RENDICONTI
del
SEMINARIO MATEMATICO
della
UNIVERSITÀ DI PADOVA

CARLO MINNAJA
GIORGIO PINI
GIOVANNI ZILLI

Un modello additivo per matrici asimmetriche di confusione

Rendiconti del Seminario Matematico della Università di Padova, tome 72 (1984), p. 357-372
<http://www.numdam.org/item?id=RSMUP_1984__72__357_0>

REND. SEM. MAT. UNIV. PADOVA, Vol. 72 (1984)

Un modello additivo per matrici asimmetriche di confusione.

CARLO MINNAJA - GIORGIO PINI - GIOVANNI ZILLI (*)

SUMMARY - Confusion matrices, widely used in audiometrics, show asymmetries which become stronger in conditions of filtering or masking of the sound. Such asymmetries were usually neglected, and the matrices were forced to be symmetric a priori. In this paper a model is presented to study the causes of such asymmetries, under the hypotheses that these asymmetries are due to additive distortion terms, correlated to the frequencies of occurrences of the concerning phonemes in the spoken language. A calculus based on statistics of occurences in spoken Italian confirms the utility of the additive model and also of the hypotheses; goodness of fit is specified.

SUNTO - Le matrici di confusione, ampiamente usate in audiometria, presentano asimmetrie che si accentuano in condizioni di filtraggio o di mascheramento del suono. Tali asimmetrie solitamente venivano trascurate, e le matrici in questione venivano simmetrizzate artificiosamente a priori. In questo lavoro viene proposto un modello per studiare i motivi di tali asimmetrie, sotto l'ipotesi che esse siano dovute a termini additivi contenenti funzioni di distorsione, correlati con la frequenza di comparizione dei fenomeni nella lingua parlata. Un computo basato su statistiche di comparizione nella lingua italiana conferma la validità di principio del modello additivo e delle ipotesi fatte, precisando i limiti di adattabilità dello stesso ai dati presi in esame.

(*) Indirizzo degli A.: Istituto di Matematica Applicata, Università di Padova.

Lavoro eseguito nell'ambito del G.N.I.M.-C.N.R. (Progetto finalizzato INFORMATICA - Sottoprogetto SOFMAT).

0. Introduzione.

Le matrici di confusione sono particolari matrici di dissimilarità usate in acustica e foniatria; esse hanno per elementi a_{ij} i risultati di un esperimento che consiste nel far riconoscere a vari ascoltatori alcuni stimoli acustici prodotti e ascoltati in varie situazioni. L'elemento a_{ij} è il numero di volte in cui lo stimolo i è stato riconosciuto come j; l'indice j viene detto *risposta* allo stimolo. Trattandosi di stimoli acustici predefiniti, l'insieme I in cui si fanno variare gli stimoli è lo stesso di quello delle possibili risposte. Solitamente I è un alfabeto fonetico.

La produzione degli stimoli può essere effettuata in situazioni diverse: ad esempio attraverso un microfono, o in ambiente rumoroso, o da persone affette da dislalia. A sua volta l'ascolto può essere effettuato attraverso cuffia telefonica, o in ambiente rumoroso, o da persone non normoacusiche.

L'utilizzazione delle matrici di confusione è plurima: essa va dalla normalizzazione di apparecchiature telefoniche, alla valutazione della fedeltà di sistemi di trasmissione, alla diagnostica e alla pratica terapeutica ai fini della rieducazione della parola o dell'udito.

A partire dalle matrici di confusione, come a partire da qualsiasi matrice di similarità o di dissimilarità, si può istituire una funzione distanza tra gli stimoli. Tale distanza può essere poi confrontata con altre distanze istituite con altri metodi, a priori o sperimentali.

Si pongono quindi due ordini di problemi: il primo è istituire un modello matematico nel quale le probabilità di confusione siano poste in relazione con una distanza tra gli stimoli; il secondo problema è studiare le inevitabili deviazioni dei risultati sperimentali da questo modello.

Particolare interesse hanno le deviazioni dovute al fatto che le matrici di confusione sono in generale asimmetriche.

0.1 L'asimmetria delle matrici di confusione è stata spesso ritenuta irrilevante ai fini della similarità degli stimoli, e si possono proporre vari metodi per eliminare tale asimmetria. Il metodo più semplice, e l'unico usato effettivamente, consiste nel porre

$$a_{ij} = \overline{a_{ij}} = \frac{a_{ij} + a_{ji}}{2}\,; \tag{0.1}$$

si agisce quindi sulle matrici simmetrizzate per istituire una distanza.

Tuttavia può essere interessante studiare l'asimmetria e non semplicemente eliminarla.

Per interpretare l'asimmetria sono stati proposti vari tipi di modelli fondamentalmente suddivisibili in due grandi categorie. Nella prima vengono classificati i modelli che considerano l'aspetto simmetrico e quello asimmetrico come parti inscindibili di uno stesso processo fondamentale; nella seconda vi sono quelli che distinguono gli aspetti simmetrici da quelli antisimmetrici tramite una analisi appropriata.

Lo scopo di un modello per le matrici di confusione è quello di predire le asimmetrie della matrice a_{ij}. I vari modelli della seconda categoria prospettati in letteratura (vd. bibliografia di [1]) utilizzano una funzione di due variabili x ed y, simmetrica rispetto ad esse, con distorsioni apportate da due funzioni, una della variabile x e l'altra della variabile y. Lo scopo di questo lavoro è di proporre un modello in cui le funzioni di distorsione siano funzioni della distanza tra gli stimoli, distanza che è a sua volta funzione delle caratteristiche fonetiche, acustiche, o articolatorie dei suoni in questione. Alcune considerazioni di minimo hanno condotto, tramite una tecnica di branch and bound, alla scelta di un modello determinato, la cui validità è stata verificata su un computer Olivetti P 6066.

1. Definizioni e risultati preliminari.

1.1 Poichè nei pochi lavori finora apparsi non sempre è ben chiaro sotto quali ipotesi sono validi i modelli proposti, e neanche le definizioni sono sempre poste in maniera univoca, diamo qui una nomenclatura a cui ci riferiremo in seguito.

1.2 Dato un esperimento di riconoscimento di stimoli acustici $x \in X$, indichiamo con $p(x, y)$ la frequenza con cui allo stimolo x è stata data la risposta y, con $y \in X$; $p(x, y)$ si intende normalizzata, con $\sum_{y \in X} p(x, y) = 1$. Gli elementi $p(x, y)$ sono collocati in una matrice p_{xy} che in generale non risulta simmetrica. I modelli che qui esamineremo dovranno pertanto adattarsi in modo che vedremo alla matrice $p(x, y)$.

1.3 Def. Dicesi *modello metrico a distorsione di risposta* di una matrice di confusione asimmetrica p_{xy} la matrice

$$b(y)\, s_{xy} \tag{1.1}$$

quando s_{xy} è una matrice simmetrica. La funzione $b(y)$ della sola risposta y è detta *funzione di distorsione di risposta.*

1.4 DEF. Dicesi *modello metrico a distorsione di stimolo* di una matrice di confusione asimmetrica p_{xy} la matrice

$$a(x)\sigma_{xy} \tag{1.2}$$

quando σ_{xy} è una matrice simmetrica. La funzione $a(x)$ del solo stimolo x è detta *funzione di distorsione di stimolo.*

1.5 OSS. I due modelli sopra definiti sono equivalenti dal punto di vista matematico. Si potrebbe proporre un modello che consideri una distorsione dipendente sia dalla risposta che dallo stimolo, cioè un modello del tipo

$$a'(x)b'(y)\sigma_{xy}\,. \tag{1.3}$$

Tuttavia tale modello non è diverso da uno a distorsione di risposta o da uno a distorsione di stimolo. Infatti (vd. [2])

$$a'(x)b'(y)\sigma_{xy} = \frac{a'(y)}{a'(y)}b'(y)a'(x)\sigma_{xy} = \frac{b'(y)}{a'(y)}[a'(y)a'(x)\sigma_{xy}] \tag{1.4}$$

e il fattore entro parentesi quadra dà luogo ad una matrice simmetrica σ'_{xy}, e quindi si ha un modello a distorsione di risposta del tipo (1.1). Pertanto la componente asimmetrica è esprimibile completamente tramite una funzione di y. Un ragionamento perfettamente analogo conduce alla conclusione che un modello del tipo (1.3) è riconducibile anche ad un modello del tipo (1.2) a distorsione di stimolo.

1.6 OSS. I modelli (1.1) e (1.2) sono equivalenti, ma le loro parti simmetriche hanno significato diverso. Tuttavia un alto valore $\sigma(x, x)$ sulla diagonale principale indica che x è un suono chiaro, che viene confuso raramente sia come stimolo che come risposta, mentre un valore basso indica una confondibilità relativamente facile.

1.7 OSS. Entrambi i modelli forniscono sulla diagonale principale $b(x)s_{xx}$. Tuttavia il modello a distorsione di risposta è più utile di quello a distorsione di stimolo ai fini di questa predizione. Infatti, sia, in una matrice di confusione, $p(x_1, x_1) > p(x_2, x_2)$. In un modello a distorsione di risposta ciò può essere predetto in uno dei seguenti due modi:

$$b(x_1) > b(x_2) \,. \tag{1.5}$$

Ciò significa che x_1 si presenta come risposta più frequentemente che non x_2; segue $s(x_1, x_1) > s(x_2, x_2)$ e quindi la disuguaglianza nella matrice simmetrica è nello stesso verso della disuguaglianza nella matrice di confusione; oppure

$$b(x_1) < b(x_2) \,; \tag{1.6}$$

in tal caso, affinchè sia

$$b(x_1)\, s(x_1, x_1) > b(x_2)\, s(x_2, x_2)$$

dovrà essere a maggior ragione

$$s(x_1, x_1) > s(x_2, x_2) \,.$$

Pertanto la disuguaglianza nella matrice simmetrica va nello stesso verso di quella della matrice di confusione.

Se invece il modello è a distorsione di stimolo, sia ancora $p(x_1, x_1) >$ $> p(x_2, x_2)$; ciò comporta che x_1 è meno confondibile di x_2; pertanto si può supporre che $b(x_1)$ abbia valori bassi. Ne seguirebbe che $s(x_1, x_1)$ è inversamente correlato con $b(x_1)$.

1.8 Def. Dicesi *modello monotono* (o *non metrico*) il prodotto

$$b(x)\, s(x, y) \tag{1.7}$$

quando vale la seguente implicazione:

$$\{s(x, y) \geqslant s(w, z),\ b(x) \leqslant b(w),\ b(y) \geqslant b(z)\} \Rightarrow \Rightarrow b(x)\, s(x, y) \geqslant b(w)\, s(w, z) \,. \tag{1.8}$$

Questo modello effettua una predizione sulla disuguaglianza tra gli elementi $p(x, y)$ sotto le condizioni indicate in (1.8); se una di queste non è verificata il modello non effettua predizioni. La funzione di distorsione

b può essere interpretata sia come distorsione di stimolo che come distorsione di risposta: semplicemente gli elementi x per i quali $b(x)$ è alto si sostituiscono spesso agli elementi y per i quali $b(y)$ è basso. Poichè il modello monotono non prende in considerazione la diagonale principale, non si pone il problema considerato con il modello metrico a distorsione di stimolo.

Il modello metrico può sembrare più appropriato di quello semplicemente monotono, e in generale lo è. Ciò tuttavia non è vero per tutti gli insiemi di dati; per alcuni infatti le funzioni b ed s fornite dal modello metrico possono rendere non valida l'implicazione(1.8) in un numero non irrilevante di casi.

Goldstein (1980, [2]) effettua un paragone tra i due modelli inserendovi i dati di Wang e Bilger (1973, [3]), e sottopone a verifica l'ipotesi che le confusioni siano correlate con la frequenza (le frequenze usate per i singoli fonemi sono quelle di Carterette e Jones [4] e di Roberts [5]). Poichè la relazione tra frequenza e distorsione non sembra lineare, Goldstein usa il τ di Kendall evidenziando così solo correlazioni tra ranghi. Tuttavia la correlazione tra le distorsioni e le frequenze lessicali non sembra sempre significativa; sembra invece potersi concludere che maggiore è l'adattamento ai dati della funzione di distorsione, maggiore è la correlazione della distorsione con la frequenza. Goldstein fornisce una ulteriore interpretazione della distorsione, ponendola in relazione con una non meglio definita « naturalezza fonologica » basandosi su uno studio di Ruhlen (1975, [6]), effettuato su 700 linguaggi; tuttavia non è sempre facile separare la « naturalezza fonologica » di certi fenomeni dalla frequenza dei fonemi o dei gruppi di fonemi coinvolti. La conclusione è che la correlazione tra distorsione e frequenza può essere significativa nel caso delle consonanti occlusive e nasali, ma è molto debole e poco interpretabile negli altri casi.

2. Il modello additivo.

2.1 DEF. Dicesi *modello additivo* per una matrice asimmetrica p_{xy} la matrice

$$s_{xy} + a(x) + b(y) \tag{2.1}$$

quando s_{xy} è una matrice simmetrica.

Le funzioni $a(x)$ e $b(y)$ sono funzioni dipendenti rispettivamente dallo stimolo e dalla risposta.

Questo modello traduce il principio di sovrapposizione delle distorsioni, che ricalca il principio della sovrapposizione degli effetti in acustica. Un'effettiva valutazione dell'attendibilità del modello è stata effettuata, ed i risultati sperimentali sono riportati nel § 4.

3. Correlazione tra s_{xy} e la distanza fra fonemi.

3.1 Per interpretare da un punto di vista matematico la matrice s_{xy} che interviene nella definizione di un modello additivo, in [7] è stata proposta l'ipotesi che tale matrice sia correlata ad una distanza tra fonemi. L'ipotesi appare particolarmente plausibile se la correlazione è espressa dall'equazione

$$s(x, y) = \frac{k_1}{k_1 + d(x, y)} \tag{3.1}$$

dove $d(x, y)$ è una funzione distanza tra i fonemi x e y.

3.2 Def. Una funzione $\delta: I \times I \to R$ si dice *dissimilarità* se soddisfa alle seguenti proprietà:

a) $\delta(x, y) \geqslant 0 \quad \forall x \in I,\ y \in I$

b) $\delta(x, x) = 0 \quad \forall x \in I$

c) $\delta(x, y) = 0 \Rightarrow x = y$.

Altre volte, ad esempio in [8], viene richiesta alla funzione dissimilarità anche la proprietà di simmetria

d) $\delta(x, y) = \delta(y, x)$

che tuttavia spesso non è soddisfatta in matrici sperimentali, e in particolare non lo è in matrici di confusione.

3.3 Def. Data una matrice p_{xy} ed un modello, si dice *sforzo* (o *stress*) la quantità

$$\frac{\{\sum_x \sum_y [p_{xy} - s_{xy}]^2\}^{\frac{1}{2}}}{\{\sum_x \sum_y s_{xy}^2\}^{\frac{1}{2}}} . \tag{3.2}$$

3.4 Il problema dell'adattabilità del modello additivo sarà pertanto un problema di minimizzazione dello sforzo. Ciò dipenderà dalla matrice s_{xy}, e quindi, se si fa l'ipotesi che s_{xy} sia quella definita in 3.1, dal tipo di distanza $d(x, y)$.

In letteratura sono stati proposti vari tipi di distanza tra fonemi; aspetti teorici e metodologici nell'interpretazione di alcune di queste distanze si trovano in [9]. In tale lavoro i fonemi sono considerati punti di uno spazio a n dimensioni, dove gli assi coordinati sono i tratti acustici, oppure, in un caso, i tratti articolatori. Le distanze che qui prenderemo in considerazione sono le seguenti.

3.4.1 Dette x_i e y_i le i-esime coordinate dei fonemi x e y nello spazio dei tratti acustici, risulta una distanza la seguente funzione

$$d_1(x, y) = \sum \alpha_i \quad \text{con } \alpha_i \begin{cases} 1 & \text{se } x_i \neq y_i \\ 0 & \text{se } x_i = y_i \, . \end{cases} \tag{3.3}$$

Tale distanza evidenzia soltanto differenze tra i fonemi sulle singole coordinate, considerando tutti i tratti acustici egualmente importanti.

3.4.2 Essendo x_i e y_i come in 3.4.1, e detta $f(a)$ la frequenza di occorrenza del fonema a, risulta una distanza la seguente funzione

$$d_2(x, y) = \sum_i |x_i f(x) - y_i f(y)| \, . \tag{3.4}$$

Calcoli delle frequenze di occorrenza dei vari fonemi nell'italiano parlato si trovano, ad es., in [10] e in [11]; qui sono state usate, seguendo [9], quelle riportate in [10].

3.4.3 Un'altra distanza, discussa in [9], prende lo spunto da una distanza proposta in [12], basata su tratti articolatori; vengono operate alcune semplificazioni e scelte di parametri per adeguarla ai fonemi della lingua italiana. Viene fissata a priori una gerarchia numerata da 1 a 4 tra quattro classi disgiunte di parametri articolatori, e una ulteriore sottogerarchia tra i parametri all'interno di una stessa classe. Il numero di parametri varia da classe a classe, da un minimo di due a un massimo di sei, con numerazione che parte dallo zero. Sia la gerarchia tra le classi di parametri sia le sottogerarchie all'interno di ciascuna classe hanno motivazioni linguistiche.

Il fonema è a sua volta individuato da un numero variabile di parametri, appartenenti a classi diverse o anche alla medesima classe. Indicata con Δ la differenza simmetrica tra insiemi, sia i l'indice della classe dei parametri articolatori che contribuiscono a individuare il fonema; sia $S_i(x)$ l'insieme dei parametri della i-esima classe che individuano x (con le scelte di [9], $S_i(\cdot)$ è vuoto oppure costituito da un solo elemento); sia $v[S_i(x)\Delta S_i(y)]$ la somma dei valori assoluti delle differenze tra gli ordini gerarchici dei parametri all'interno della classe i-esima; se uno degli insiemi $S_i(x)$, $S_j(y)$ è vuoto, si pone $v[S_i(x)\Delta S_i(y)] = 0$. Posto ancora

$$\varphi_i(x, y) = \begin{cases} 0 & \text{se } S_i(x) = S_i(y) \\ 1 & \text{se } S_i(x) \neq S_i(y) \end{cases} \tag{3.5}$$

la seguente funzione risulta una distanza su ogni asse coordinato

$$\beta_i(x, y) = i\varphi_i(x, y) + \frac{\operatorname{card}\left(S_i(x)\Delta S_i(y)\right)}{10} + \frac{v[S_i(x)\Delta S_i(y)]}{10^2}. \tag{3.6}$$

Il primo termine indica se i fonemi sono individuati, nella classe i-esima, dallo stesso insieme di parametri, nel qual caso tale primo termine è nullo; se invece i due insiemi di parametri nella classe i-esima sono diversi, il primo termine indica l'importanza gerarchica della classe in cui si verifica tale diversità. Il secondo termine indica la differenza del numero di elementi tra l'insieme di parametri individuante x nella i-esima classe e quello individuante y nella stessa classe. Nelle scelte operate in [9] per l'adeguamento dei parametri ai fonemi della lingua italiana, se è $\varphi_i = 0$ è anche $\operatorname{card}\left(S_i(x)\Delta S_i(y)\right) = 0$, mentre se è $\varphi_i(x, y) = 1$, $\operatorname{card}\left(S_i(x)\Delta S_i(y)\right)$ può valere 1 o 2. Il terzo addendo indica l'importanza gerarchica all'interno della classe i-esima degli elementi sui quali differiscono i due fonemi. La distanza complessiva risulta definita così:

$$d_3(x, y) = \sum_{i=1}^{n} \beta_i(x, y) \tag{3.6}$$

dove la sommatoria è effettuata sulle n classi di parametri articolatori (per l'italiano la classificazione dei parametri di [9] ha $n = 4$).

3.4.4 Un'altra distanza $d_4(x, y)$ è quella ottenuta tramite l'MDS (Multi-Dimensional Scaling), introdotto da Kruskal (vd. [13], [14], [15]).

Questa distanza è ottenuta simmetrizzando le matrici di confusione (ad esempio ponendo $a_{ij} = a_{ji} = \overline{a_{ij}}$), e quindi costruendo una configurazione di fonemi in uno spazio a dimensione R opportunamente predeterminata le cui reciproche distanze siano simili agli elementi della matrice simmetrizzata.

Esistono vari studi su come si possono costruire tali configurazioni, e su quale sia il numero ottimale di dimensioni dello spazio. Il procedimento di Kruskal consiste nel minimizzare lo sforzo tra la matrice delle distanze da costruire $[d_4(x, y)]_{ij}$ e la matrice simmetrizzata $\overline{[p(x, y)]_{ij}}$; si tratta pertanto di minimizzare

$$\frac{\left\{\sum_i \sum_j \{\overline{[p(x, y)]_{ij}} - [d_4(x, y)]_{ij}\}^2\right\}^{\frac{1}{2}}}{\sum_i \sum_j \{[d_4(x, y)]_{ij}\}^2}. \tag{3.7}$$

Vari algoritmi di MDS si trovano in [8].

4. Sviluppo del modello additivo.

4.1 Viene ora proposta una ulteriore ipotesi per una successiva approssimazione del modello additivo definito in 2.1. Supponiamo che le funzioni $a(x)$ e $b(y)$ che compaiono in (2.1) siano fortemente correlate con le rispettive frequenze di x e di y. La correlazione più semplice è la proporzionalità (diretta o inversa). È plausibile che la probabilità di fraintendimento dello stimolo x sia inversamente proporzionale alla sua frequenza $f(x)$, mentre d'altra parte è plausibile che la probabilità di credere di aver sentito y sia proporzionale alla frequenza $f(y)$ di y. Pertanto viene proposto il seguente modello

$$p(x, y) = \frac{k_1}{k_1 + d(x, y)} + \frac{k_2}{f(x)} + k_3 f(y) \tag{4.1}$$

dove $d(x, y)$ è una delle distanze considerate nel § 3. Si calcola dapprima quale distanza minimizza lo sforzo avendo posto

$$s(x, y) = \frac{k_1}{k_1 + d(x, y)} \tag{4.2}$$

ottenendo quindi anche la migliore parte simmetrica del modello; una semplificazione che non comporta modifiche teoriche essenziali si ha ponendo $k_1 = 1$. Il calcolo di k_2 e di k_3 in (4.1) si ottiene successivamente con una regressione.

4.2 Il computo è stato effettuato su 18 matrici di confusione che costituivano tre gruppi da 6 matrici ciascuno. Il primo gruppo era costituito da matrici di confusione costruite ciascuna come media delle risposte di 24 ascoltatori in situazioni di ascolto disturbato da rumore (vd. [16]): le sei condizioni di disturbo erano con i seguenti rapporti segnale/rumore (S/R) in decibel: 20; 15; 5; 0; -5.

È stata quindi effettuata la media aritmetica di queste 6 matrici medie, costituendo così un'unica matrice $p(x, y)$. Ponendo $k_1 = 1$ sono poi state calcolate le singole matrici simmetriche $s(x, y)$ di (4.2), una per ciascuna delle distanze $d_i(x, y)$ presentate in 3.4. Per la distanza d_4, la motivazione che rende giustificato $R = 5$ come numero di dimensioni della configurazione è dettata dal fatto che quel numero di dimensioni permette di dare un'interpretazione acustica dei cinque assi coordinati; vd. [17].

Il computo ha fornito per gli sforzi i risultati riportati in Tab. I. Lo sforzo risulta pertanto minimo prendendo come distanza la d_2, ma ha un valore vicinissimo al minimo anche considerando d_3. Il risultato è di notevole importanza, perchè indica l'equivalenza di fatto, ai fini del modello, tra una valutazione di distanza basata su tratti acustici, cioè su caratteristiche della percezione del fonema, e una valutazione basata su tratti articolatori, cioè su caratteristiche di produzione del fonema stesso. Si noti che invece le distanze teoriche calcolate nei due modi erano piuttosto diverse (vd. [9], tabelle 6 e 5).

Si può invece notare che la distanza d_4 definita in 3.4.4 dà uno sforzo ben più marcato. Ciò segue dal fatto che la distanza ottenuta tramite l'MDS opera una simmetrizzazione a priori, che risulta quindi arbitraria. Per il trattamento ulteriore del modello è stata assunta la d_2.

4.3 Sono stati quindi calcolati i coefficienti k_2 e k_3 della (4.1) tramite una regressione lineare; è risultato

$$k_2 = -4{,}761\ \text{E-03} \qquad k_3 = -1{,}202\ .$$

4.4 È stata effettuata una indagine più approfondita nelle varie condizioni di rumore citate in 4.2, per lo studio della variazione dello sforzo e dei coefficienti k_1 e k_2 al variare del rumore. I risultati sono nella Tab. II.

Tab. I *Variazione dello sforzo in funzione delle varie distanze.*

Matrice R/N (*media*)	d_1 $\sum \lvert x_i f(x) - y_i f(y) \rvert$	d_2 $\sum \alpha_i;\ \alpha_i = \begin{cases} 1 & x_i \neq y_i \\ 0 & x_i = y_i \end{cases}$	d_3 tratti articolatori	d_4 *MDS* $R = 5$
k_1	1	1	1	1
Sforzo	0,937	0,678	0,679	0,885

Tab. II *Variazione dello sforzo con la distanza d_2 in funzione delle condizioni di mascheramento.*

Matrice	*S/R* (*media*)	*S/R* 20 dB	*S/R* 15 dB	*S/R* 10 dB	*S/R* 5 dB	*S/R* 0 dB	*S/R* – 5 dB
Sforzo	0,678	0,638	0,660	0,694	0,706	0,762	0,751
k_2	– 4,761 E-03	– 5,245 E-03	– 5,216 E-03	– 4,881 E-03	– 4,815 E-03	– 4,398 E-03	– 4,011 E-03
k_3	– 1,202	– 0,989	– 1,001	– 1,150	– 1,178	– 1,362	– 1,533

Tab. III *Filtraggio PA: i numeri in testa alle colonne indicano le frequenze al di sotto delle quali il suono è tagliato.*

Matrice	*Matrice mediata*	0,47 kHz	1 kHz	1,5 kHz	2,2 kHz	2,7 kHz	3,3 kHz
Sforzo	0,623	0,626	0,626	0,627	0,624	0,623	0,625
k_2	– 5,446 E-03	– 5,476 E-03	– 5,482 E-03	– 5,480 E-03	– 5,421 E-03	– 5,421 E-03	– 5,427 E-03
k_3	– 0,901	– 0,888	– 0,885	– 0,886	– 0,912	– 0,925	– 0,909

Tab. IV *Filtraggio PB: i numeri in testa alle colonne indicano le frequenze al di sopra delle quali il suono è tagliato.*

Matrice	*Matrice mediata*	0,39 kHz	0,47 kHz	0,56 kHz	1 kHz	1,5 kHz	2,2 kHz
Sforzo	0,646	0,626	0,626	0,678	0,693	0,712	0,696
k_2	– 4,980 E-03	– 5,479 E-03	– 5,469 E-03	– 5,018 E-03	– 4,563 E-03	– 4,703 E-03	
k_3	– 1,106	– 0,886	– 0,891	– 1,089	– 1,251	– 1,289	– 1,228

Si noti come, eccezion fatta per la situazione in ultima colonna, dove il mascheramento rumoroso supera in intensità il segnale, lo sforzo cresce; anche k_2 cresce, ma rimane sempre molto vicino allo zero; invece k_3 decresce. Il rumore pertanto rende sempre più asimmetriche le matrici, oltre che, come era facilmente intuibile, aumentare la dispersione dei dati attorno alla diagonale principale. Ciò significa che, anche tra i normoacusici, l'asimmetria ha un significato ed un peso preciso, e quindi non può essere eliminata con una semplice media aritmetica.

4.5 Gli studi su segnali mascherati da rumore hanno generalmente interesse per i normoacusici. Altro tipo di interesse nasce in vista delle terapie per coloro che hanno difetti di udito, come ad esempio coloro che non odono suoni al di sopra o al di sotto di certe frequenze. Anche per questo tipo di difetti sono state studiate le matrici di confusione. In particolare sono state prese in considerazione sei situazioni di filtraggio passa-alto (PA) e sei situazioni di filtraggio passa-basso (PB) (vd. [16]). Una applicazione a queste situazioni del modello proposto in 4.1 con $k_1 = 1$ ha portato ai risultati raccolti nella Tab. III per il caso PA e nella Tab. IV per il caso PB.

Dall'esame della Tab. III si nota come il modello si adatti senza praticamente risentire del taglio delle basse frequenze: lo sforzo è sempre praticamente costante e assai minore che nelle situazioni di mascheramento con rumore; k_2 continua ad essere insignificante, k_3 ha qualche oscillazione un po' più marcata.

Il taglio delle alte frequenze ha una maggiore incidenza sul modello che non il taglio delle basse; k_2 è sempre praticamente nullo.

Alcuni interessanti commenti sulla maggiore o minore intelligibilità per le consonanti dell'italiano in relazione alle loro caratteristiche acustiche e basate sulle matrici di confusione si trovano in [18].

5. Conclusioni.

La dipendenza di $a(x)$ e $b(y)$ dalle rispettive frequenze secondo il modello qui proposto è soddisfacente soltanto per quanto riguarda $b(y)$. La funzione $a(x)$ è sempre poco significativa; viceversa $b(y)$ ha una buona correlazione con la frequenza di y, e tale correlazione cresce al crescere del rumore. Nel caso di situazioni con filtraggio, il modello si adatta meglio al filtraggio passa-alto.

BIBLIOGRAFIA

[1] E. W. HOLMAN, *Monotonic Models for Asymmetric Proximities*, Journ. Math. Psych., **20** (1979), pp. 1-15.

[2] L. GOLDSTEIN, *Bias and Asymmetry in Speech Perception*, in V. FROMKIN (ed.), *Errors in linguistic performance*, Academic Press, N. Y. (1980), pp. 241-261.

[3] M. D. WANG - R. C. BILGER, *Consonant Confusion in Noise: a Study of Perceptual Features*, JASA, **54**, pp. 1248-1266.

[4] E. C. CARTERETTE - M. H. JONES, *Informal Speech*, Univ. Calif. Press, Los Angeles (1974).

[5] A. M. ROBERTS, *A Statistical Linguistic Analysis of American English*, Mouton, The Hague (1965).

[6] M. RUHLEN, *A Guide to the Languages of the World*, Stanford, Palo Alto (1975).

[7] U. BORTOLINI - C. MINNAJA - G. PINI - G. ZILLI, *Evaluation of Asymmetry in Similarity Matrices*, Proc. IV F.A.S.E. Symposium on Acoustics and Speech, ESA, Roma (1981).

[8] G. PINI, - G. ZILLI *Algoritmi di scaling multidimensionale con programmi orientati a mini-computer*, IAC-CNA, Monografie di Software Matematico, n. 14 (1983).

[9] U. BORTOLINI - C. MINNAJA - L. PACCAGNELLA, *Aspetti teorici e metodologici della nozione di distanza*, in AA. VV., *Linguaggi e formalizzazioni*, Bulzoni, Roma (1979), pp. 151-171.

[10] R. BUSA - C. CROATTO-MARTINOLLI - L. CROATTO - C. TAGLIAVINI - A. ZAMPOLLI, *Una ricerca statistica sulla composizione fonologica della lingua italiana parlata eseguita con un sistema IBM a schede perforate*, Proc. XII Int. Speech and Voice Therapy Conference, Padova (1973), pp. 542-562.

[11] U. BORTOLINI - F. DEGAN - C. MINNAJA - L. G. PACCAGNELLA - G. ZILLI, *Statistics for a Stochastic Model of Spoken Italian*, Proc. XII Int. Congr. Ling., Wien (1977), pp. 580-586.

[12] G. E. PETERSON - F. HARARY, *Foundations of Phonemic Theory*, Proc. Symp. Appl. Math., **12** (1961), pp. 139-165.

[13] J. B. KRUSKAL, *Multidimensional Scaling by Optimizing Goodness of Fit to a Nonmetric Hypothesis*, Psychometrika, **29** (2) (1964), pp. 1-27.

[14] J. B. KRUSKAL, *Multidimensional Scaling and Other Methods for Discovering Structure*, in K. ENSLEIN - A. RALSTON - H. S. WILF, *Statistical Methods for Digital Computers*, Vol. III, Wiley, New York (1977).

[15] J. B. KRUSKAL, *Nonmetric Multidimensional Scaling: A Numerical Method*, Psychometrika, **29** (2) (1964), pp. 115-129.

[16] E. Magno-Caldognetto - F. E. Ferreo - K. Vagges, *Un test di confusione tra le consonanti dell'italiano: primi risultati*, C.N.R., Quad. Centro Studio Ric. Fonetica, I (1982), pp. 238-296.
[16] U. Bortolini - G. Pini - G. Zilli - F. E. Ferrero, *Dimension of Perception for Italian Consonants: Multidimensional Analysis*, C.N.R., Quad. Centro Studio Ric. Fonetica, I (1982), pp. 77-90.
[18] E. Magno-Caldognetto - F. E. Ferrero - K. Vagges - D. Domenghini, *Intelligibilità delle consonanti dell'italiano in condizioni di mascheramento (S/R), di filtraggio passa-alto (PA) e passa-basso (PB)*, Il Valsalva, Boll. It. Audiol. Foniatr., IV (1982), pp. 163-172.

Manoscritto pervenuto in redazione il 3 novembre 1983.