

Valutazione dell'intelligibilità della parola: metodi a confronto.

A.Cocchi, A.Farina, R.Vezzali

Istituto di Fisica Tecnica - Facoltà di Ingegneria
Università di Bologna

Le ricerche e i numerosi metodi che sono stati elaborati e proposti per il calcolo dell'intelligibilità del parlato hanno come scopo quello di fornire ai tecnici, nell'ambito del processo progettuale, un mezzo attendibile di previsione, e quindi di ottimizzazione, delle caratteristiche acustiche dell'auditorio da realizzare; né sono da trascurare i tentativi di mettere a punto procedure che, sempre muovendo da tale calcolo, consentano di effettuare valutazioni di qualità sonora, basandosi sulla base misurazione di un parametro oggettivo non direttamente o completamente rappresentativo del problema specifico. Intensità del segnale, rumore di fondo, riverberazione e modalità di riflessione sonora in ambiente, sono tutti fattori generalmente accettati come determinanti sull'intelligibilità risultante del parlato: le controversie si concentrano invece sul contributo relativo di questi parametri e sulla loro interdipendenza nel complesso ambito del fenomeno uditivo.

Knudsen e Harris (1), con ricerche di natura molto empirica, misero in evidenza una relazione inversa tra intelligibilità e valore del tempo di riverberazione alla frequenza di 512 Hz, nel senso che quella decresceva all'aumentare di questo, risultando in generale buona in ambienti con tempo di riverberazione inferiore a 1,3 s. Una riverberazione adeguata non risultava comunque garanzia di buona intelligibilità, essendo necessario considerare anche gli altri fattori sopra indicati.

Il successivo lavoro di Kryter (2) delinea una procedura nella quale in sostanza gli spettri del segnale e del rumore di fondo sono utilizzati per definire un Indice di Articolazione i cui valori calcolati sono stati sperimentalmente collegati ai diversi livelli di intelligibilità attraverso PB-words tests.

Alle diverse bande di frequenza, in ottave o terzi d'ottava, i vari livelli di pressione di picco, opportunamente definiti, vengono depurati dei corrispondenti livelli del rumore di fondo. I contributi in tal modo determinati vengono moltiplicati per opportuni fattori di peso, tabulati in funzione della frequenza, e infine sommati per determinare il valore dell'Indice di Articolazione globale.

La limitazione del metodo AI, sulla quale maggiormente si è dibattuto, riguarda la sua applicabilità al caso di ambienti riverberanti. L'accuratezza delle previsioni ottenibili con questo metodo per distanze inferiori a quella critica (campo sonoro diretto) è infatti ormai ben documentata, mentre esso è risultato essere meno accurato nella previsione dell'intellegibilità all'interno del campo riverberante.

Barducci (3), esaminando le differenze tra i risultati del calcolo dell'AI eseguito secondo le proposte di Kryter e Lehmann (4) che riguardano essenzialmente le bande di frequenza da utilizzare nella procedura, propone anche un metodo grafico per il calcolo dell'AI, e quindi dell'intellegibilità, secondo le due tecniche, evidenziando la maggiore cautelatività del metodo di Lehmann e la sua migliore attendibilità, soprattutto in riferimento alla lingua italiana. Viene inoltre presa in esame una valutazione di tipo grafico dell'AI di un ambiente: in pratica questo procedimento consente di prevedere con buona approssimazione il livello di intellegibilità per una conversazione che si svolga all'interno di un ambiente delimitato da pareti con isolamento conforme a quello previsto dalle norme, quando all'esterno vi sia un rumore noto, ma non generato da uno speaker, perché in questo caso il comportamento filtrante delle pareti può differire dalla curva limite ISO e quindi non essere adeguatamente attenuato da un isolamento a norma.

Per valutare l'intellegibilità di un messaggio parlato che perviene all'ascoltatore attenuato per l'attraversamento di una parete, Brosio (5) propone un metodo grafico basato sulla conoscenza delle caratteristiche acustiche dell'ambiente, del livello sonoro della sorgente e del rumore di fondo. L'AI così calcolato rappresenta quindi un valido elemento di qualificazione dell'isolamento acustico di una parete divisoria.

Un diverso approccio al problema dibattuto è quello proposto da Houtgast e Steeneken (6), che notano come l'effetto smorzante dell'ambiente sull'involuppo del segnale, tradizionalmente quantificato attraverso la curva di riverberazione, non è più ben rappresentato dal valore del tempo di riverberazione qualora l'andamento della curva devii sostanzialmente da quello ideale esponenziale. Questi Autori hanno suggerito di descrivere l'azione dell'ambiente come quella di un filtro attivo sul segnale in ingresso. Questa caratteristica filtrante (MTF-Modulation Transfer Function) è specifica di un dato ambiente e applicabile a qualunque segnale in ingresso. Poiché il metodo proposto si riferisce in particolare all'intellegibilità del parlato, il segnale in ingresso preso in considerazione è sinusoidale modulato: la riverberazione di un ambiente causa un decremento della profondità di modulazione in funzione della frequenza di

modulazione applicata: la linea di trasmissione è allora caratterizzata da questo decremento attraverso l'Indice di Modulazione $m(f)$ il quale, tenendo implicitamente conto anche del rapporto segnale/disturbo caratteristico dell'ambiente, può essere espresso matematicamente con la relazione:

$$m(f) = \frac{1}{1 + 0,455 \left(\frac{f}{F} \right)^2} \quad (1)$$

nella quale F è la frequenza di modulazione e T il tempo di riverberazione. Se, come accade normalmente, le azioni disturbanti sono funzione della frequenza, la determinazione della MTF va condotta per bande di frequenze. Stabilito le frequenze della portante sinusoidale e quelle di modulazione da esaminare, l'ambiente risulta così caratterizzato da una griglia di valori di m . Ciascuno di questi viene convertito in un S/N apparente, indipendentemente dal reale tipo di disturbo che causa il valore di m , secondo la relazione:

$$(S/N)_{app} = 10 \log(m/1-m)$$

Questi valori, o meglio il loro valore medio opportunamente normalizzato, vengono infine ricollegati in un singolo "Speech Transmission Index STI" dalla relazione:

$$STI = 1/90 (S/N)_{mapp} + 15$$

il cui valore risulta, per definizione, compreso fra 0 e 1. Il passaggio finale, cioè la correlazione fra il valore univoco STI e il giudizio soggettivo di intellegibilità, è stato effettuato servendosi dei già citati PB-words tests.

Per una più rapida determinazione dell'indice STI, essendo del resto stata notata una certa ridondanza del numero di dati connessi con il suo calcolo, è stata messa a punto una apparecchiatura elettronica che calcola automaticamente l'indice riferendosi a due sole ottave e ad un set ridotto di frequenze di modulazione: accettata la scelta metodologica, questa soluzione strumentale non è, ovviamente, l'unica disponibile, anche se sicuramente risulta la più pratica per operare in campo. Il valore così ottenuto è detto RASTI (Rapid STI) (7). Sono state altresì messe a punto (8) delle procedure diagnostiche delle cause dei principali difetti acustici, basate sul confronto ragionato fra le curve MTF tracciate per il caso ideale di decadimento esponenziale e quelle realmente rilevate.

Qui di seguito vengono presentati i risultati ottenuti dalla applicazione dei due metodi di calcolo di cui sopra ad un medesimo ambiente, precisamente alla chiesa di S.Gabriele in località Isola del Gran Sasso (AQ).

Le tabelle allegate riportano una mappatura dei valori AI e RASTI nelle diverse posizioni di prova individuate in un braccio della chiesa, a forma di croce greca, oltre ai valori del tempo di riverberazione rilevati con tecnica impulsiva nelle medesime posizioni.

Si può notare un accordo sostanzialmente buono fra i risultati, tenendo conto anche delle caratteristiche riverberanti dell'ambiente che, come già evidenziato, penalizzano la qualità del metodo AI.

Dopo una serie di indagini condotte su un certo numero di sale canadesi, Bradley (9) ha messo in evidenza la centralità di quattro tipi di parametri nella definizione delle caratteristiche di intellegibilità di una sala, tutti ricavabili dal rilevamento della risposta d'impulso dell'ambiente stesso: il livello di pressione sonora globale G, il rapporto energia diretta/energia ritardata C80 (come indicatore della chiarezza), uno fra i valori EDT o TR e infine un misuratore dell'intensità dell'energia laterale riflessa (LF o IADD). Lo studio citato propone anche una semplificazione del parametro esprimente il rapporto fra energia utile e energia sonora sviluppato da Lochner e Burger (10), che potrebbe essere sostituito da una grandezza denominata U80 così definita:

$$U80 = 10 \log_{10} \left(\frac{C80}{1 + (C80 + 1) I_r / I_d} \right) \quad \text{dB}$$

dove I_r e I_d sono rispettivamente le intensità sonore del rumore di fondo e del parlato. Dopo prolungate campagne di prova, risulterebbe dimostrato che questa grandezza è in grado di fornire previsioni della intellegibilità con una accuratezza almeno pari a quella dello STI, senza scatto nella semplicità di calcolo, tanto più che oggi giorno l'elaborazione dei dati sperimentali viene di norma affidata ad idonei calcolatori o a mano dedicati.

L'aspetto particolarmente interessante di questo risultato sta nel fatto che U80 si presta indifferentemente alla valutazione di sale sia per il parlato che per la musica.

Un'altra proposta interessante in tal senso riguarda la possibilità di utilizzare come descrittore la differenza fra i valori del D80 misurato e valutato ipotizzando all'interno della sala un decadimento esponenziale: questo parametro consentirebbe infatti di quantificare il grado di diffusione del suono nella sala. Ambienti dove questa differenza è grande (maggiore di +20dB) presentano forte energia diretta e di prima riflessione e sono da considerare buoni per il parlato ma non per la musica (direct sound halls). Se il valore del parametro è compreso fra 0 e +20dB la sala ha caratteristiche che meglio si adattano all'ascolto della musica (diffuse

sound hall); infine, valori molto negativi del parametro denotano situazioni di forte carenza (muddy hall).

Al crescere del valore della grandezza citata, il valore dell'EDT tende a diminuire rispetto a quello del classico TR.

Da quanto visto fino ad ora risulta chiaro che non esiste un singolo metodo perfetto per prevedere l'intelligibilità. Ogni procedura sembra presentare delle limitazioni relativamente alle condizioni sotto le quali viene sviluppata e alle curve empiriche che mettono in relazione i parametri fisici e l'articolazione stessa.

La procedura AI, sviluppata in condizioni non riverberanti, fornisce i risultati più accurati a distanze dalla sorgente sonora inferiori a quella critica; il metodo STI invece prevede condizioni anche riverberanti, mentre nessuno dei metodi qui citati considera le caratteristiche spettrali della riverberazione del rumore, che anche in condizioni poco riverberanti risultano molto importanti.

Un altro aspetto di considerevole disaccordo tra i vari A. è quello riguardante il contributo all'intelligibilità delle riflessioni discrete: Lochner e Burger (10) stabilirono che tutte le riflessioni con tempo di ritardo inferiore a 30 ms vengono pienamente integrate dall'orecchio al suono diretto, mentre quelle "inferiori" a 95 ms si integrano solo parzialmente; Peutz sostiene invece (11) che qualunque riflessione, a prescindere dal suo tempo di ritardo, va a deprimere l'intelligibilità.

Infine, non va dimenticato che l'attuale tecnologia prevede sempre per il parlato, sempre più spesso anche per la prosa, l'assistenza di un apparato di amplificazione e di diffusione sonora, dal quale deriva un valido valore del segnale in ogni ordine di posti: in quest'ottica, il rapporto segnale/rumore viene a perdere molto del suo significato originario, mentre permane l'importanza della riverberazione, alla quale si aggiunge un fatto contingente, vale a dire l'introduzione di linee di ritardo nella distribuzione del suono. Accade così che impianti mal fatti consentano ad un ascoltatore di ricevere due distinti messaggi sonori con tempi di ritardo inaccettabili: di fronte a situazioni del genere, l'impiego della tecnica STI automatizzata, con segnale trasmesso direttamente attraverso l'impianto di amplificazione, consente di valutare nel migliore dei modi la qualità di quest'ultimo, proprio in quanto la funzione di trasferimento viene così a risentire automaticamente, ed in modo che rispecchia fedelmente la realtà, del filtro attivo artatamente introdotto dall'uomo lungo la linea di trasmissione sorgente-ascoltatore.

A conclusione, si richiama l'attenzione sull'attuale avvio della procedura di revisione della norma ISO 3382-1975(E) che dovrà recepire i più recenti risultati della ricerca nel settore della qualificazione delle sale da spettacolo in generale e degli auditori in particolare: in questo frangente, il parametro "percettibilità della parola" non dovrà essere trascurato, ma anzi il suo impiego dovrà essere enfatizzato e sfruttato al meglio. Lo stato di stallo in cui si trova poi lo studio per una normativa in tema di valutazione delle caratteristiche di isolamento acustico con metodi rapidi potrebbe dare spazio a metodi, quale quello già ricordato (5), che utilizzano in tal senso valutazioni sperimentali di "percettibilità della parola".

BIBLIOGRAFIA

- (1) V.O.Knudsen,C.M.Harris - Acoustical designing in architecture - Wiley,New York,1950;
- (2) K.D.Kryter - Methods for the calculation and use of the articulation index - J.A.S.A.,1689,34,1964;
- (3) I.Barducci,A.De Lieto Vollaro,M.Felli - La valutazione dell'indice di articolazione in acustica ambientale - Rivista Italiana di Acustica, vol.VII-2,1983;
- (4) R.Lehmann - Effect de masquage et indice de netteté - Revue d'Acoustique,167,1,1968;
- (5) E.Bosio - L'indice di articolazione come parametro di valutazione dell'isolamento acustico fra ambienti - Rivista Italiana di Acustica, vol.VII-1,1983;
- (6) T.Houtgast,H.J.M.Steneeken - Envelope spectrum and intelligibility of speech in enclosures - IEEE-AFCRL Speech Conference,1972;
- (7) S.J.Orfield - The RASTI method of testing relative intelligibility - Sound and Vibration,12,1987;
- (8) Bruel & Kjaer Ed. - Manuale d'uso del RASTI;
- (9) J.S.Bradley - Progress in auditorium acoustics measurements - 12th. ICA,1986;
- (10) J.P.A.Lochner,J.F.Burger - Influence of reflections on auditorium acoustics - J.S.V.,1,4,1964;
- (11) V.M.A.Reutz - Articulation loss of consonants influenced by echo ,noise and riverberation - Symposium on Speech Intelligibility,Liege,1973.

Lavoro eseguito con il contributo finanziario M.P.I. 60% .

A.I. - S. GABRIELE

STAMPA RA.S.T.I. - S. GABRIELE

21	14	7
0.53	0.47	0.51
20	13	6
0.35	0.31	0.32
19	12	5
0.34	0.30	0.30
18	11	4
0.32	0.29	0.30
17	10	3
0.31	0.28	0.28
16	9	2
0.30	0.29	0.29
15	8	1
0.30	0.29	0.29

21	14	7
0.47	0.52	0.38
20	13	6
0.48	0.46	0.44
19	12	5
0.37	0.37	0.37
18	11	4
0.34	0.33	0.37
17	10	3
0.36	0.35	0.35
16	9	2
0.34	0.32	0.32
15	8	1
0.35	0.35	0.33

STAMPA TEMPI DI RIVER

L5100 1 2 3

1 SX 5.76 5.66 5.70
DX SEC SEC SEC

2 SX 5.7 5.75 5.67
DX SEC SEC SEC

3 SX 5.70 5.65 5.68
DX SEC SEC SEC

4 SX 5.73 5.71 5.7
DX SEC SEC SEC

5 SX 5.76 5.7 5.7
DX SEC SEC SEC

6 SX 5.7 5.7 5.7
DX SEC SEC SEC

7 SX 5.68 5.7 5.7
DX SEC SEC SEC

STAMPA RA.S.T.I. - S. GABRIELE

21	14	7
0.47	0.52	0.38
20	13	6
0.48	0.46	0.44
19	12	5
0.37	0.37	0.37
18	11	4
0.34	0.33	0.37
17	10	3
0.36	0.35	0.35
16	9	2
0.34	0.32	0.32
15	8	1
0.35	0.35	0.33

A.I. - S GABRIELE

21 0.53	14 0.47	7 0.51
20 0.35	13 0.31	6 0.32
19 0.34	12 0.30	5 0.30
18 0.32	11 0.29	4 0.30
17 0.31	10 0.28	3 0.28
15 0.30	9 0.29	2 0.29
15 0.30	8 0.29	1 0.29

STAMPA TEMPI DI RIVER

LS100	1	2	3
1	SX 5.76	5.66	5.70
	DX SEC	SEC	SEC
2	SX 5.7	5.75	5.67
	DX SEC	SEC	SEC
3	SX 5.70	5.65	5.68
	DX SEC	SEC	SEC
4	SX 5.73	5.71	5.7
	DX SEC	SEC	SEC
5	SX 5.76	5.7	5.7
	DX SEC	SEC	SEC
6	SX 5.7	5.7	5.7
	DX SEC	SEC	SEC
7	SX 5.68	5.7	5.7
	DX SEC	SEC	SEC