

小特集—音声の明瞭度と認識率—

空間内における音声情報伝達*

戸井田義徳 (東京都立大学工学部建築学科)**

1. はじめに

音声情報伝達を評価するための尺度は、まず電話の発達に伴って研究され、1940年代には明瞭度指数¹⁾が提案され、引き続き種々の客観的評価指標が提案された。その後、建築音響の分野でホール等空間内の評価用として、これらの尺度の適用拡大が工夫された。更に空間内で必然的に生じる残響や反射音による妨害に対応するため、STI等新たな尺度も提案され現代に至っている。本稿では未だ解決されたとはいえない空間内特有の音声情報伝達特性を解説し、合わせて現状の問題点について紹介する。

2. 音声情報伝達の評価尺度とその定義

音声情報伝達の評価尺度には、明瞭度を初めとして幾つかの尺度が提案されている。これらの尺度は被験者を用いて直接評価するものと、音場や音声等の物理特性から客観的に算定するものに分けられる。本章ではこれらの尺度について簡単に紹介する。なお、これまでに書かれた解説書も多く、詳しくはそれらを参照されたい。

2.1 被験者を用いた評価尺度

i) 明瞭度

明瞭度は無意味な音節を発声し、受聴者がその何%を正しく聞き取れたかを示す尺度である。日本語の場合音節を1音節ずつ区切って発声したものを単音節明瞭度、2音節又は3音節ずつ続けて発声したものを2(連)音節明瞭度、3(連)音節明瞭度等という。

ii) 了解度

了解度とは「発声・伝送された章句又は単語の

正しく聴取される場合²⁾であり、明瞭度との最大の相違点は、有意味の単語又は文章フレーズを用いる点である。発声又は集計の単位が単語の場合は単語了解度、文章の場合は文章了解度と呼ぶ。

単語了解度は、発声された単語のうち、正しく聞き取られた単語の割合である。文章了解度は質問文又は命令文を読み上げ、正しい解答を得られた文章数の割合で表す。また、文章了解度は、文章中の特定の単語(key wordと呼ぶ)の正聴率で表す場合もある。この場合、正確には「文章を用いた単語了解度」と呼ぶべきだが、外国文献を含め用語は必ずしも統一されていない。同一音場の場合、一般的には了解度の値が明瞭度よりも高い。これは有意味であることやコンテキストにより聴者側で類推が可能のためである。

iii) カテゴリ評価尺度

5~7段階の尺度を用いて、主観的に評価する尺度である。「聴き取り易さ」「明瞭性」などの用語をあてはめる場合が多い。この尺度は、音声の情報伝達以外の音質やアメニティをも表す尺度と考えられるが、他の尺度に比べばらつきが大きくなりがちであり、注意が必要である。

この尺度を用いた研究例は少ないが、図-1は「聴き取り易さ」とSTIの関係³⁾、図-2はシェッフエの一対比較法により求めた同じくSTIとの関係⁴⁾である。

2.2 客観的算定手法

音声のレベル、騒音レベル、残響時間等の物理量と明瞭度・了解度の関係は、古くはKnudsenとHarrisにより求められ、日本では久我により更に詳しく検討された。これらについては解説書もあり⁵⁾本節ではそれ以外の主な算定手法について述べる。また AL_{CONS} 等主としてアメリカで使用されることの多い尺度についても、別書⁶⁾を参照されたい。

* Speech intelligibility in sound fields.

** Yoshinori Toida (Department of Architecture, Tokyo Metropolitan University, Hachioji, 192-03)

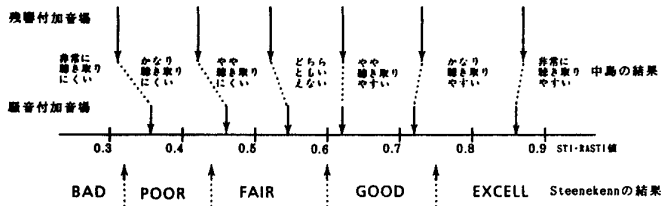


図-1 日本語音声の「聞き取り易さ」と STI の関係³⁾

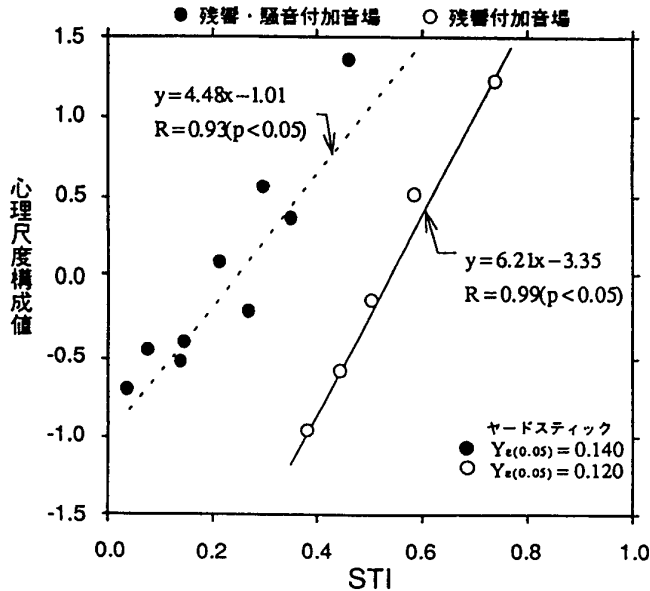


図-2 心理尺度構成値と STI の関係⁴⁾

i) 明瞭度指数

明瞭度指数は、本来電話受聴を対象として開発された。French と Steinberg¹⁾は、ベル電話研究所で行われた膨大な研究をまとめ、スピーチレベル、ノイズレベル、伝送周波数特性等の物理量から明瞭度を計算で求める方法を開発した。また Kryter はその適用範囲をスピーカ受聴にまで広げ、更に簡易計算法を提案した⁷⁾。

計算法は、図-3⁷⁾のようにチャート上に音声と騒音のレベル、システムの周波数特性等をプロットして求められる。計算原理は「音声周波数を幾つかのバンドに分割したとき、各バンドはそれぞれ独立して音声認識に貢献し、各バンドの貢献の総和が全体の音声認識に等しい」との仮説に基づいている。

実測値との相関は、非常に高ことが知られているが、原理的に残響音場には適用できない。ただし文献 7) には残響時間に対する修正量が提案されている。また音声の受聴レベルにもあまり対応していないらしく、特に音声のレベルが低い場合

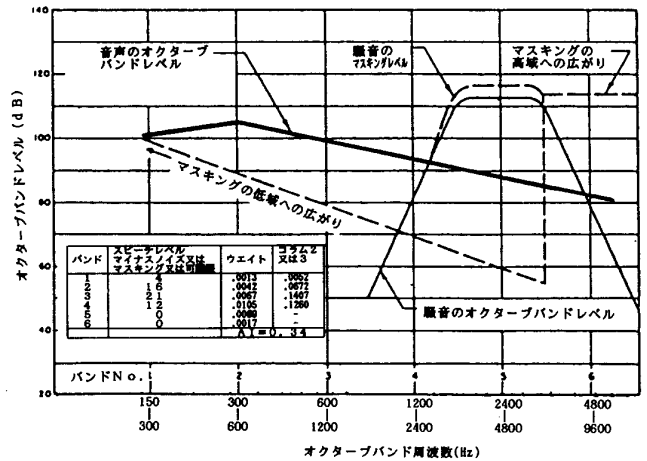


図-3 明瞭度指数計算例 (1/1 オクターブバンド法)⁷⁾

に誤差が多い傾向が見られる。低レベルスピーチに対する明瞭度指数は文献 8) を参照されたい。

ii) 初期反射音の貢献度を考慮した手法

Lochner と Burger⁹⁾はエコーと残響のある音場で明瞭度を測定し、ハース効果により delay time 95 ms 以内のエコーは明瞭度に貢献することを見出した。Latham¹⁰⁾は、それまでの研究を基に、95 ms までの初期反射音を useful、それ以降の反射音と暗騒音を有害 (detrimental) として両者の S/N から次式により明瞭度を求める方法を提案した。

$$S/N' = 10 \log_{10} \frac{\int_0^{95} \alpha_{(p,t)} p^2(f,t) dt}{\int_{95}^{\infty} p^2(f,t) dt + p_{PNC}^2 T} \text{ dB}$$

ただし、 f は 1 kHz (オクターブバンドの中心周波数)、 $\alpha_{(p,t)}$ は反射音の遅れ時間とレベルによる貢献度 (原文献では図示)、 $p(t)$ は瞬時音圧 (N/m²)、 t は直接音到達後の時間 (ms)、 T は明瞭度テスト音節の継続時間、 P_{PNC} は暗騒音の PNC 値 (L_{10})、 ∞ は実用上 1 秒間である。この予測式は多数のホールについて実測値との相関が高いことが指摘されている。

iii) STI, RASTI

Houtgast と Steeneken¹¹⁾は音場内では音声波形の包絡線が残響や騒音により変形することに着目し、100% 振幅変調した試験音を用いて MTF (modulation transfer function) の変化量から STI (speech transmission index) を求める手法を開発した。その後簡易測定法である RASTI が

提案され測定器も数種類が市販されている。STI, RASTI に関する詳しい解説は文献^{12),13)}を参照されたい。

この方式は理論的に明快なこと、残響と騒音の影響を同時に評価できること、コンピュータ処理に適し、現場で被験者なしに測定できる物理量であること等、実用上の利点が多く、現在では多くの機関で採用され、データ蓄積も多い。

3. 評価用音表・音源と音声伝達特性の関係

明瞭度や了解度を測定する場合、被験者に聴かせる音節や文章を集め表にする。これが明瞭度(了解度)測定用音表であり、それを読み上げたものが明瞭度(了解度)測定用音源と呼ばれる。

3.1 明瞭度測定用音表・音源

i) 単音節音表・音源

日本語の音節数は鼻濁音、撥音(ん)を除きほぼ100音節であり、これらをランダムに並べた音表を100音節平等率音表と呼ぶ。単音節で発声した場合、1音節の継続時間は200~500ms程度と短く、空間内で発声すると、語頭の子音は反射音による妨害を受けにくい。このため単音節音源は空間内の音声伝達特性評価には不向きである。

ii) 連音節音表・音源

連音節を空間内で発声した場合、第2音節以降の音節が先行音節の反射音によりマスクされる。しかし無意味な連音節では最大3連音節程度が限度であり、それ以上は聞こえても記入できず、長い継続時間の音源は作れない。発声レートも4~6シラブル/秒のため、大空間や残響時間の長い室内では、反射音の影響を受けにくく誤差が大となる。特定の発声レートであるための弊害を改善するた

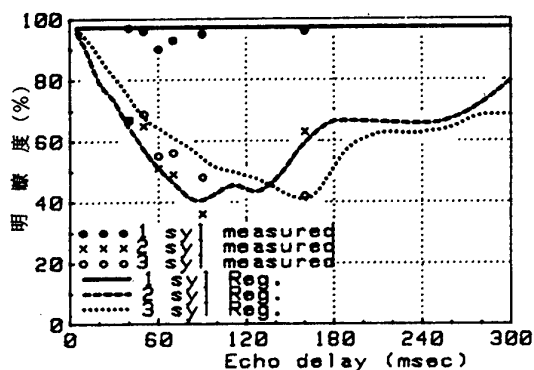


図-4 エコー妨害による明瞭度低下の例

めに、数種のレートを持つ音源も提案されている¹⁴⁾。図-4はディスクリートな3個の等間隔エコーがあるとき、3連音節中の各音節が受ける低下特性である。第2音節以降がエコーによる影響を受けていることが分かる。

iii) キャリアフレーズ、マスキング音源

キャリアフレーズとは音節の前後に続けて発声される章句のことであり、本来は発声時のレベル合わせと自然さを保つたものであったが、その後空間内の測定に用いられるようになった。マスキングはキャリアフレーズを人工的なノイズに置き換えたものである。音声と同一の周波数特性を持つランダムノイズを用いて、継続時間や包絡線も音声に近似させたものが多い。

3.2 了解度測定用音表・音源

単語を用いた音表には、2音節単語、3音節単語等があり、文章音表には、前述した質問文・命令

表-1 文章了解度音表の例

No.102	
1.	日本にあるスキー場の名を一つ書いて下さい
2.	数字の2と6をたすといくつですか
3.	食べ物の野菜の名を一つ書いて下さい
4.	日本では田植えは何月ごろに行いますか
5.	動物の豚の足の数は何本ですか
6.	あなたの知っている動物の名を一つ書いて下さい
7.	日本では札幌と鹿児島ではどちらが南にありますか
8.	動物の馬の鳴き声を書いて下さい
9.	今これを書いているあなたは男性ですか女性ですか
10.	食べ物のパンの名を一つ書いて下さい
.	.

表-2 無意味文の例¹⁴⁾

恐ろしい声	が散る
ほっそりした攻撃	が生える
太い建築	が終わる
細い電気	がめくる
近いタバコ	が揺れる
遠い痛み	が燃える
激しい事務所	が断わる
黒いコンサート	が飛ぶ
空いているパーティー	が落ちる
柔らかい戦争	が上がる
見えない俳優	が去る
弱い航海	が書く
.	.

空間内における音声情報伝達

文を用いたものと普通文を用いたものがある。また、コンテキストの影響をなくすため敢えて意味が通らないように作った無意味文・不自然な文も用いられる。表-1は質問文の例である。表-2¹⁴⁾は無意味文の1例で、文中の名詞の正答率を単語了解度とする。文章音源は、測定に時間を要すること、被験者の学習効果が大きい等、測定上の制約は多い。しかし実際の使用状況に最も近い音源であり、評価指標としては最も実情に近いものと考えられる。

4. 音声情報伝達と空間・音声特性の関係

空間内における音声の妨害は、音声反射音により互いに重なりあって起こるため、低下特性は空間のインパルスレスポンスと音声の物理特性の双方に影響される。

i) 音声特性

音声側の物理特性では、子音や母音のレベルや継続時間、連続音声では発声レート（話す速さ）やアクセントの位置が影響する。空間側の修正（ディレイマシ等）が不可能な場合（防災用屋外スピーカ等）は、音声の運用法が音声伝達にとって重要となるケースもある¹⁸⁾。

ii) 空間特性

反射音の時間あたりの密度、残響時間、ディスクリートのエコーの有無が関係する。反射音の密度が高く、かつ残響時間があまり長くない音場ではSTI・RASTIによる評価に適した例が多い。

iii) 音声と空間の相互関係

音声の時系列と反射音の時系列に相関があるとき、特定の音節や特定の発声レートだけ妨害（貢献）を受ける、特異点的な現象が起こるときがある。反射音による妨害は本来、特定の音節が特定の反射音により妨害を受けた結果であり、本質的に特異点の集合であるといつて良い。この種の妨害は、大空間やディスクリートのエコーのある時に起き、普通の閉空間内ではまれな現象ではあるが、音声と空間の特性がたまたまシンクロナイズしたときに起こる現象である。

5. STI・RASTI 測定・評価における問題点

STI・RASTIは音声情報伝達の客観的な予測・評価指標であり、実用的な利点も多く、現在特に

民間企業を中心にして幅広く利用され測定が行われている。しかしSTIの評価基準が必ずしも統一されていないにもかかわらず、設計仕様書にRASTIの目標値が記入される例があり、数値のひとり歩きの傾向が見え始めている。また、測定法の規格が不統一であるという問題点もある。本章（特に5.2, 5.3節）は、これらの問題点を踏まえて先頃行われた、建築学会明瞭度研究ワーキング・グループ（以後明瞭度WGと略す）の調査結果¹⁵⁾に基づき、筆者の意見を加えたものである。

5.1 STI・RASTIの原理的な問題点

STI・RASTIは、全ての反射音を有害としており、反射音による貢献については考慮されていない。また空間内に限定された物理量のため、発声側と受聴側の条件、例えば発声レベル、発声レート、受聴者の聴力等による違いには対応しない。また試験音の性質上、大空間のようにディレイタイムの長いエコーや、ディスクリートのエコーによる妨害にも対応できない。

5.2 測定法・評価法の問題点

i) 測定法の不統一

測定法のうち、音源用スピーカは、その指向性や拡声レベル、特にホールの電気音響設備を用いて測定する際の測定法が不統一であると言われている。また、測定マイクロホンについても指向性や設置高さ等に違いが見られる。STIを測定する際の音源は、定義どおり100%変調ノイズを用いる方法のほかに、インパルスレスポンスからSTIを求める方法があるが、インパルスレスポンス測定用音源や、取り込み時間等も、各機関によりまちまちである。このように現段階では、各機関で測定されたSTI・RASTI値を横並びにして評価するには疑問点が多い。

ii) 評価基準について

SteenekenによるSTI評価基準は図-1中に示した。しかしこの評価基準を実用的なすべての空間に適用するには問題点が多い。また、明瞭度・了解度とSTI・RASTIの相関関係を求めようとすると、更に問題点が多くなる。これは前述したSTI・RASTIの原理的な適用限界と、明瞭度・了解度の持つ適用限界の二つの原因が重なるためである。

実際のホールのSTI・RASTIの測定値は

0.45~0.6 (fair) の間に入るものが大半である¹⁵⁾。しかし聴感的には明瞭性に大差がある¹⁵⁾ことが分かっており、評価尺度の見直し又はケース分けが必要であろう。

3連音節明瞭度と STI の関係では、多種類の空間の測定値をまとめて相関図を作ると、ほとんど無相関となってしまう場合が多い。これは明瞭度測定用音源が大空間や残響過多の音場に適應しないことと STI の適用限界が重なったためと推定される。この点に関しては空間別、ケース別に測定評価法の見直し又は何等かのガイドラインの作成が必要であろう。

5.3 STI・RASTI の適用限界と新たな尺度への展望

明瞭度 WG の調査では、STI・RASTI の適用限界として、①音場の時間構造、空間構造を反映しない。②音声の物理特性、スピーチレートを反映しない、という指摘がなされ、これらを反映する新たな指標が必要であるという意見が出された。このような新たな指標実現の可能性について筆者の意見を述べてみたい。

空間内の音声伝達に関係する要素には、①音源(音声及び拡声系)、②空間、③受聴者、の3要素がある。これからの指標はこれら3要素のすべてに対応していることが望まれる。時間的に変動の多い音声と、不連続な時系列である空間のインパルスレスポンスの関係に対し、これまでのように平均化又は統計処理を用いた指標では、対応できないケースがどうしても生じてしまう。今後は空間内で発声された個々のフレーズの評価ができるような指標、つまり各音節の伝達特性にまで踏み込んだ新たな尺度(又は評価手法)が工夫されるべきであろう。このような視点に基づいた研究例はないわけではなく、古くは斎藤¹⁶⁾による個々の音韻特性と明瞭度の関係、最近では春田¹⁷⁾による音節の分類があり、またディスクリートのエコーのある音場を対象とした筆者の提案¹⁸⁾等がある。

6. おわりに

STI・RASTI は物理量であり、本来なら客観的な測定値が得られる点に特徴がある。しかし測定法の不統一等によるばらつきや、原理的に音源(音

声) と受聴側との対応に不備がある等問題点も多い。一部に起こりつつある数値のひとり歩き等には特に注意した対応が望まれる。本稿では受聴者側の要因についてはあまり触れることができなかったが、高齢者や災害時を対象とした研究等急務と言えよう。

文 献

- 1) N.R. French and J.C. Steinberg, "Factors governing the intelligibility of speech sounds," J. Acoust. Soc. Am. **19**, 90-119 (1947).
- 2) 日本音響学会明瞭度委員会, 明瞭度試験法の基準 (1956).
- 3) 中島立美, 前田節雄, "室内での音声の「聴き取り易さ」と STI," 音響学会建築音響研資 AA 84-30 (1984).
- 4) 佐藤 洋, 長友宗重, 矢島吉紀, "残響と騒音の明瞭度におよぼす影響とその評価に関する研究," 音響学会建築音響研資 AA 94-1 (1994).
- 5) 日本音材協会編, 建築音響工学ハンドブック (技報堂, 東京, 1961), pp. 181.
- 6) 持丸 聡, "アメリカにおける明瞭度評価のための予測計算, 聴感試験, 並びに物理測定手法に関する概要—第 33 回音シンポ," 音響学会建築音響研資 AA 90-17, 11-18 (1990.9).
- 7) K.D. Kryter, "Method for the calculation and use of the articulation index," J. Acoust. Soc. Am. **34**, 1689-1697 (1962).
- 8) Chaslav V. Pavlovic, "Derivation of primary parameters and procedures for use in speech intelligibility predictions," J. Acoust. Soc. Am. **82**, 413-422 (1987).
- 9) J.P.A. Lochner and J.F. Burger, "The intelligibility of re-inforced speech," Acustica **9**, 31-38 (1959).
- 10) H.G. Latham, "The signal-to-noise ratio for speech intelligibility—An auditorium acoustics design index," Appl. Acoust. **12**, 253-320 (1979).
- 11) T. Houtgast and H.J.M. Steeneken, "The modulation transfer function in room acoustics as a predictor of speech intelligibility," Acustica **28**, 66-73 (1973).
- 12) 小椋靖夫, 浜田晴夫, 三浦種敏, "音場における音声伝送品質のための MTF と STI," 音響学会誌 **40**, 181-191 (1984).
- 13) 中島立視, "音声の明瞭度指標 (STI) の測定," 音響学会誌 **49**, 103-110 (1993).
- 14) 橋本 修, 木村 翔, 宇津木淳一, "会話音声の発声レートを考慮した三連音節明瞭度試験用音源による室内音場の明瞭度評価について," 日建学会計画系論文集 **456**, 1-8 (1994).
- 15) 中島立視, "RASTI, STI の利用と現状の問題点," 日建学会室内音場小委資料, 音響学会建築音響研資 AA 94-17, 31-38 (1994.3).
- 16) 斎藤修三, "明瞭度試験法と明瞭度の評価に関する考察," 研究実用化報告 **8**, 845-853 (1959).
- 17) 春田千秋, "騒音場における音声伝達特性に関する研究," 学位請求論文 (1987).
- 18) 戸井田義徳, "長大なディレイタイムのエコーによる音声聴取妨害," 音響学会建築音響研資 AA-90-5 (1990).