1 sec としたときの D 値を $D(t_1)$ で表わせば次の関係が得られる。

$$1 - D(t_1) = e^{-\alpha t_1} \tag{4.6}$$

式 (4.6) における左辺は t_1 sec 以後を間接音と考えたときそのエネルギー比を表わしており、これが右辺のように指数関数で表現できることから式 (4.4) は更に次のように書きかえることができる。

$$P = -10 \log_{10} (1 - D(37 \text{ ms})) \tag{4.7}$$

つまり等プレファレンス SN 比である P は変形された D 値である $D(37 \text{ ms})$ を用いて表現することが可能である。したがって式 (4.7) から時間にして約 40 ms で二分したときの間接音パワ比が大きくなるほど等プレファレンスとなる信号対雑音比は小さくなるといえる。

このようにプレファレンスを介して残響音声の通話への好ましさを測定した結果、間接音成分が多い残響をとまなう音声は通話にとって望ましくないことがわかる。そして受聴者が好ましさの面で許容できる遅延信号には限度がありその値は以前から言われているように 30 ~ 50 ms 程度であることが等プレファレンスによる S/N 評価についてもあてはまる。

更に式 (4.7) は $I(t)$ を用いて次のように変形される。

$$P = 10 \log_{10} \left\{ \int_0^{\infty} I(t) dt / \int_{37 \text{ ms}}^{\infty} I(t) dt \right\} \quad [\text{dB}] \tag{4.8}$$

このため等プレファレンス SN 比である P は残響音全体を信号成分とし、インパルスレスポンスにおける 37 ms 以後の成分を雑音成分と考えたときの信号対雑音比に帰着される。

従ってプレファレンスの観点から求めた残響付加音声の電話伝送における品質 P が室内音響学で用いられている D 値の変形された値、つまり $D(37 \text{ ms})$ と実験的に結びつけることができる。

ここで導びいた関係は残響時間にして約200～300ms以下の場合に有効であり、かつ電話伝送系に限定されたものであるが、このようなプレファレンス評価とD値の関連性は非常に密接である。それは音の歯切れの良さ、明りょうさなどはほぼD値に相当すると⁽⁵³⁾といわれており音声通信に先ず要求される歯切れの良さ、明りょうの観点からD値つまり直接音としてのエネルギー比と対応関係があると思われる。

式(4.8)で表わされたように約40ms以上の遅延信号、すなわちエコーを含む場合に等プレファレンスで定められる通話の品質が劣化することを実験的に裏付ける目的で次節では多数のエコーを付加した音声を用いプレファレンステストによって検証する。

4.4 多重エコーを伴う音声のプレファレンス⁽⁵¹⁾

前節で述べたように残響音パワが平均的に指数曲線に沿って減衰すると仮定し、電話伝送系について求めた品質は約30～50msで二分されたインパルスレスポンスと関係があると思われ、最終的に式(4.8)に帰着されると述べた。ここでは式(4.8)で用いた積分区間を検証する目的で多重エコーを付加した音声を用いてその通話品質を求める。実験方法は3章で述べたものと同じ手法で行ったが、人工残響音にかえて多重エコーを付加した音声を用いた。この多重エコーの構造は遅延信号の大きさが時間とともに指数関数的に減衰し、かつある時間以後の成分が0であるようなエコー群を用いるべきと思われるが、ここでは簡単のため指数減衰がない場合について実験を行った。使用した試験音のエコー付加は図4.2に示すように原音声を遅延させ、単一エコーを付加する

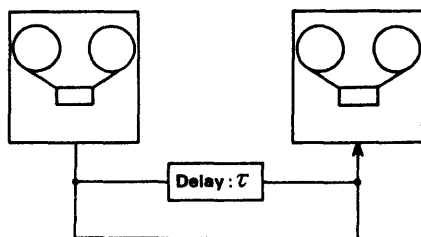


図 4.2 単一エコーの付加

操作を遅延時間を半減しつくりかえし、多数のエコーが付加された試験音を作成した。この試験音は図4.3に示されるような構造であり、直接音に相当する信号からそれぞれ τ の整数倍の遅延を受けた合計 n 個の信号の和であり、 n 個のエコーが付加されたことに相当する。ここでエコーのレベルはすべて直接音に相当する信号と同一に保たれており、遅延時間 τ およびエコーの個数 n は表4.1に示した値を用いて実験を行った。

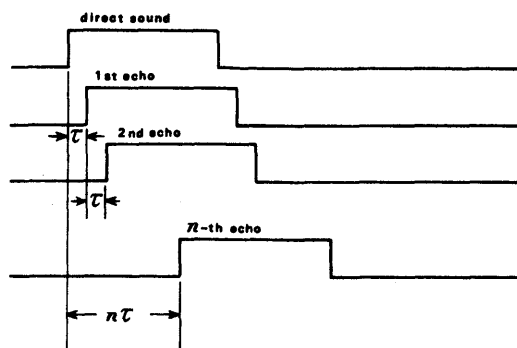


図 4.3 マルチエコーの時間構造。すべてのエコーレベルは直接音のレベルと同一である。

τ (ms)	n
4	1
10	3
12.5	5
20	7
25	15
50	31
100	
200	

表 4.1 実験に用いられた遅延時間 τ 及びエコーの個数 n の値

さて受聴テストはこのようなエコーを含む音声（白雑音を加えたもの）と図 4.4 のように一組にして受聴者に呈示する。受聴者は 3 章で述べたように

「どちらが通話に好ましいか」

の判断をし、プレファレンスが等しい信号対雑音比を合計 40 回の受聴テストの結果から求めた。受聴者はこのようなテストに習熟しているそれぞれ 2 名の男子学生および男子職員であり、その他の受聴条件は 3 章で用いたものと同様であって受聴レベルはここでも約 70 dB である。

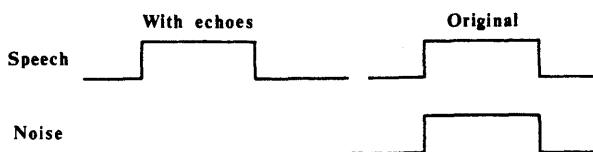


図 4.4 呈示信号の構成

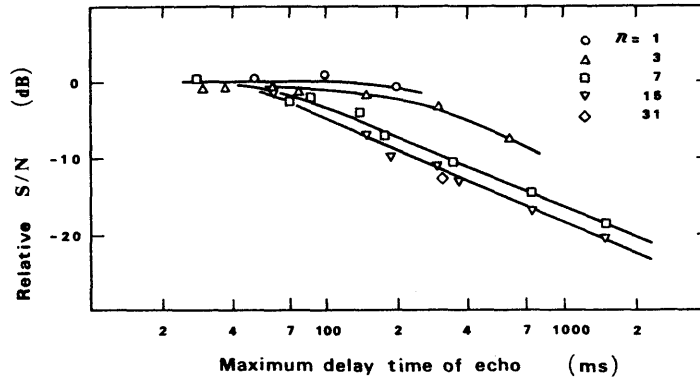


図 4.5 等プレファレンス S/N とエコーの最大遅延時間の関係

このようなエコーが多数付加された音声の品質について受聴結果からプレファレンスが等しい信号対雑音比を求めた結果を図 4.5 に示す。横軸は $n\tau$ 、すなわちエコーの最大遅延時間に相当する値である。この結果から $n\tau$ が約 40 ms 以下ではエコーを付加したことによる品質の変化はほとんど認められない。しかしエコーの最大遅延時間 $n\tau$ が 40 ~ 50 ms を越えると徐々に品質が劣化しはじめる。ところが $n=1$ の場合、すなわち単一エコーの場合では品質の変化をほとんど認めることができない。単一エコーの検知限をスピーカを用いて測定した田原ら⁽⁵⁴⁾の結果によれば約 30 ms 以上のエコーは分離して検知されると述べており、これを考えあわせると図 4.5 の $n=1$ に相当する曲線の示す傾向は $n\tau$ が 30 ms 以上でエコーの検知はされているものの、通話における好ましきの点では顕著な影響は現われないものと考えられる。他方 n が 7 以上の場合は $n\tau$ が 50 ms 以上で急激に品質が劣化しはじめ、エコーによる妨害効果を認めることができる。つまり多くの反射音を伴う音声を受聴者に呈示された場合、反射音の最大遅延時間が約 40 ms 以内であれば通話の好ましきの点では影響が少なく、反面 50 ms 以上にわたる遅延を伴う反射音の存在は等プレファレンスとなる信号対雑音比でながめた場合にも明らかに品質低下をもたらすと解釈できる。

このようにプレファレンスを介して測定した品質の劣化がはじまる $n\tau$ の限界値は約 40 ~ 50 ms であることからこれらの値を上まわる遅延によって生じたエコーは n が大きい場合、妨害効果を伴うといえる。そのため従来から広くいわれているように^(1~4) 直接音から約 40 ~ 50 ms 以内の反射音は品質の劣化に結びつかないことがプレファレンスを介して測定した場合についてもあてはまり、約 40 ms 以上の遅延信号を妨害とする式 (4.8) の有効性を示唆するものといえる。

4.5 ま と め

このように人工残響音を用いて室内残響音をシミュレートし、プレファレンステストによって残

響音の品質を整理すれば、残響時間が約200ms以下で等プレファレンスとなる信号対雑音比は残響時間と直線的な関係となる。そしていくつかの仮定のもとではプレファレンスが等しい信号対雑音比はあたかもD値と同様にインパルスレスポンスを40ms近くで二分することにより示されることを述べた。

また直接音から約40～50ms以内の遅延音はすべて有効な信号エネルギーとして作用し、遅れが大なる音声信号は信号成分としての作用とともに雑音としての妨害効果も伴うと考え、これを多重エコーをともなう音声によって検証した。

これらのより詳細な議論は音声の時間的なレベル変動特性や母音、子音の継続時間の変化を組み入れたマスキング現象によって説明が可能と思われ、これらについては5章で検討を加える。

第5章 残響の明りょう度に及ぼす影響^(55,56)

5.1 はし が き

残響と明りょう度の関係を論じた研究として Bolt⁽⁵⁷⁾は音声信号を130msの active な部分と100msの silent な部分に分割し、明りょう度計算を行っている。また Janssen は50ms 経過時における残響特性をもとに明りょう度を予測しているが、⁽⁵⁸⁾本章では音声学の見地から連続音声を母音と子音に分類し、これらの平均的な性質を組み込んだ音声モデルをもとに残響が子音に及ぼす妨害効果について考察する。

通話の基本となる音声信号は発声器官の複雑な動きによってつくられるが、調音時の特徴やスペクトルの変化から母音と子音に大別される。そしてわれわれが日常用いる日本語は母音もしくは子音と母音によって大半の音節を構成しているため会話音声は子音と母音の繰り返しによって近似的に表現することが可能であろう。一方残響等遅延歪による音声への妨害を論ずる場合、これら子音、母音間の時間的性質も考慮することが必要と思われる。

このような観点から本章では会話音声の平均的な性質を調査し、連続音声のモデル化を試み、このモデルの有効性を確かめた上モデルにもとづく擬似音声簡単な残響を伴う場合の検知限から残響の限界の予備的な推定を行う。

次に実際の音声に指数減衰をとまう簡単な残響を付加し、通話品質を表わす尺度の1つである明りょう度によって残響の評価を行うとともに明りょう度計算によって通話の速度と明りょう度の関係についても考察する。

5.2 音声の時間的性質

音声は周知のように母音と子音から成り立っている。日本語の連続音声では母音または子音+母音の繰り返しかえしであって、これらの継続時間の性質は比企⁽⁵⁹⁾によって詳細な報告がなされており、例えば子音+*a*では平均132ms、母音*a*では平均75msとされている。また連続音声には種々の休止時間が存在するがその中で吸気を伴う休止もしくは構文間に出現する休止を除いた連続部分は近似的に子音と母音が交互に存在するものと見なすことができる。そしてこれら子音と母音は時間的な側面からそれぞれ平均52ms、75msの時間だけ存在し、これらが周期132msで繰り返されるものとして扱うことができよう。

一方これら子音、母音の継続時間を t_c 、 t_v とすればこれらは発話速度に依存しており

$$t_c + 2 t_v = m$$

$$m = 160 \sim 250 \text{ ms}$$

の関係があるとされている。⁽⁵⁹⁾ また音節中の平均継続時間は子音が約 100 ms , 母音が約 150 ms であろうと思われる。⁽⁶⁰⁾ このように音声の継続時間を母音, 子音に分割して考えたが, 次に音声のレベル変動も同様に母音と子音に分割し, その相対的なレベル変化から母音に対する子音の平均的なレベルについて考察する。

母音は声帯の振動による喉頭原音が声道の共振を受けて変化したものであり, 一方子音の音源としては呼気流が調音部のせばめを通過するときを生ずる乱流音源や, 呼気圧の開放による破裂性の音源などがあり, 大部分の子音は母音に比して過渡的であるためレベル測定には注意を要する。そのためあらかじめ録音された音声サンプルを $\frac{1}{2}$ のスピードで再生し, 記録速度を最大にした B & K 社のレベルレコーダに描かせることによりレベルの測定を行った。音声サンプルは男女各 1 名による無意味 3 音節の中央の音節から母音に対する子音レベルを合計 468 回測定した。しかし無声破裂音については測定が不可能であったため系のノイズレベル (平均母音レベルから約 -3.8 dB) で置きかえて整理した。

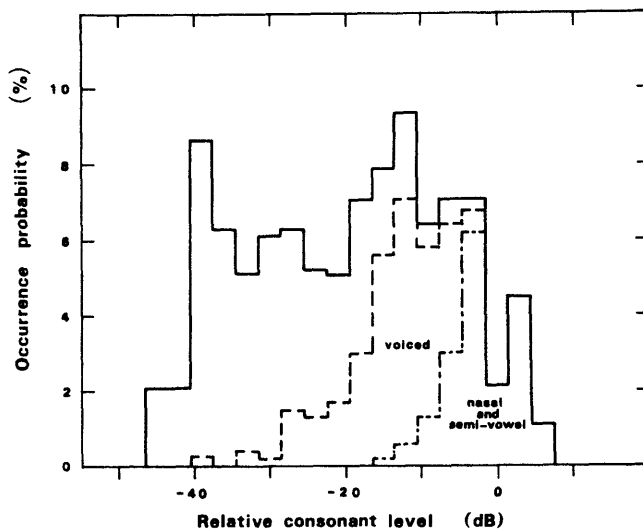


図 5.1 子音レベルの頻度分布

これらの測定結果を子音レベルを 3 dB ごとにまとめて頻度分布として表せば図 5.1 の実線のようにになる。ここで 0 dB は平均母音レベルを表わしており, すべての子音の平均は -19.9 dB となった。また有声子音のみについて分類すれば破線で表わされる分布となり平均は -9.4 dB , 更に鼻音, 半母音に限定した分布は鎖線のようになり平均 -5.2 dB である。つまり鼻音や半母音は母音より約 5 dB 低く, 有声子音中でもレベルが大きいグループを形成しており, 全子音の平均レベルは母音から約 20 dB 低いといえる。一方レベルレコーダで描かれた曲線から視察で求めた子音・母音の平均継続時間はそれぞれ 91 ms , 173 ms であった。

これらの結果から子音レベルについて主な結果を記せば次のようになる。

- 1) 子音の平均レベルは母音に比して約 - 20 dB である。
- 2) 子音をレベルの面から分類するとレベルの大きい方から順に次の3グループに分けることができる。
 - (1) 鼻音および半母音
 - (2) 有声子音から(1)を除いたもの
 - (3) 無声子音

したがって連続音声についてもこの関係が成り立つものとすれば継続時間と子音レベルの平均値から連続音声は図 5.2 のような構成として表現することができる。

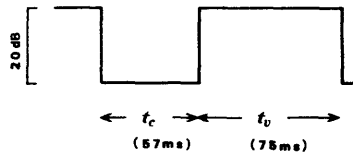


図 5.2 平均的な会話音声のレベル変化

さてこのように音声の子音，母音の継続時間とレベル比の平均値を用いて表現した場合，音声の物理的特性に矛盾が生じないことを確認しておかねばならない。音声の長時間統計量の一つとして良く知られているものに瞬時振巾の確率密度関数がある⁽⁶¹⁾これを $W(x)$ とすれば次式が比較的適合するといわれている。

$$W(x) = \frac{0.4}{\sqrt{2\pi} \cdot 0.118} e^{-\frac{1}{2} \left(\frac{x}{0.118}\right)^2} + \frac{0.6}{\sqrt{2} \cdot 1.23} e^{-\frac{\sqrt{2}|x|}{1.23}} \quad (5.1)$$

ここで x は音声の長時間実効値で正規化した瞬時振巾であり， $W(x)$ の第 1 項は子音に対応するガウス分布，また第 2 項は母音についての分布を定めた両側指数分布である。子音，および母音の実効値をそれぞれ σ_c ， σ_v とすれば

$$\begin{aligned} \sigma_c &= 0.118 \\ \sigma_v &= 1.23 \end{aligned}$$

であり，母音に対する子音レベルは

$$20 \log_{10} \frac{\sigma_c}{\sigma_v} = -20.3 \quad [\text{dB}]$$

である。これは米語についての平均的な関係を表わしているが日本語についての報告⁽⁶²⁾は約 - 2.8

dB であって今回求めたレベルは約 8 dB 大きい値を示している。この原因はノズルレベルを下まわる無声破裂音のレベルをノイズレベルで置きかえて評価したため録音時における暗騒音等の雑音によって平均子音レベルの上昇をもたらしたためと思われる。

連続音声のレベル変化がこのように約 20 dB のパルス状波形として表わすことができ、かつ子音・母音の瞬時振巾がそれぞれガウス分布、両側指数分布に従うと考えれば振巾の確率密度関数は次式で定めることができる。

$$W(x) = \left\{ \frac{t_c}{\sqrt{2\pi} \sigma_c} e^{-\left(\frac{x}{\sigma_c}\right)^2} + \frac{t_v}{\sqrt{2} \sigma_v} e^{-\frac{\sqrt{2}|x|}{\sigma_v}} \right\} / (t_c + t_v)$$

ただし t_c ; 子音の平均継続時間
 t_v ; 母音の平均継続時間
 σ_c ; 子音の平均レベル
 σ_v ; 母音の平均レベル
 x ; 長時間実効値で正規化した音声の瞬時振巾

ここで $t_c = 52 \text{ ms}$

$t_v = 75 \text{ ms}$

$$20 \log_{10} \frac{\sigma_c}{\sigma_v} = -20 \text{ [dB]}$$

とし、更に $W(x)$ の正規化条件を考慮すれば、

$$W(x) = \frac{0.43}{\sqrt{2\pi} 0.132} e^{-\left(\frac{x}{0.132}\right)^2} + \frac{0.57}{\sqrt{2} 1.32} e^{-\frac{\sqrt{2}|x|}{1.32}} \quad (5.2)$$

となる。

このように連続音声の子音、母音の継続時間とレベル比を組み込んだ瞬時振巾確率密度関数は、音声波形から直接求められた式(5.1)に近い結果を示していることがわかる。

したがって連続音声の子音および母音の平均的なレベルが図5.2のようなパルス状変化として扱うことができ、このモデルを用いて残響による妨害を考察する。

5.3 擬似音声による検知限測定

連続音声のレベルが時間および振巾の変化を定めることによってパルス状に変化するものとして扱うことができるため、この平均的な性質をそなえたモデルにもとづく擬似音声が残響付加系を通過した場合の妨害について考察する。

すでに述べたように平均子音レベルは母音に比較して約 20 dB 低いため連続音声に及ぼす残響妨害が最も顕著な場合として話者の発声が次の音節に移行した状態、つまり母音から子音へ移行し

た際、先行母音の残響成分が後続子音の聴取に与える妨害が考えられる。他に考えられる妨害としては先行母音の子音をはきんで後続母音に影響を及ぼす場合や、母音が連続する際先行母音が残響によって後続母音に与える影響等が考えられる。しかし母音のレベルは一般に子音のレベルより大であることから、また子音の継続時間が母音のそれに比して短いため、子音や母音の残響成分が後続母音を妨害することよりもレベルの小さい子音が母音によって妨害される方がより顕著であろうと考えられる。

このような観点から残響が付加された連続音声を通話に用いる場合に予想される妨害として以後先行母音の後続子音に及ぼす妨害効果を取りあげて考察する。そのため先ず母音、子音に相当する音声信号を電氣的に発生させ、5.2で述べた継続時間および振幅の変化を与えることによって連続音声の平均的性質をそなえた擬似音声を模擬することからはじめる。そして予備的実験としてこの擬似音声に残響が付加された場合を母音および子音の成分を指数的に減衰させることにより模擬し、子音に相当する成分についての検知限特性の変化から子音を受ける残響妨害の推定を試みる。

さてこの実験に用いる信号はそれぞれ母音、子音としての特徴をそなえた2種類の信号が必要とされる。音声の中でも有声音、特に母音はほぼ周期的な振動であるため、これを基本周波数 f_1 をもつ調和構造の複合音とみなして高調波成分の周波数が調和構造を保ったまま変動するモデル音が提案されている。⁽⁶³⁾ これは基本周波数 $f_1(t)$ を250 Hz とし、第 m 高調波の瞬時周波数 $f_m(t)$ が

$$\begin{aligned}
 f_m(t) &= m f_1(t) \\
 &= m \left(250 + \frac{\sqrt{2}}{6} \sin 4\pi t \right) \quad [\text{Hz}] \quad (5.3)
 \end{aligned}$$

に従って変化する複合音であり、ここでも人工的な母音としてこの複合音を用いた。

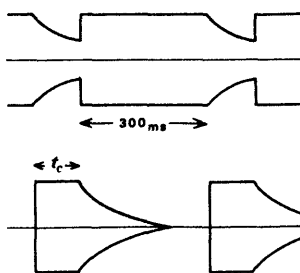


図5.3 測定に用いた信号音の構成

一方子音は調音時における種々の乱流によってつくられるため雑音源によって近似できるが簡単のため子音の周波数スペクトルが平坦なものとして白雑音を用いた。それぞれの試験音は図5.3に示すように母音および子音に相当する信号が300 ms および t_c の時間だけ存在し、かつ包絡の

立ち下り部分はそれぞれ定められた指数減衰を行うようにして残響付加状態を模擬する。これらの試験音は150～6000 Hzの帯域フィルタを経て70 dBの音圧で被験者に与えられる。なお受話器はコンデンサ型であり音圧は6ccカップラによって校正した。

被験者は聴力が健全な成人男子であって調整法によって白雑音の検知限レベルを決定する。実験は t_c が50、70および100 msの3種類について行い、各々5回の測定から得られた結果を図5.4に示す。縦軸のレベルは複合音レベル(70 dB)に対する白雑音レベルを表わしており、横軸は指数減衰から換算した残響時間である。

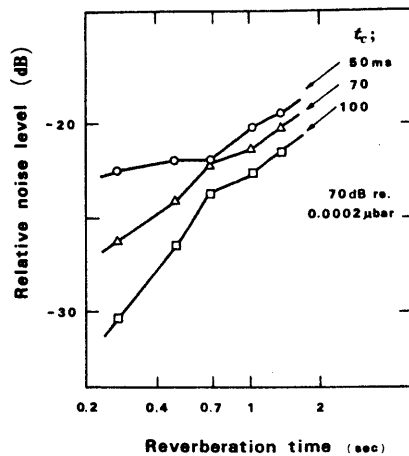


図 5.4 測定結果

この予備実験の結果から白雑音の呈示時間 t_c が長いほど、また複合音の指数減衰が急なほど白雑音の検知レベルが低くなり、先行音の残留成分に起因する妨害効果が少くなる傾向がうかがえる。この擬似音声による実験結果を実際の音声にあてはめることは大胆であるが、およその傾向を知る手がかりとして複合音、白雑音をそれぞれ母音、子音とみなせば次の推定ができる。

すなわち先にのべたように子音は母音より20 dB低く、連続音声の場合として子音継続時間 t_c を50 msとみなせば残響時間は約1秒、また音節単位で発声された場合には t_c が100 msとすれば約2秒が限界値と考えられる。そして子音の平均レベルである-20 dBを境界として子音群が等分されると強引に考えればこの限界値が調整法の閾値であることから半数の子音がマスクされることになり、更にこれらのマスクされた子音がすべて異聴されるという仮定を強いれば子音明りよう度は50%と推定される。

この予備実験の結果および5.2の結果から残響の通話に及ぼす影響をまとめてみると次のようになろう。

- (1) 音声のレベルから残響の影響はまず子音の妨害となって現われる。
- (2) 子音の中でも鼻子音、半母音などはレベルが母音のそれに近いいため他の有声子音や無声子音に

比して妨害には強いと思われる。

- (3) 子音に及ばず妨害は子音の継続時間が長いほど小さくなるといえる。一方子音、母音の継続時間は発話速度に依存するため残響にみる妨害は発話速度が遅いほど小さくなると思われる。
- (4) 大胆な推定として連続発声では残響時間にして約1秒また音節単位で発話された場合には約2秒で子音の半数がマスクされ、子音明りょう度は50%になるとと思われる。

これらの推定を裏付けるため音声に簡単な残響を付加し、残響が及ぼす妨害を定量的に評価することを試みる。

5.4 人工残響付加による明りょう度の低下

通話品質を客観的に表わす尺度の一つに明りょう度があり、残響による明りょう度の低下について河原田ら⁽⁶⁴⁾は残響の物理尺度の一つである残響時間と関連づけられることを示し、また明りょう度計算の適用も行っている。先に述べたように残響妨害は子音に与える影響が特に顕著であろうと推察されることから先行母音の残響成分がそれに続く子音及び母音に及ぼす影響を表現する尺度としてここでも明りょう度を選び、簡単な残響を人工的にほどこした音声について明りょう度を測定する。

試験音節はV C V形式の無意味2音節であり、300~1000 Hzにおける残響時間が70 ms以下の特性をもつ防音室であらかじめ録音した。発声者は男女各1名である。この試験音節にほどこす人工的な残響付加は音源停止以後の残響曲線が音源停止直前の状態から連続して指数減衰に移行する場合として図5.5に示すものを用いた。

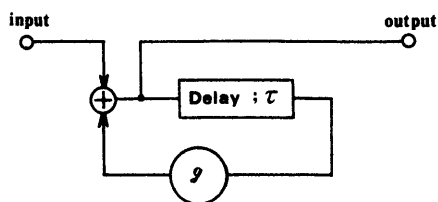


図 5.5 指数減衰を模擬させた人工残響付加回路

受聴条件は2章で述べたものと同じで300~3400 Hz に制限された試験音声を標準受話器DR-305によって片耳受聴する。受聴者は聴力が健全な成人男子4名であり、十分な練習のち各々4回の受聴を行った。明りょう度は受聴者にV C Vで与えられた2音節のうち、後半の音節について受聴された結果を整理することによって求め、これを図5.6に示す。この図からわかるように人工残響を付加する際に用いた素子の遅延時間 τ が200 ms の場合を除いてほとんど明りょう度の差がなく、残響時間にして約0.6秒以上で単音明りょう度は徐々に低下する。遅延時間 τ が200 ms の場合は人工残響に現われる遅延信号の密度、いわゆるエコーの密度が疎であり、明り

よう度測定の対象とした音節の平均が 250 ms であることを考えあわせると試験音節に与える妨害信号が少なくなったことにより明りよう度はさほど低下しないものと出される。

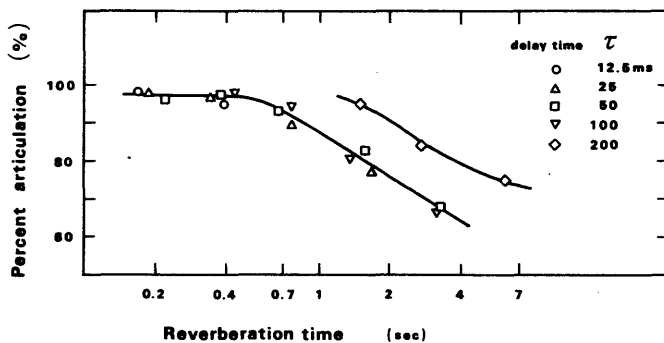


図 5.6 人工残響付加音声による明りよう度低下

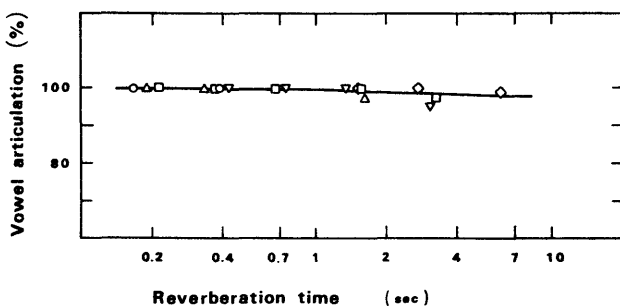


図 5.7 人工残響付加音声の母音明りよう度

さてこの結果から母音明りよう度について整理すれば図 5.7 のようになり 5.3 で推定したように母音が受ける残響妨害は極めて少いといえる。したがって図 5.6 に見られる明りよう度低下はほとんどが子音の異聴によるものであり、子音は残響によっても他の妨害と同様、非常に妨害され易いことを裏付けている。

また 5.3 において子音をレベルから分類した結果鼻音、半母音はレベルが大きいため残響による妨害を受け難いと想像したが、鼻音、半母音のみの明りよう度を C_n とし、鼻音、半母音を除いた有声子音、および無声子音の明りよう度をそれぞれ C_v 、 C_l としこれらを 99% 以下についてプロットすれば図 5.8 のようになる。このようにレベルから分類した子音の種類によって明りよう度は比較のために示した 45° の直線の附近に位置づけられることから明確な差を見いだすことができず各子音は残響によってほぼ同様の割合で明りよう度が低下するものと思われる。そのため子音の種類をレベルの大小にもとづいて分類し、残響妨害を論ずることは母音・子音という大まかな分

類には適用されるものの子音間の分類まで拡張することは適切でないといえよう。

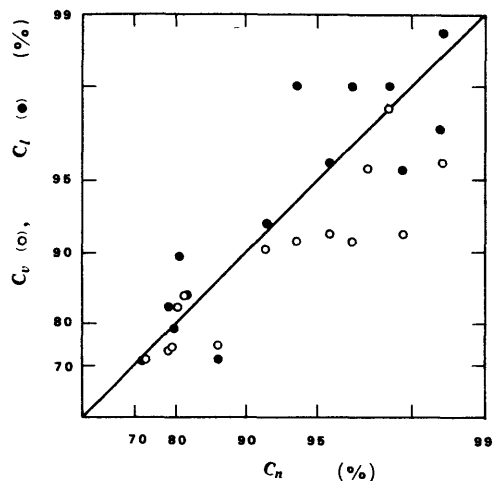


図 5.8 子音の種類による明りよう度

さて今回用いた試験音節は約 250 ms の継続時間であり、連続音声の場合より長いので次に明りよう度と発話速度の関係を明りよう度計算⁽⁶⁴⁾によって推定する。この計算はまず測定の対象となる音節の中央における先行音の残留成分を B_R とし、また対象とした音節の中央においてその音節が上昇したレベルを B_P とし、それぞれを残響特性から求めることから始める。そうすれば実効感覚レベル E_i は B_R 、 B_P を用いて

$$E_i = B_P + P + R - (X (+) B_R)$$

によって与えられ、明りよう度指数を定めることができる。

ここで P ; 音声のピークファクタ

R ; 受聴レベル

X ; 耳内残留雑音

であり、記号(+)はパワ和の演算である。また i は分割した周波数帯域の番号であるがここでは簡単のため残響特性は周波数に依存しないものとし、音節継続時間が 250 ms および 100 ms の 2 種類について受聴レベルを 70 dB かつ受聴者の熟練度を 1 と置き明りよう度を求めこれを図 5.9 に示す。図中の記号は遅延時間が 12.5 ~ 100 ms によってつくられた人工残響音の明りよう度測定値であり図 5.6 と同一である。

また計算による明りよう度は残響を模擬する際の遅延時間により傾向が異なるが 1 例として 25 ms と 100 ms の場合を示した。

計算結果は音節継続時間が 250 ms としたとき今回の実験結果に近い傾向を示しているが残響時間が 1.5 秒以上では実験値の方が計算値を上まわる。この原因は実験結果の母音明りよう度がほとんど低下せずに 100% 近くを保つため、子音、および母音明りよう度が各々独立であれば

の時の単音明りょう度は51%を下まわらないためである⁽⁶⁵⁾

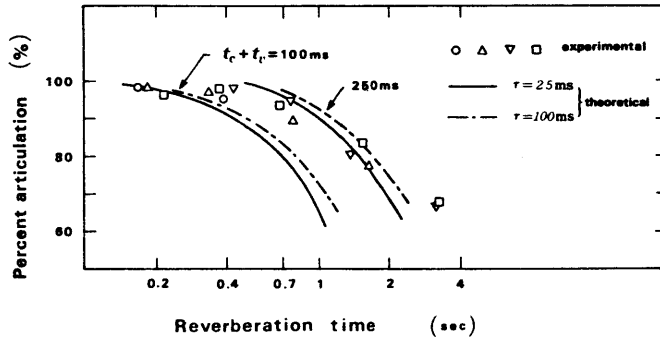


図 5.9 明りょう度の計算結果と実測値

さてこの計算結果から発話速度が大きい場合の明りょう度として $t_c + t_v$ が 100 ms の場合によって推定すれば残響時間が 0.2 ~ 0.3 秒を越えると明りょう度が低下するものと思われ発話速度が大きいほど残響による妨害が顕著になる傾向を示唆している。この限界となる残響時間はアナウンス用として許容される値の 0.25 秒や 3 章で述べたようにプレファレンスの観点から推定された 0.22 秒に近く、連続音声の受聴における共通した性質とも思われるが断定はできず今後の研究課題であろう。

一方子音および母音明りょう度がそれぞれ 50%、100% としたとき単音明りょう度は 76% となり、⁽⁶⁹⁾ この条件となる残響時間は発話速度によって異なるが図 5.9 から約 0.85 ~ 1.8 秒と推定される。また実験によって測定された値も遅延時間が 200 ms の場合を除いて約 2 秒であり、これらの値は 5.3 で強引に子音明りょう度が 50% になると予想した残響時間がそれぞれ 1 秒、2 秒であることを考えあわせると興味深い。

5.5 ま と め

残響が通話に及ぼす影響を音声の時間的側面から考察した結果、次の結論が導びきだせる。

- (1) 残響成分が音源の消滅とともに連続して指数的に減衰する場合を想定し、人工的に残響を付加した音声の単音明りょう度を測定した結果、残響時間と関連づけることができる。
- (2) 音節の構成を母音と子音に分けて考えた場合、先行音の残響成分によって後続子音は妨害を受けやすく簡単な残響を付加して行った実験によっても母音明りょう度がほとんど低下せず、子音明りょう度の低下が認められた。
- (3) 子音の中でも鼻子音や半母音などレベルの比較的大きなものは妨害され難いと思われたが明りょう度試験ではレベルの大きさから分類した子音間に差は見られなかった。

- (4) 残響妨害による明りょう度低下を計算によって推定する場合、明りょう度が80%以下では測定値より下まわるが、80%以上では簡単な残響妨害における明りょう度を計算できる。
- (5) 子音に及ぼす残響妨害は発話速度に依存し、連続音声で明りょう度80%を満足する残響時間の限界は0.75秒であろう。一方音節の時間長が250ms程度となるように速度を遅くした場合の限界は約1.6秒であろうと思われる。

第6章 多重音声による通話妨害⁽⁶⁶⁾

6.1 はし が き

われわれが日常行っている会話はある意味で最も単純な音声通信系といえるが、このような通信方法では外来雑音の大きさによって通話品質が大きく左右される。特に雑踏の中における会話などでは受聴者は一般に特定の送話者からの音声に注目しているが他から混入する会話音声⁽⁶⁶⁾が妨害の要因として作用することは日常よく経験することである。

このように雑踏の音声⁽⁶⁶⁾が妨害雑音として音声通話に及ぼす影響の一つはカクテルパーティ効果⁽⁶⁶⁾として知られているが、本章ではこのような妨害要因としての音声⁽⁶⁶⁾が通話に及ぼす影響を議論する上での基礎データを得るため片耳受聴に限定していくつかの考察を試みる。

まず音声の基本的な一性質を表わす瞬時振幅確率密度関数⁽⁶¹⁾から出発し、多くの音声⁽⁶¹⁾が同時に存在する場合の推定を試み、妨害音発声話者数についての一般的な性質を論ずる。そして妨害音声を変動騒音としてながめ、妨害音発声話者数とレベル変動の関係を実験によって示す。

またこのような妨害音声⁽⁶¹⁾が実際の音声の聴取に及ぼす妨害効果を明りょう度試験によって評価する。その際、音声レベルに対する妨害音声の等価レベルを信号対雑音比とみて整理し、妨害音発声話者が明りょう度に与える影響について考察し、更にこの妨害を音声の長時間スペクトルと等価な定常雑音や、変動する母音等による妨害と比較検討する。

6.2 多重音声の性質

まずはじめに雑踏の中で行われる会話状態を正確には握するため次のようなモデル⁽²²⁾を考える。

今、雑踏において形成されている会話のグループ数を n とする。このとき i 番目のグループ($i = 1, 2, \dots, n$)に着目すればこのグループに属する m 名($m \geq 2$)のうち話者は一般に1名と考えるとよく、そのため受話者数は $m-1$ となるが、この受話者の各々はグループ i の話者による音声を信号とし、 i 以外のグループの話者による音声、つまり $n-1$ 名の話者による音声はこのときすべて雑音として作用することになる。つまり妨害信号は $n-1$ 名の話者による会話音声であってこれを多重度 $n-1$ の音声と呼ぶことにする。このようなモデルにおいてグループ i の受話者が受ける妨害音レベルは各妨害音発声話者との距離およびこれらの雑踏が囲まれている壁面を考慮すれば定めることができ MacLean⁽²²⁾はこのモデルについて各話者の発話レベルの変化に着目して解析を行っている。

ここでは多重度のちがいによる妨害音声の性質に着目し、この問題について議論を進めてゆく。

前章で述べたように音声の基本的な性質を表現する手段の一つに瞬時振幅確率密度があり、これは国語による本質的な差がないとされているためこの確率密度によって会話音声の一般的な特徴を

論ずることができよう。

さて音声の多重度が n のとき、この音声の瞬時確率密度関数を $W_n(x)$ によって表現する。ここで x は長時間実効値で正規化した瞬時振幅である。話者が1名による会話音声の瞬時振幅確率密度関数 $W_1(x)$ は前章(5.2)式で用いたような形となり次式に示す。

$$W_1(x) = 0.57 \frac{e^{-\sqrt{2} \frac{|x|}{1.32}}}{1.32 \sqrt{2}} + 0.43 \frac{e^{-\frac{1}{2} \left(\frac{x}{0.132}\right)^2}}{0.132 \sqrt{2} \pi} \quad (6.1)$$

この式の第1項は両側指数分布、また第2項はガウス分布であってそれぞれ母音及び子音に対応していると考えられているが、連続音声の場合、音声の休止部分も付随するため⁽⁶⁷⁾このガウス分布には子音、および音声の休止部分が含まれていると考えることができる。

一方、妨害雑音として連続音声をとらえれば前章でも検討したように子音等小振幅部分による妨害よりもエネルギーの大きい母音等大振幅部分によって主要な妨害がなされるものと考えられるため、連続音声による通話妨害を論ずる場合、これらの出現頻度も考慮する必要が生ずる。

さて $W_1(x)$ は $x=0$ に対称な関数であり、正領域について記せば図6.1に示す曲線となる。なおこの図にはサンプリング間隔1msで100秒間にわたって測定したいくつかの会話音声の瞬時振幅確率密度も同時に示した。この実験に用いた会話音声は表6.1に示すような内容であり、すべてあらかじめFM放送から録音されている。図6.1からわかるように式(6.1)による値は異なる

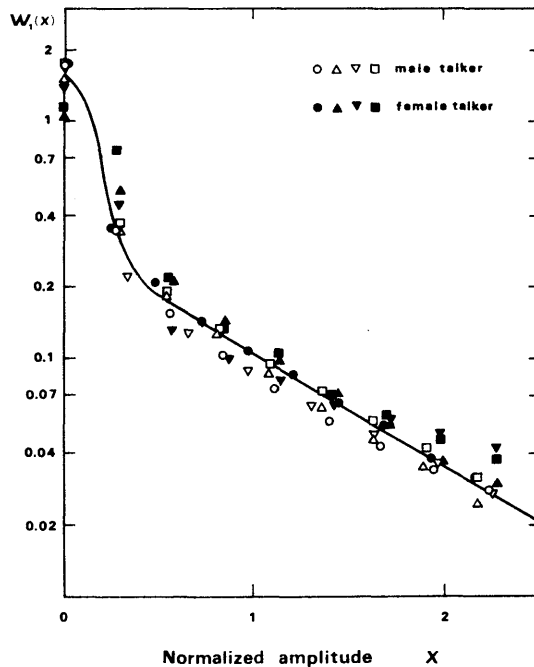


図 6.1 測定された確率密度と式(6.1)の比較

記号	話者の性別	音声の内容
○	男 性	ニュース朗読
△		ニュース朗読
□		文章朗読
▽		音楽解説
●	女 性	音楽解説
▲		
■		
▼		

表 6.1 分析に用いた音声の内容

会話音声についての瞬時振幅確率密度も適確に表現しているため式(6.1)を以後会話音声の基本として用い、多重度 n に対する $W_n(x)$ を導出する。

さて式(6.1)を一般化すれば式(6.2)のようになる。

$$W_1(x) = p \frac{1}{\sqrt{2}\sigma_v} e^{-\sqrt{2}\frac{|x|}{\sigma_v}} + (1-p) \frac{1}{\sqrt{2\pi}\sigma_c} e^{-\frac{1}{2}\left(\frac{x}{\sigma_c}\right)^2} \quad (6.2)$$

前章で用いたように p 、 $1-p$ は母音、子音のレベル差およびそれぞれの継続時間を考慮した結果母音および子音もしくは音声休止の出現頻度とも解釈でき、前章で用いた図6.2の連続音声モデルによって多重音声の性質を明らかにしてゆく。

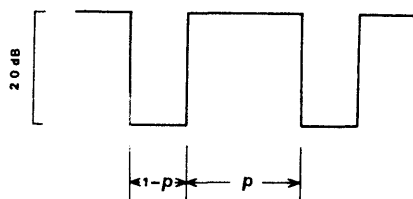


図 6.2 音声モデルのレベル変化

さて n が2のとき、レベルが大きくなる場合として

- i) 2名の母音が重畳する場合、および
- ii) 1名の母音のみが存在する場合、

の2通りが考えられるが、これらのレベル差は約3 dBであることから大振幅部分に3 dBのレベ

ル変動を許せば母音の出現頻度は

$$1 - (1 - p)^2$$

となる。

このような考えから n が 2 および 4 についての確率密度 $W_2(x)$, $W_4(x)$ は母音, 子音のレベル差が

$$20 \log_{10} \frac{\sigma_v}{\sigma_c} = 20 \quad [\text{dB}]$$

であったことを考慮すれば次式で与えられる。

$$W_2(x) = 0.815 \frac{e^{-\sqrt{2} \frac{|x|}{1.11}}}{1.11 \sqrt{2}} + 0.185 \frac{e^{-\frac{1}{2} \left(\frac{x}{0.111}\right)^2}}{0.111 \sqrt{2\pi}} \quad (6.3)$$

$$W_4(x) = 0.966 \frac{e^{-\sqrt{2} \frac{|x|}{1.02}}}{1.02 \sqrt{2}} + 0.034 \frac{e^{-\frac{1}{2} \left(\frac{x}{0.102}\right)^2}}{0.102 \sqrt{2\pi}} \quad (6.4)$$

これらの曲線をそれぞれ図 6.3, 6.4 に実線で示す。またこれらの図には表 6.1 にある会話音声信号を基本にそれぞれの長時間実効値を等しく加え合わせて作成した多重音声による実測値もあわせて示したが, 式 (6.3), (6.4) によって実測値をうまく説明できると思われる。

このように音声の多重度 n が 4 までについての確率密度を音声の空白部分, すなわち子音もしくは休止区間に相当する時間率の減少をもとに解析したが, この方法が非常に単純にもかかわらず図 6.3, 6.4 に示すように現実の多重音声のふるまいはかなり適確に表現されていることがわかる。

したがって音声の多重化にともなう一つの特徴は会話音声の有していた子音および休止の時間率, つまり空白の割合が n の増加につれて減少することによって代表されるものと思われる。

次に多重度が 4 以上の場合について考えてみる。多重度 n が 4 の場合は式 (6.4) からわかるように第 2 項の係数がほとんど 0 に等しくなって瞬時振幅確率密度は第 1 項の両側指数分布の部分で定められることであり, 図 6.4 の実測値もそのような傾向を示している。したがって多重音声は n が 4 において瞬時振幅確率密度が近似的に両側指数分布として扱うことができることから, n が 4 以上では新しく両側指数分布を基本として議論を進める。

さて多重度が 4 以上では両側指数分布の重畳として扱えるため付録に示すような方法によって次式の確率密度が得られる。

$$W_8(x) = \frac{1}{2} e^{-2|x|} (2|x| + 1) \quad (6.5)$$

$$W_{16}(x) = \frac{3\sqrt{6}}{16} e^{-\sqrt{6}|x|} (2x^2 + \sqrt{6}|x| + 1) \quad (6.6)$$

これらの式についても先程と同様に実測値とともに図 6.5, 6.6 に示したが実際の多重音声のふ

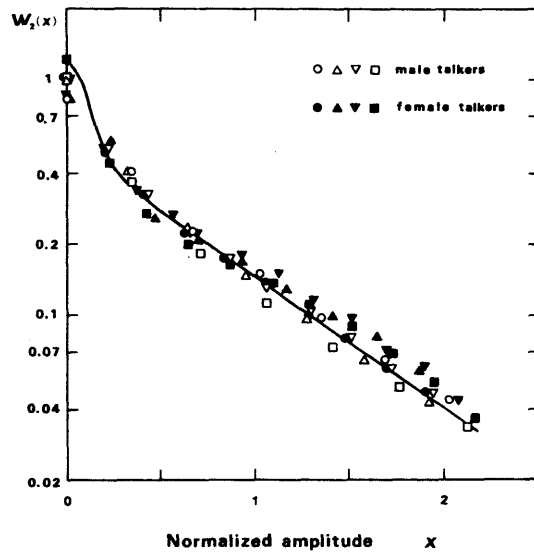


図 6.3 多重度 $n = 2$ に対する実測値と計算値の比較

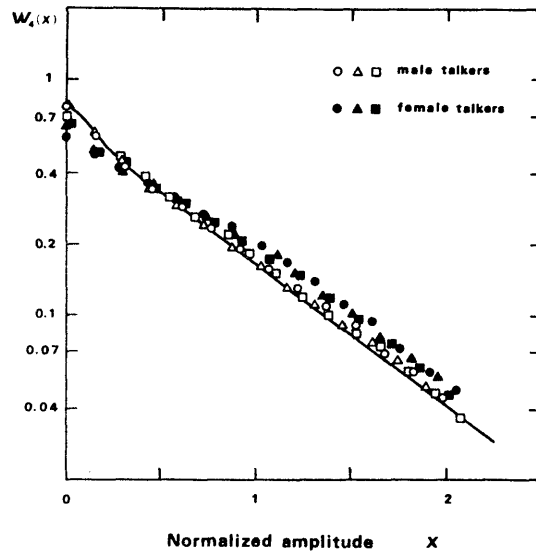


図 6.4 $n = 4$ における比較

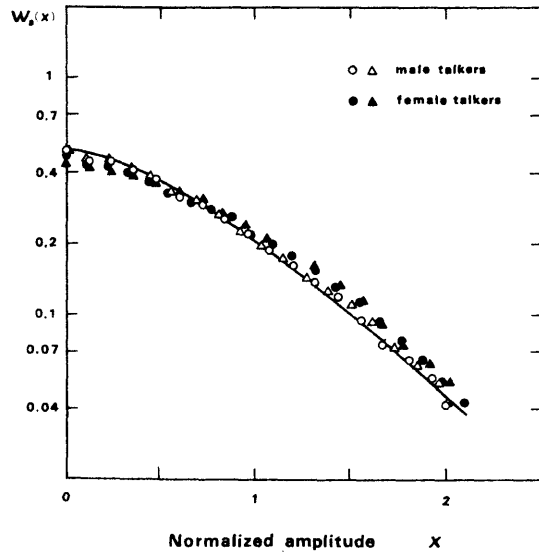


図 6.5 $n = 8$ における実測値と理論値の比較

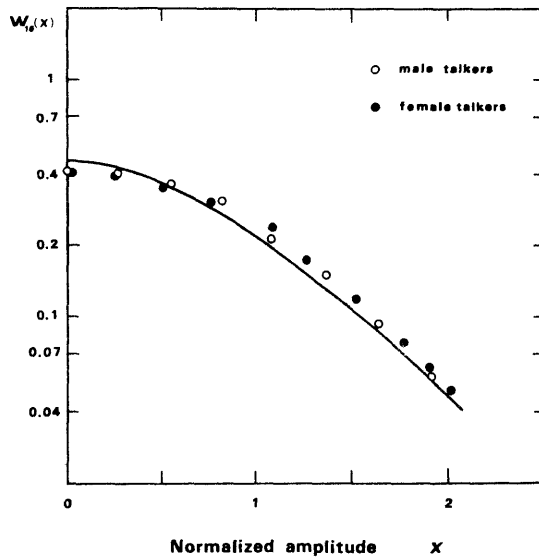


図 6.6 多重度 16 における実測値と理論値

るまいを忠実に表現していると思われる。

このように瞬時振巾確率密度を用いて評価した多重音声の性質をまとめると次のようになる。

- (1) 音声の多重度 n が4を境界とする2通りの解析方法によって説明することができる。
- (2) n が4以下の場合、休止部分を代表しているガウス分布の割合 $(1-p)$ が n によって $(1-p)^n$ とすることによって近似できるため、多重度 n の増加は主として音声の休止確率の減少効果をもたらす。
- (3) そして $n=4$ で上記の割合はほぼ0となり、瞬時振巾確率密度は両側指数分布で代表させることができる。
- (4) このため n が4以上では母音による大振巾部分を表現していた両側指数分布の重畳によって瞬時振巾確率密度を計算することができ、 n が大きくなるにつれて中心極限定理によってガウス分布に漸近するであろう。⁽⁶⁸⁾

このように瞬時振巾確率密度による議論からは多重度が増加するにつれて音声は両側指数分布の重畳に近づくものと推察される。

6.3 多重音声の騒音レベル分布

前節でみたように多重音声のふるまいの一つとして多重度が4以下では音声の休止割合が減少することを明らかにしたが、通話にこのような多重音声を与える妨害結果を調べるには先ず多重音声の妨害雑音としての性質を見いだす必要がある。特にこの多重音声は多重度 n が小さい場合、レベル変動が大きいと予想されることから変動雑音としての取り扱いが必要となる。

一方、生活環境などに影響を及ぼす騒音が時間とともに変動する場合、これを定量的に表示する手段として等価騒音レベル L_{eq} によって評価することが多い。そのためここでも通話に及ぼす多重音声の妨害を定量的に論じる手段として L_{eq} およびレベルの累積確率分布によって変動する雑音としての性質を明らかにする。

多重音声の等価レベル L_{eq} は騒音における測定と同様にA特性を経由した信号を分析することにより行った。そしてこの信号はアナログ計算機によって模擬された騒音計回路によって騒音レベルに変換され、更にデジタル計算機によって処理される。測定はサンプリング間隔30 ms で100秒間にわたる多重音声について L_{eq} 及びレベルの累積確率を計算した。この結果を L_{eq} を規準に正規確率紙に表示したものが図6.7である。

さてこの図によれば多重度 n が1、つまり1名による会話音声のレベル変動は L_{eq} 以下の部分および $L_{eq} + 10$ dB 以上の部分がそれぞれほぼ直線とみなせる分布となる。レベルが $L_{eq} - 25$ dB 以下では測定系のノイズのため正確な値を示していないが n の増加とともにこれらの曲線は傾きがより急な直線に徐々に近づき、またレベル分布は L_{eq} を中心とするガウス分布に漸近しかつ変動も小さくなる様子がうかがえる。そして音声の多重度が小さい場合、レベルが L_{eq} を大きく下まわることに特徴を見いだすことができ、このレベル変動は音声の休止部分の存在に置か

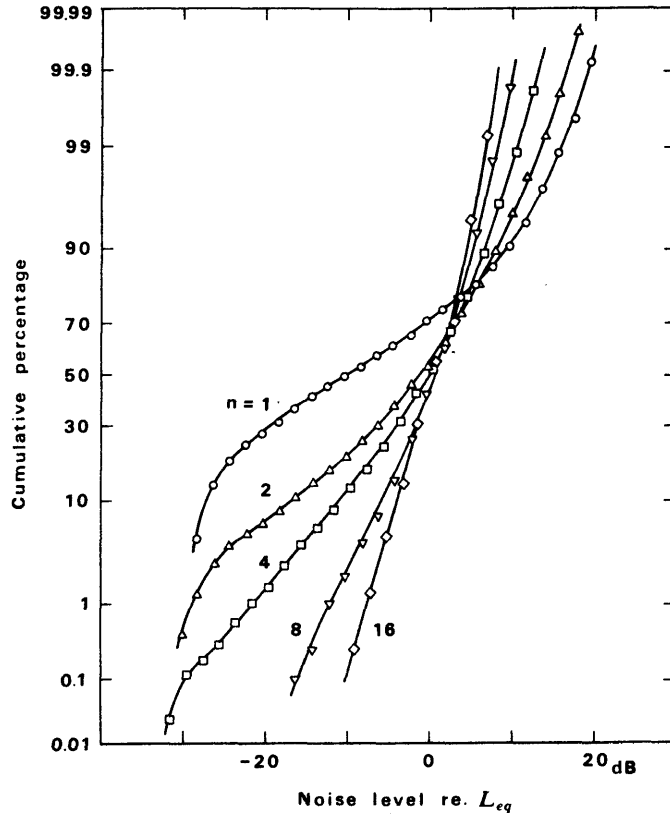


図 6.7 多重音声の累積分布

えて議論できる。

一方 n が 4 を越えるような多重音声ではレベルの変動巾も小さくなり L_{eq} のまわりにはほぼ対称にゆらぐことがわかる。これらのレベル変動の速度は音声の発話速度に関連しているものと思われるが、一例として変動する騒音レベルが L_{eq} を交叉する平均時間間隔は図 6.8 のようになって前述の変動巾についての性質を考えあわせると多重度が増すことによってレベル変動も小さく、かつ速くなることがわかる。

このように音声の多重度が増加するにつれてレベル変動巾が変化の様子がうかがえ、特に $n = 1$ においてはレベルが L_{eq} から大巾に低下する割合がかなり大きいことが推定される。この変動をより明確にするため累積確率が 5% および 95% となるレベルと L_{eq} の差をそれぞれ L_5, L_{95} とし、また変動巾の目安として 90% レンジ ($L_5 - L_{95}$) で代表させてみると図 6.9 のようになる。この図から先ず音声の多重度 n が増加するにつれて変動巾を代表させた 90% レンジが急激に減少してゆく様子がうかがえる。また L_5, L_{95} を比べると多重度 n が小さい場合、 L_{95} つまり 0.95 以上の時間率で越えているレベルは L_{eq} を大きく下まわっていることになるが、 n の増加につれて L_{eq} に漸近する様子が認められる。しかし L_5 の方は L_{95} に比して n による変化が少い

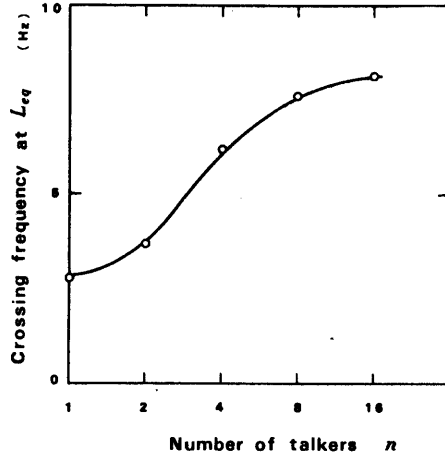


図 6.8 多重音声のレベルが L_{eq} を
交叉する平均周波数

ことから前節で述べたように多重度の増加にもなって音声休止の割合が減少する性質をこのようにレベル変動を測定することによって確認することができる。

また図 6.9 には音声の長時間スペクトルと近似のフィルタを経由させた白雑音によるレベル変動も併せて記した。このフィルタは図 6.10 に示すような特性であって、このフィルタを通過した定常等スペクトル雑音は n が 16 の多重音声に比してもレベル変動はかなり小さいことがわかる。

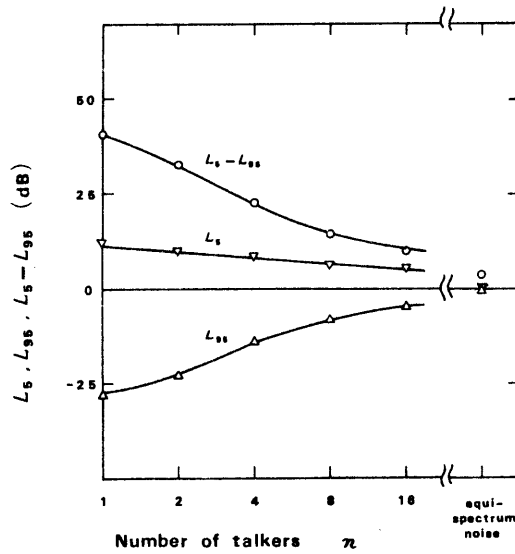


図 6.9 L_5 , L_{95} , $L_5 - L_{95}$ によって
表わしたレベル変動

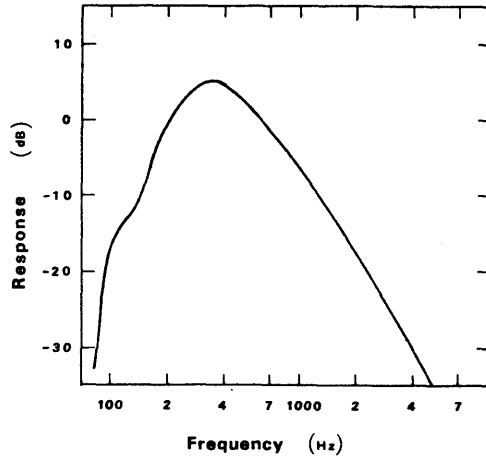


図 6.10 音声等価フィルタの特性

6.4 多重音声による明りよう度低下

前節で議論したように多重音声は多重度 n の増加につれてレベル変動が小さくなることが一つの特徴であると述べたが、次にこのような多重音声を実際の音声の聴取に与える妨害効果を調べてみる。通話の品質を表現する尺度は前章と同じ明りよう度により評価を行う。

さて妨害音としての多重音声は表 6.1 に掲げたもののうち男性話者による会話音声を基本にし、その長時間実効値を規準にして等レベルで混合したものを用いた。一方明りよう度試験は女性により読みあげられた単音節を用いた。これらの信号は 150～6000 Hz の帯域に制限されたのち受話者の片耳に与えられたが、このとき明りよう度試験音節は常に約 64 dB のレベルに固定されている。受聴者は聴力が健全な成人男子 2 名であり、のべ 4 回の受聴結果から平均と標準偏差をあわせて図 6.11 に表示する。ここで横軸は SN 比で表わしてあり、信号レベルとしては明りよう度試験音声のレベルであって、これは 200 音節にわたり測定した平均から算出した。なおこの標準偏差は 1.9 dB であった。また雑音としての多重音声のレベルは変動する雑音としての性質を有していることから前節で用いた等価騒音レベル L_{eq} によった。

さて図 6.11 の受聴結果からまず SN 比の低下とともに単音明りよう度も低下する傾向がみられる。この傾向は音声の多重度 n によって異り n が増加するにつれて急激な低下となる。この原因の一つは前節で考察したように多重音声が増加することによるものと思われる、妨害音声に存在する音声休止の割合によって明りよう度に及ぼす影響が異なるものと推察される。

一方、SN 比の減少ともなう明りよう度の変化は多重度 n が増加するにつれてより急激に劣化する傾向を見せるが、 n が 8 および 16 ではこの差が明確でなく、 n の増加とともに明りよう度変化の性質が一定の傾向に漸近するものと思われる。

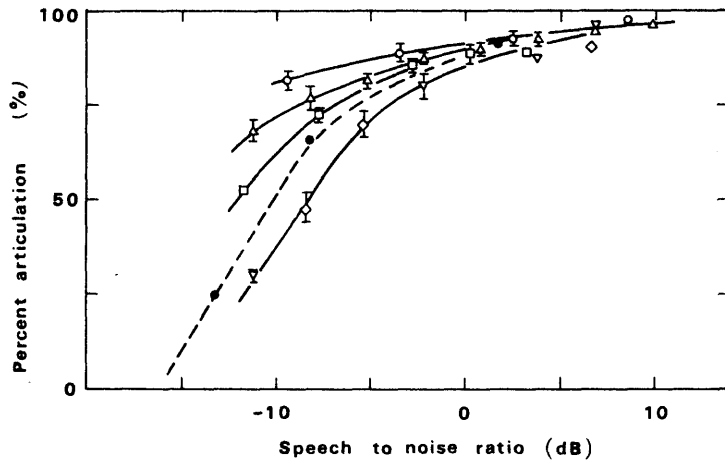


図 6.11 多重音妨害による単音明りょう度の低下

また SN 比が同一の場合では多重度が大きいほど通話に与える妨害効果も大きいと推定され、SN 比が負つまり多重音声の等価騒音レベルが音声レベルを上まわるような悪条件では特に顕著である。

さてこれら多重音声による通話妨害を定常雑音による妨害と比較してみよう。これは前節で用いたように音声の長時間スペクトルに近似させたフィルタによって音声と等スペクトルな定常雑音により明りょう度の劣化を測定したもので、この結果は図 6.11 に破線で示されている。この結果は三浦⁽⁶⁾による報告とはほぼ同様の傾向を示し、SN 比が負のときで明りょう度劣化特性は多重音声による実験結果と類似の傾向を示すものの、多重度が 8 以上では多重音声による妨害の方が大きくなることを示しておりこの原因について考察してみる。多重音声と等スペクトル定常雑音の物理的なちがいの一つは図 6.9 でみたように等スペクトル定常雑音のレベルがほとんど変動せず 90% レンジが約 3.5 dB であったのに対して多重度が 16 の音声では約 10 dB にも達し、そのため先ず考えられることはこのレベル変動のちがいであろう。そこで等スペクトル定常雑音に多重音声と同様のレベル変動を与えたときの明りょう度を測定することによって比較してみる。定常雑音にレベル変動を与えるには図 6.12 の回路によった。これは多重音声の二乗電圧を 40 ms の時定数回路によって平滑したのち、⁽⁷⁾ その平方根に相当する電圧により等スペクトル定常雑音の振巾を変調さ

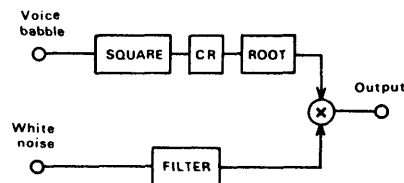


図 6.12 多重音声と同様のレベル変動を与える回路

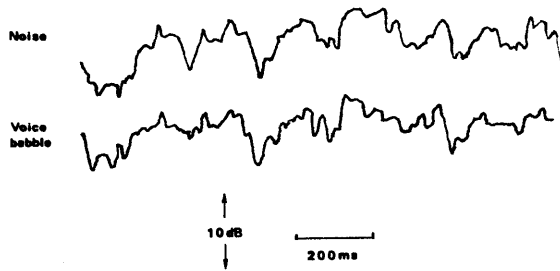
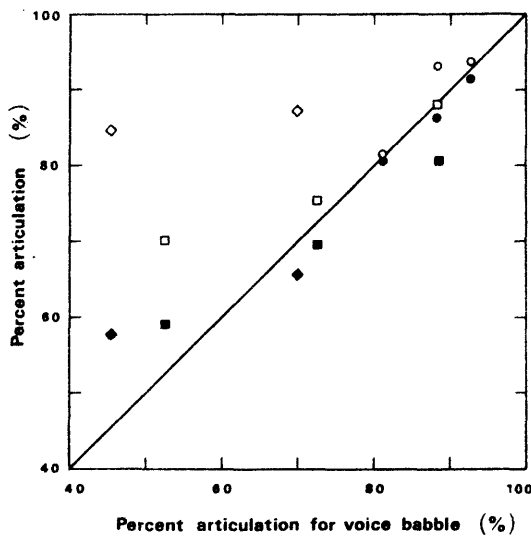


図 6.13 $n = 16$ の多重音声とそれによる変動を与えた雑音の比較
 B & K 2305 lower limiting frequency : 50 Hz
 writing speed : 500 mm/sec

せ変動雑音を擬似的につくりだした。図 6.13 は n が 16 における多重音声及び、これによって図 6.12 の回路で変動を与えた等スペクトル変動雑音のレベル変動の様子を示している。両者はほとんど一致しており、多重音声とはほぼ同様のレベル変動を等スペクトル定常雑音に与えることができる。

さてこの変動雑音を多重音声に替えて明りょう度低下を測定した結果を図 6.14 で比較して示した。両者による妨害の程度が同じならこれらの結果は比較のために示した直線上に並ぶが図からわかるように n が 4, 16 と大きくなるにつれて多重音声が示した明りょう度よりも増加しており、そのため等スペクトル変動雑音による妨害は多重音声より少ないことがわかる。言いかえれば n が 16 の多重音声および同様の変動特性を有する等スペクトル変動雑音が明りょう度に及ぼす妨害効



noise	n		
	1	4	16
fluctuated noise	○	□	◇
fluctuated complex tone	●	■	◆

図 6.14 等スペクトル変動雑音および変動複合音による単音明りょう度の比較

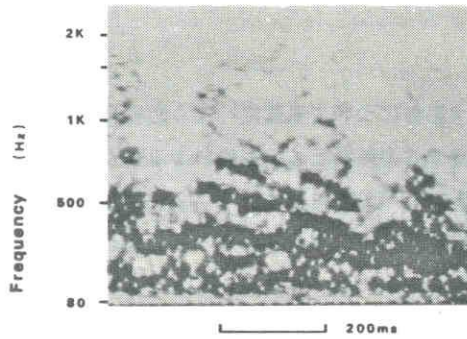


図 6.15 $n = 16$ における多重音声のソナグラム

果は同一でない結論づけられる。そこで他に考えられる原因としてスペクトルをとりあげ更に検討を加える。

図 6.15 に多重度が 16 の音声を分析したソナグラムを示す。先の実験に用いた等スペクトル変動雑音ではこのパターンが一定スペクトルの強弱になるが多重音声ではこのように母音と思われる明確なホルマントの部分を見いだすことができる。したがって n が 16 の多重音声では母音としての性質がかなり残存していると考えられることからこのスペクトルを図 6.10 で示されるような音声の長時間スペクトルで近似することは危険であろう。そこで次に母音をスペクトル的に模した複合音を用いて実験を重ねる。この複合音は前章 (5.3) 式で示したものであり図 6.12 の白雑音の代りに用いた。したがってこの複合音は図 6.12 の回路によって多重音声と類似のレベル変動をともなうがスペクトルの面からは母音を単純に模擬したものと考えることができよう。

さてこの変動複合音による明りょう度変化は図 6.14 に併記されており図に示した直線に近い位置にあることから明りょう度に関して多重音声とふるまいが似ていることがわかる。そのため多重音声は等スペクトル定常雑音あるいは等スペクトル変動雑音よりも、等スペクトル変動複合音に近い妨害特性を有していると結論づけることができる。6.2 における瞬時振幅確率密度の解析において n が 4 以上で両側指数分布の重畳、つまり母音の和によって多重音声は近似的に表現されたと考えたが、このように実際の妨害についても母音を近似させた複合音の示す妨害効果と同様の傾向を見いだすことができた。

ここでの議論で明らかにされた点をまとめるとまず SN 比が 0 dB 以上では多重音声の妨害にともなう明りょう度は 90 % 以上を保ち妨害効果は少いことがわかる。

更に SN 比が 0 dB 以下では SN 比の減少につれて明りょう度に及ぼす妨害量も増加し、これは音声の多重度にも大きく依存することから SN 比および妨害音声の多重度が明りょう度低下の要因となることが結論づけられる。

また多重度 n が 8 以上における妨害の性質は多重音声と同様のレベル変動を与えた複合音による明りょう度劣化と性質が類似であり、音声とスペクトルが等しい定常あるいは変動雑音よりも妨害効果が大であることが判明した。

6.5 ま と め

本章では妨害音声に通話に及ぼす影響について多重度つまり妨害音発声話者数による効果を論じた。そのため先ず音声の性質を表現する物理量として言語による差がないとされている瞬時振巾の確率密度をとりあげて考察した結果、音声の多重度によって2つの性質が認められた。その1つは多重度が4以下の場合、妨害音声の瞬時振巾確率密度はガウス分布および両側指数分布の合成で表現されることから妨害音声のレベル変動が大きい性質をとりあげ、これはレベル分布を測定した結果からも確かめられた。

そして多重度が4において妨害音声はレベルが連続する性質に移行し瞬時振巾確率密度は単一の両側指数分布で近似されることから、2番目の性質として4以上の多重度では確率密度が両側指数分布の重畳によって表現され、連続した母音としての性質を見いだすことができた。

また多重音声のレベル変動は多重度の増加につれて小さくなると推定され、これは実験によっても裏付けることができた。

このような多重音声が実際の通話に与える影響を明りょう度によって評価した結果は、多重度が小さい場合、瞬時振巾確率密度を用いて議論したように妨害音声の休止時間率によって妨害の確率が減少するため明りょう度の低下はさほど顕著でない。しかし多重度の増加につれて妨害音声は連続性となり、多重度が8もしくは16では擬似母音を妨害雑音としたときと同様の明りょう度低下特性を示した。この点に関しても瞬時振巾確率密度の解析において推定したように多重度が多くなるにつれて妨害音声は母音の重畳で近似されると述べたことに一致した。

6.6 付 録

音声の多重度 n が4の場合、瞬時振巾確率密度関数 $W_4(x)$ を両側指数分布とし

$$W_4(x) = P_1(x) = \frac{\lambda}{2} e^{-\lambda|x|}$$

から出発する。この特性関数を $g(\theta)$ とすれば

$$g(\theta) = \frac{\lambda^2}{\lambda^2 + \theta^2}$$

となり両側指数分布が m 回重畳されたときの特性関数は $g(\theta)^m$ となることから Fourier 変換の理論を適用すれば次式が得られる。

$$P_m(x) = \left(\frac{\lambda}{2}\right)^m |x|^{m-1} \frac{e^{-\lambda|x|}}{(m-1)!} \sum_{r=0}^{m-1} \frac{(m+r-1)!}{r!(m-r-1)!} \cdot \frac{1}{(2\lambda|x|)^r}$$

この式を実効値で規準化すれば

$$P_m(x) = \frac{\sqrt{2m}}{2} \left(\frac{\sqrt{2m}|x|}{2} \right)^{m-1} \frac{e^{-\sqrt{2m}|x|}}{(m-1)!} \sum_{r=0}^{m-1} \frac{(m+r-1)!}{r!(m-r-1)!} \cdot \frac{1}{(2\sqrt{2m}|x|)^r}$$

となる。これより

$$W_8(x) = P_2(x) = \frac{1}{2} e^{-2|x|} (2|x| + 1)$$

$$W_{16}(x) = P_4(x) = \frac{3\sqrt{6}}{16} e^{-\sqrt{6}|x|} (2x^2 + \sqrt{6}|x| + 1)$$

等が導びかれる。

第7章 あ と が き

本論文では音声通信の品質を劣化させる要因として残響の影響及び音声を主体とする騒音の2点について、これらが通話に与える妨害を定量的に評価し、種々の角度から考察を行った。

残響を伴う音声を主観的に評価することは従来困難であったが残響音を等プレファレンスとなる信号対雑音比に置きかえることにより尺度化できることを示した。一方信号対雑音比は通信系の品質を表示する量として最も一般的であるため、残響音の主観の評価をこれに結びつけることによって通話品質の比較が極めて簡単に行うことができる。

この方法によって人工残響音の品質を測定した結果、等プレファレンスとなる信号対雑音比は残響発生に用いた素子の遅延時間が30 msにおける結果を中心として分布する。この原因は人工的残響付加に見られる規則的遅延が聴覚に作用するためと考えた。

そのため人工残響音の通話に対する評価に新しく規準化を導入し、30 ms以内の遅延信号に負の妨害、つまり品質の上昇をもたらす傾向を見いだした。この規準化については%ディスタージの観点から考察した結果からも従来発表されている単一エコーと同様の傾向を示すことからその有効性を確かめることができた。このようにプレファレンスの点から残響音声の品質を信号対雑音比に置きかえたが、残響時間が200 ms以下では信号対雑音比との関係がほぼ直線とみられることから次に残響付加系のインパルスレスポンスとの関係を考察した。

それによれば等プレファレンスとなる信号対雑音比は従来から利用されているD値を用いた表現に帰着できることが明らかとなった。つまり聴覚にとって有効な信号は直接音及びそれに続く間接音成分の全エネルギーであり、他方40 ms以後に到着した間接音のエネルギーは妨害雑音として作用する。これを裏付けるためエコーを伴う音声と雑音を加えた音声との一対比較によるプレファレンステストを行ったが、単一エコーの場合を除き直接音から40～50 ms以上遅れた反射音の存在が音声品質の劣化をもたらすことがわかり、これは上記の結果とよく合っている。このため片耳受聴についてもD値と同様の関数を用いることにより、通話にとっての好ましきの尺度を部屋の物理特性であるインパルスレスポンスと実験的に結びつけることが可能となった。ここで得られた40～50 msの遅延時間の値は室内音響学では経験的に知られていたものの定量的な扱いが不足していたが、本論文では主観的な品質、つまり通話にとっての好ましきを物理量である信号対雑音比と関連させることができ、上述の遅延時間を限度としてそれ以上の遅れを伴う間接音エネルギーがあたかも雑音として働かせることをいくつかの実験を通して裏付けることができた。したがって残響の存在による品質の劣化を防止するには40～50 ms以上の遅れをもつエネルギーを抑制することが有効であることが結論づけられる。

音声通信における主観的な好ましきについての以上の議論に引き続き通話品質の客観的な尺度である明りょう度によって残響の影響を論じた。

音声通信に用いられる符号は母音と子音に大きく分類でき、これら母音・子音は通話の客観的な

尺度である明りょう度と結びつけられることから通話を母音と子音によって代表させる音声モデルを提案し、明りょう度との関係をマスキング現象により解明した。そこでは先ず音声の統計的性質を調査及び実験により求め、これらの値を反映した音声モデルを定めた。そして連続音声に残響が及ぼす影響が最も顕著な例として母音の残響成分が後続子音をマスクする現象をとりあげ次の点について考察を行った。第一に擬似的な音声信号に簡単な残響が付加された状態を実現し、受聴実験により子音に相当する部分の検知限を求め残響時間との関係を論じた。第二に人工的に残響を付加した試験音声の明りょう度を測定した結果を音声モデル及び擬似音声の検知限測定結果をもとに議論した。そして残響の存在が音声通信に及ぼす影響は他の多くの雑音妨害と同様に子音明りょう度の低下となって現われる特徴を結論づけ、明りょう度が音声の継続時間つまり発話速度に依存することを実験及び明りょう度計算によって示した。したがって残響下での通話は発話速度を小さくすることによって明りょう度低下をかなり防止できることがマスキング理論を用いて結論づけることができることも音声モデルの有効性も確認できた。

最後にここで提案した音声モデルを用い、音声通信に及ぼす妨害要因である音声騒音による影響を検討した。そこでは妨害音発生話者数による音声騒音の瞬時振巾に対する確率的な特徴を前述のモデルによって理論的に算出し、実験結果との一致をみた。また本論文で提案した音声モデルの有効性をここでも確かめることができ、非常に簡単な構造であるにもかかわらず音声の時間的な現象を端的に表現するモデルであると考えられ、今後これらの問題解決に役立つと思われる。音声騒音について上記の特徴を考慮しつつ実際の妨害量を測定したが、通話妨害は妨害音発生話者数に比例し、4名以下の話者による妨害ではその音声が休止することにより実効的な妨害は少くなる。しかし4名以上の妨害話者が存在する場合には妨害音声に母音としての特徴が顕著になり、16名の話者による妨害量は比較のために選んだ白雑音による妨害を上まわることが明らかとなった。そしてこのような多重音声騒音として通話に与える影響は騒音レベル及び妨害音発生話者数が主な要因となることが知られた。

以上のように本論文は最も基本的な通信系、すなわち送受信端がともに人間系であって音声という最も古くから用いられてきた符号化方式によって実現される情報伝達系の効率を劣化させる要因として残響及び多重音声による妨害をとりあげ、これらを工学的立場から定量的に評価し音声通信のより円滑な利用に際して音声信号が人間の聴覚に与える影響についての基礎資料を提供することができた。

謝 辞

本研究の遂行にあたり最後まで熱心な御指導をいただいた三重大学工学部故渡辺真吾教授に心から感謝の意を表するとともにつつしんで御冥福を祈る。

本研究の遂行にあたり終始懇切なる御指導をいただき、また本論文をまとめるにあたっては数多くの御意見、御指導を賜わった名古屋大学工学部池谷和夫教授に謹んで謝意を表する。

また本論文をまとめるにあたり懇切な御指導を賜わった名古屋大学工学部福村晃夫教授に深く謝意を表する。

更に名古屋大学工学部久野和宏助教授には研究遂行および論文の作成に対して機会あるごとに種々御教示いただいた。ここに謹んで謝意を表する。

本研究をまとめるにあたり、多くの貴重な御意見を賜わり、便宜をはかっていただいた三重大学工学部服部昭三教授に深く謝意を表する。

また実験に多大の協力をいただいた三重大学工学部山本好弘技官、ならびに名古屋大学工学部池谷研究室の諸氏に感謝する。

最後に種々の御指導を賜わった三重大学工学部電気系の諸先生方、ならびに実験に参加された各位に感謝する。

参 考 文 献

- (1) H. Haas; Über den Einfluss eines Einfachechos auf die Hörsamkeit von Sprache. *ACUSTICA*, Vol. No.2, p. 1 (1951)
- (2) J. P. A. Lochner, J. F. Burger; The Influence of Reflections on Auditorium Acoustics, *J. of Sound Vib.*, Vol. 1, No.4, p. 426 (1964)
- (3) J. P. A. Lochner, J. F. Burger; The Subjective Masking of Short Time Delayed Echoes by their Primary Sound and their Contribution to the Intelligibility of Speech, *ACUSTICA*, Vol. 8, No.1, p. 1 (1958)
- (4) 柳沢武三郎, 二村忠元; 合成音場における音質評価に関する実験, *日本音響学会誌*, Vol. 26, No.2, p. 67 (1970)
- (5) 柳沢武三郎, 二村忠元; 室内音響に関する物理量と心理量の対応に関する実験結果について(I), *日本音響学会誌*, Vol. 24, No.3, p. 111 (1968)
- (6) 柳沢武三郎, 二村忠元; 室内音響に関する物理量と心理量の対応に関する実験結果について(III), *日本音響学会誌*, Vol. 24, No.4, p. 187 (1968)
- (7) E. Meyer; Definition and Diffusion in Rooms, *J. A. S. A.*, Vol. 26, No.5, p. 630 (1954)
- (8) M. R. Schroeder, D. Gottlob, K. F. Siebrasse; Comparative Study of European Concert Halls, *J. A. S. A.*, Vol. 56, No.4, p. 1195 (1974)
- (9) H. Fletcher; *Speech and Hearing in Communication*, p. 278, Krieger (1972)
- (10) 三浦宏康; 拡声電話機, *電気通信学会誌*, Vol. 43, No.9, p. 1043 (昭35)
- (11) J. W. Emling; General Aspects of Hands-Free Telephony, *Comm. & Elect.*, Vol. 30, p. 201 (May 1957)
- (12) 大賀寿郎, 吉川昭吉郎; ハンドフリー電話機の鳴音条件に関する検討, *電気通信研究所研究実用化報告*, Vol. 19, No.9, p. 1755 (1970)
- (13) 大泉充郎監修; *音声科学*, p. 95, 東京大学出版会 (1972)
- (14) L. L. Beranek; *Acoustics*, McGraw-Hill (1954)
- (15) V. O. Knudsen, C. M. Harris; *Acoustical Designing in Architecture*, John Wiley & Sons (1950)
- (16) 境 久雄, 中山 剛; 聴覚と音響心理, p. 183, コロナ社 (1978)
- (17) J. P. Egan, E. C. Carterette, E. J. Thwing; Some Factors Affecting Multi-Channel Listening, *J. A. S. A.*, Vol. 26, No.5, p. 774 (1954)
- (18) J. W. Black; The Reception of Repeated and Overlapping Speech Patterns; *J. A. S. A.*, Vol. 27, No.3, p. 494 (1955)

- (19) W. Spieth, J. C. Webster; Listening to Differentially Filtered Competing Voice Messages, J. A. S. A., Vol. 27, No.5, p. 866 (1955)
- (20) J. C. Webster, L. N. Solomon; Effects of Response Complexity upon Listening to Competing Messages, J. A. S. A., Vol. 27, No.6, p. 1199 (1957)
- (21) J. C. Webster, L. Sharpe; Improvements in Message Reception Resulting from "Sequencing" Competing Messages; J. A. S. A., Vol. 27, No.6, p. 1194 (1955)
- (22) W. R. MacLean; On the Acoustics of Cocktail Parties; J. A. S. A., Vol. 31, No.1, p. 79 (1959)
- (23) R. Plomp; Acoustical Aspects of Cocktail Parties; ACUSTICA, Vol. 38, No.3, p. 186 (1977)
- (24) I. Pollack, J. M. Pickett; Stereophonic Listening and Speech Intelligibility against Voice Babble, J. A. S. A., Vol. 30, No.2, p. 131 (1958)
- (25) 電子通信学会編; 聴覚と音声, p. 245 コロナ社 (昭48)
- (26) 斎藤収三; 明りょう度試験法と明りょう度値の評価に関する考察, 電気通信研究所研究実用化報告, Vol. 8, No.6, p. 845 (1959)
- (27) J. P. Guilford; 精神測定法 (秋重義治訳) p. 146, 培風館 (1959)
- (28) 文献 (25) p. 393
- (29) 文献 (25) p. 412
- (30) 井 研治, 渡辺真吾; 人工残響音のプレファレンス評価について, 日本音響学会講演論文集, p. 369 (昭51, 5)
- (31) 井 研治, 渡辺真吾; 等プレファレンスによる残響音の通話品質について, 電子通信学会電気音響研究会, EA76-6 (1976)
- (32) 井 研治, 渡辺真吾; 残響の音声品質に及ぼす影響について, 電子通信学会論文誌, Vol. 60A, No.3, p. 316 (1977)
- (33) 坂本吉弘, 山口照男; 鉄板の自由振動を利用した人工残響付加装置について, NHK技術研究, Vol. 10, No.1, p. 39 (昭33)
- (34) G. R. Crane, G. A. Brookes; Artificial Reverberation Facilities for Auditoriums and Audio Systems, J. of Audio Eng. Soc., Vol. 9, No.2, p. 198 (1961)
- (35) M. R. Schroeder, R. F. Logan; 'Colorless' Artificial Reverberation, J. of Audio Eng. Soc., Vol. 9, No.3, p. 192 (1961)
- (36) M. R. Schroeder; Improved Quasi-Stereophony and "Colorless"

- Artificial Reverberation, J. A. S. A., Vol. 33, No. 8, p. 1061 (1961)
- (37) M. R. Schroeder; Natural Sounding Artificial Reverberation, J. of Audio Eng. Soc., Vol. 10, No. 3, p. 219 (1962)
- (38) 伊達 玄, 戸塚良則; 伝送特性と残響時間の周波数特性とを独立に制御できる残響付加方式, NHK技術研究, Vol. 17, No. 5, p. 309 (昭40)
- (39) S. Watanabe, K. Inomoto; Effect of Reverberation on Speech Quality Evaluated by Equi-Preference, 9-th International Congress on Acoustics p. 432 Madrid (1977)
- (40) W. A. Munson, J. E. Karlin; Isopreference Method for Evaluating Speech-Transmission Circuits, J. A. S. A., Vol. 34, No. 6, p. 762 (1962)
- (41) M. R. Schroeder; New Method of Measuring Reverberation Time, J. A. S. A., Vol. 37, No. 3, p. 409 (1965)
- (42) 渡辺真吾; 電話伝送品質のオピニオン評価についての検討, 電気通信研究所研究実用化報告, Vol. 21, No. 2, p. 271 (1972)
- (43) R. H. Bolt, P. E. Doak; A Tentative Criterion for the Short-Term Transient Response of Auditoriums, J. A. S. A., Vol. 22, No. 4, p. 507 (1950)
- (44) M. R. Schroeder; Computers in Acoustics, 6-th International Congress on Acoustics GP6-1, Tokyo (1968)
- (45) 山本照二; 複合反射音を対象としたエコーの検知限, 日本音響学会誌, Vol. 26, No. 3, p. 153 (1971)
- (46) 山本照二; スタジオの残響特性と聴感特性との対応, NHK技術研究, Vol. 15, No. 6, p. 429 (昭38)
- (47) 井 研治, 渡辺真吾; 残響による音声の品質劣化についての一考察, 電子通信学会電気音響研究会, EA77-7, p. 1 (1977)
- (48) 井 研治, 渡辺真吾; 残響による音声の品質劣化についての一考察, 日本音響学会誌, Vol. 34, No. 4, p. 242 (1978)
- (49) 井 研治, 渡辺真吾; 遅延時間による人工残響音の品質劣化について, 日本音響学会講演論文集, p. 249 (昭51. 10)
- (50) 井 研治, 渡辺真吾; 残響による音声の品質劣化と物理量の関連について, 日本音響学会講演論文集, p. 95 (昭52. 4)
- (51) 井 研治, 渡辺真吾; マルチエコーによる通話品質の劣化について, 日本音響学会講演論文集, p. 97 (昭52. 4)
- (52) 例えば早坂寿雄, 石井鈴枝, 吉川昭吉郎; 音響工学概論, p. 295, 日刊工業新聞社 (昭47)

- (53) 二村忠元, 奥田襄介, 城戸健一, 曾根敏夫; 電気音響工学Ⅱ p. 147, オーム社 (1963)
- (54) 田原靖彦, 曾根敏夫, 柳沢武三郎, 二村忠元; 直接音と同一方向から到来する単一エコーの検知限について, 日本音響学会講演論文集, p. 237 (昭43. 11)
- (55) 井 研治; 残響の明りょう度に及ぼす影響について, 電子通信学会電気音響研究会, Vol. 78, No.32, p. 9 (1978)
- (56) 渡辺真吾, 井 研治; 残響の明りょう度に及ぼす影響について, 日本音響学会誌, Vol. 35, No.3, p. 30 (1979)
- (57) R. H. Bolt, A. D. MacDonald; Theory of Speech Masking by Reverberation, J. A. S. A., Vol. 21, No.6, p. 577 (1949)
- (58) J. H. Janssen; A Method for the Calculating of the Speech Intelligibility under conditions of Reverberation and Noise, ACUSTICA, Vol. 7, p. 305 (1957)
- (59) 比企静雄; 連続音声中の各種の区分の持続時間の性質, 電子通信学会誌, Vol. 50, No.8, p. 1485 (昭42)
- (60) 文献 (25) p. 232
- (61) W. B. Davenport; An Experimental Study of Speech-Wave Probability Distributions, J. A. S. A., Vol. 24, No.4, p.390 (1952)
- (62) 文献 (25) p. 247
- (63) 松田亮一; 周波数特性に不規則性をもつ伝送系の通話品質に関する研究, 電気通信研究所研究実用化報告, Vol. 17, No.9, p. 49 (1968)
- (64) 河原田邦康, 石井直樹, 吉川昭吉郎; 残響音場の通話品質に関する2, 3の考察, 電子通信学会電気音響研究会, EA67-6 (1967)
- (65) 文献 (25) p. 349
- (66) 井 研治, 池谷和夫; 多重音声の統計的性質と通話妨害, 日本音響学会誌, Vol. 35, No.1, p. 9 (1979)
- (67) 藤崎博也, 大村俊郎; 連続音声における発声区間及び休止の性質について, 日本音響学会講演論文集, p. 221 (昭46. 11)
- (68) 例えば宮脇一男; 雑音解析, p. 52, 朝倉書店 (1961)
- (69) 三浦種敏; 日本語に対する定量的な伝送品質と伝送特性との関係, 電気通信研究所研究実用化報告, Vol. 3, No.4, p. 468 (1954)
- (70) Y. Horii, A. S. House, G. W. Hughes; A Masking Noise with Speech-Envelope Characteristics for Studying Intelligibility, J. A. S. A., Vol. 49, No.6, p. 1849 (1971)

以上