



LO SPAZIO DELLE VOCALI

Bas o a de ome, S a a a a^{*}
Sc o a M a r S e o r e, f s a, l n r e s a d e S d d S r e n a s e d e d A r z z o
b. c a d e o m e s n s. i; c a s a g t s c a . i

(in corso di stampa negli Atti del III Convegno Nazionale AISV "Scienze Vocali e del Linguaggio"
Metodologie di Valutazione e Risorse Linguistiche, Trento 29.IX-I.XII.2007)

*Always end the name of your child with a vowel,
so that when you yell the name will carry*
(B osby)

1. SOMMARIO

l'esperienza a o d'ia co r e f n a à o s d o d e r e a a b d e s c . . . r e a o r e n e
c o n o r e n t p a a s b r e (r e a o t a c a) d i f f e r e n z a z o m e, n c a r e t t e f o m e t t e, d
s e r e n t o c a c . f n a c o a r e, a n o s a n d a m e r e f o c a z z a a s ' a n a s d e a
a n a t d e s c . . . a m e c c e s s a a a a r e s e n t a r e a s . . . a r e a d s o z o m e d e r e r e n t
o c a c i ' n r e n o d p s s e a r e o t s a n o t t e r e s e n a a c p a s r e t t z z a n t
(c o r e ' a b b a s s a r e n o d e t o c a o n c i r e r e d o b a s s e r e t a r e a z z a z o m e d / a /).
r e t t e r e o d s c o t t e c o r e s a o s s ' b r e r e d e n z a r e a n a o r e r e s a à a ' n r e n o d
p s e t d o c a a d o a n d o p a o s t e a n o n - s u p e r v i s i o n a t a, n c c o d e n o n t s a
r e s e n z a d p s r e t t s o r e r e s e n o t t e r e n o s s e a s r e c o r e r e o d a à d e a
s r e t t a z o m e d e d a o n c a r e o r e: a c o n a o, t o c e s s o d o a n z z a z o m e d e r e
o c a, s a n t e a f i n e a r e n d n o, a t e m e r e a r e n t s a b a s e d r e o a t e
c o r e n z e s a s t t e d e c o r p u s d i n p u t .

r e r e s t t a z o n e s a o r e a o p o d e o m e a r e a d a r e n d r e n o c o r e o
r e n o n s r e s o n a o, à c o s d d e a S M (S e l f - O r g a n i z i n g M a p s) . t a o c o r p o r a
o c a c, s d d s t e o o a a c c e n a r e (r e s e n z a / a s s e n z a d a c c e n o) r e s t t e (r e o v s.
s o n a n o) s o n o s a s c e c o r e t r a i n i n g s e t .

l' n o s o n t e n a o t o c a c o r e d e n o d a d i f f e r e n t o o r e d a a b : a o
s r e a, d i f f e r e n z e s r e a, f r e r e n z a f o n d a r e n a r e, t d t a a. A f i n e d r e f c a r e
r e c o n o a d e s c . . . a m e c c e s s a a a a n t e n a c o r e n t e r e c a t e o a r e n e b e n d e f n a
o a n z z a z o m e (c l u s t e r i n g) s a z a r e d e r e t o c a, d e t t o a o t a c a e n o n - s u p e r v i s i o n a t a,
d i f f e r e n t a d d e s a r e n s o n o s a c o n d o a n d o t t e r e o r e t a o o a d e r e a a b
r e s e n e s a r e r e a t r e s e n a z o m e d e t t a o a c s c o. A r e o c e d a s r e r e n a r e r e s a
a s c o r e a d e s c z o m e n t a r e d e d a o, o r e o t a r e a n t e a a b d e s c t t e s o n o
s c r e n t a a a n t e p o a n z z a z o m e d e d a t o c a c n c l u s t e r s d s n t .

2. CATEGORIE, GRANULARITÀ DESCRITTIVA, AUTO-ORGANIZZAZIONE

o n r e m e ' a o o a n z z a z o m e ' s n e n d e a c a a c à r e s b a d a s s e a d
a p r e r e m e r e o, p o s a o d r e b o t o, n r e m e a r e, p a s t t a z o m e

* l' a o o r e f o d e a c o a b o a z o m e d e d r e a o . A f n c o n c o s a, s o n o d a
a b s a B a a a f 2, 3, 5, 2; a S a a a f 4, 5. l' a a a f o f n a r e r e o r e a d
r e n t a b.
b o m e n (2 0 0).

La seconda parte del capitolo, intitolata "SOM (Self-Organizing Map)", descrive il funzionamento di questo algoritmo di clustering. Si parte dalla definizione di SOM come una rete neurale artificiale che mappa dati di alta dimensionalità su un numero ridotto di neuroni (tipicamente 10-100). Il processo di apprendimento avviene attraverso la competizione tra i neuroni per attrarre i dati di input, con il risultato di formare una mappa topologica dei dati. La SOM è particolarmente utile per visualizzare dati multidimensionali in uno spazio bidimensionale, consentendo di osservare la struttura intrinseca dei dati e di identificare gruppi di punti correlati. Il testo discute anche le applicazioni della SOM in vari campi, come l'analisi di dati di mercato, la classificazione di documenti e la scoperta di nuove conoscenze in campi come la biologia e la medicina. Infine, vengono presentati alcuni esempi di implementazione della SOM e i risultati ottenuti.

3. LA SOM (SELF-ORGANIZING MAP)

3.1 Caratteristiche e funzionamento

Una SOM (Self-Organizing Map) è una rete neurale artificiale che mappa dati di alta dimensionalità su un numero ridotto di neuroni (tipicamente 10-100). Il processo di apprendimento avviene attraverso la competizione tra i neuroni per attrarre i dati di input, con il risultato di formare una mappa topologica dei dati. La SOM è particolarmente utile per visualizzare dati multidimensionali in uno spazio bidimensionale, consentendo di osservare la struttura intrinseca dei dati e di identificare gruppi di punti correlati. Il testo discute anche le applicazioni della SOM in vari campi, come l'analisi di dati di mercato, la classificazione di documenti e la scoperta di nuove conoscenze in campi come la biologia e la medicina. Infine, vengono presentati alcuni esempi di implementazione della SOM e i risultati ottenuti.

me omne de a re co s ondebbe n cluster re co è n a a re no o o meo de da d input, oca zza n re a zona re n a re c re o d s a à nd d a o da a re².

Il processo di adde s a re no (training process) con ce nd con pa fase d ada a re no s a re a t input d na a re o are. La re ada a re no re co o da a S M a re od f ca de o con f c re n d t conness o me.

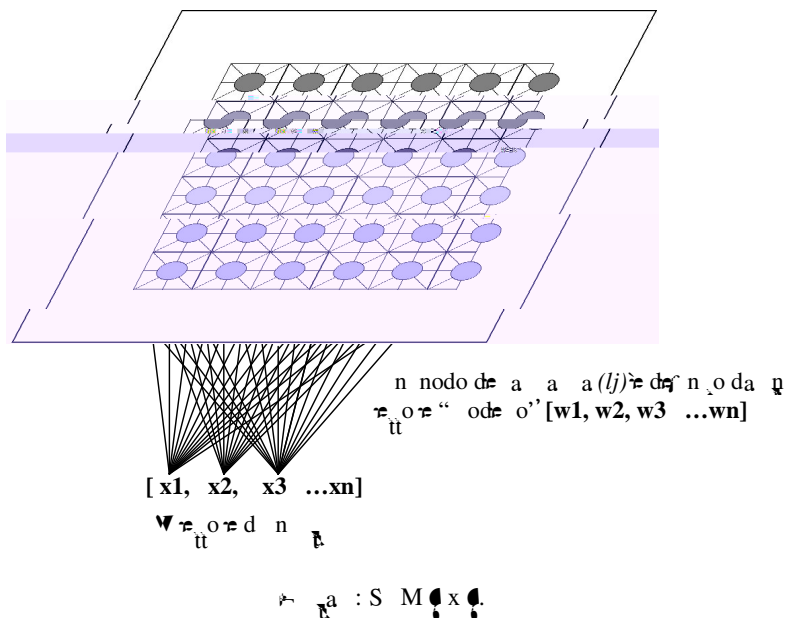
o re t a re a o, o ccesso d a re nd re no re non s re s ona o, n a re a o re non re de n z o me n an c o s a con f a z o me s a re c re re à a S M. La a a z o me de con f c re n d t conness o me, re nd a con f a z o me de a a a, d re de à so a o da re ca a re s re s a s re de ca o me d da re co s re a base d conness o n z a s c re s n o re a o t t o ccesso d a re nd re no.

n a o a s re se n a n pattern d n re sso, re ne n a re s e t o n matching a re pattern (de n o co re t re o re a n d re n s o n) re re t t n a me o n a f o an t a a a b d re n s o n a re (de n re a n c re sse n re n re t o a a n d re n s o n). I c re o ado a o re s a re a d re s t a a d re t o f a ca t o a a no z o me d 'd s a n z a'. La 'd s a n z a' n t re s o me o re sse re ca co a a n a od (so a re re s co n s de a a d s a n z a re c de a). o an re no a re re t re s o ca co o re re s o s s b re da f a o re d re re o, re o re a re se n a t pattern n n re sso re re o re a re se n a t, n a de a t a a, con re n o n o re de s t o n re o d co o me n.

S con s de a t o a co re n a d re s c a ' n c ce t re a n c a 'd s a n z a' re ce de n re re ca co a a as s re n o a o re. L'a o o d re o me n ad o a a s a re a de a t b o a: non re n a so a n a ' n c ce', a s con s de a n o ' n c t c' (o re o s re c re a n o) a n c re re re a re n a d re s c a re s o a n o re n o pa de o ' n o n o' da ' n a ' n c ce, i n c a re. La re n o n o, re a re t n o o a d re o me n, re de t o r a g g i o d e l l a b o l l a.

o o a re nd d a o re n a ' n c c', s od f ca n o so o con f c re n d conness o me de re me re f a n n o c a o a re s re n a, n o d o da d n re a 'd s a n z a', a o re se n a re re se n a o n n re sso. S a s s a o a re o re d input s c c e s s o. M a n t a n o re o ccesso d a re nd re no a a a n, ' a t t o' de a b o a re a s s o d od f ca de con f c re n t re n o n o f a t t o re s s a re n re de c re s c e re, con n a re re d e n a da o s re re n a o re.

² Non re sc sa re o a o s s b a, n re o f re re n re, re s a n a s s e b a o d me on a de n re n n c o cluster. o re do o a a so f o me de a a a, re re n re a z o me con a s a d re n s o me: a re n a n d o, n a, n o d re nd a n c re a d re n s o me), re s, n d d re a n n o n a ca re o a t s re t a t t a n a re.



Quando l'output del nodo i è maggiore di quello del nodo j , il nodo i è detto "nodo vincitore" [w1, w2, w3 ... wn].

3.2 L'algoritmo di addestramento

La set di vettori d'input è descritto da $x(t) \in R^n$ dove t indica il passo di tempo. Il nodo i della SOM con il vettore $w_i(t)$ è detto "nodo vincitore" se $w_i(t) \cdot x(t) > w_j(t) \cdot x(t)$.

L'algoritmo della SOM si basa sul concetto di "cluster", creando gruppi di nodi (clusters), in base alla loro posizione spaziale.

1. Il vettore d'input $x(t)$ è confrontato con i vettori $w_i(t)$ a fine di trovare il "Best-Matching Unit" (BMU), il nodo che ha il prodotto scalare con $x(t)$ maggiore di tutti gli altri.
2. Il vettore $w_i(t)$ del nodo "vincitore" viene aggiornato con un certo peso α (dove $0 < \alpha < 1$) e il vettore $w_j(t)$ del nodo "vicino" viene aggiornato con un certo peso β (dove $0 < \beta < \alpha < 1$).

a tendimento) e' p'ncipalmente se ne ricerca la minima differenza tra il modello adattato e l'osservazione $x(t)$, e' opportuno anzichè d'indicare il processo di apprendimento, si può dire che il processo di apprendimento è descritto da una successione di vettori $m_i(t)$ che si avvicina a $x(t)$.
 Se il valore di α è basso, il processo di apprendimento è lento e se è alto è veloce.
 La differenza tra $x(t)$ e $m_i(t)$ è data dalla differenza di adattamento c che si ottiene dalla seguente equazione:

$$c = m_c(t) = \min \|x(t) - m_i(t)\| \quad (1)$$

L'adattamento è un valore che si avvicina a $x(t)$ con un certo passo α che dipende dal valore di c . L'adattamento è un valore che si avvicina a $x(t)$ con un certo passo α che dipende dal valore di c .

$$m_i(t+1) = m_i(t) + \alpha(t)[x(t) - m_i(t)] \quad \text{per } i \in N_c(t) \quad (2)$$

do t è l'indice di apprendimento e $\alpha(t) \in [0, 1]$ è il passo di apprendimento, che è un valore che si avvicina a $x(t)$ con un certo passo α . La differenza tra $x(t)$ e $m_i(t)$ è data dalla differenza di adattamento c che si ottiene dalla seguente equazione:
 L'adattamento è un valore che si avvicina a $x(t)$ con un certo passo α che dipende dal valore di c .
 La differenza tra $x(t)$ e $m_i(t)$ è data dalla differenza di adattamento c che si ottiene dalla seguente equazione:
 L'adattamento è un valore che si avvicina a $x(t)$ con un certo passo α che dipende dal valore di c .

$$r(t) = A - B \cdot t$$

nei casi in cui $A=4$ e $B=0.002$, si ha:

$$r(t) = A / (1 + B \cdot t)$$

nei casi in cui $A=4$ e $B=0.003$.

Le fasi di apprendimento sono descritte dalla seguente equazione:
 La differenza tra $x(t)$ e $m_i(t)$ è data dalla differenza di adattamento c che si ottiene dalla seguente equazione:
 L'adattamento è un valore che si avvicina a $x(t)$ con un certo passo α che dipende dal valore di c .

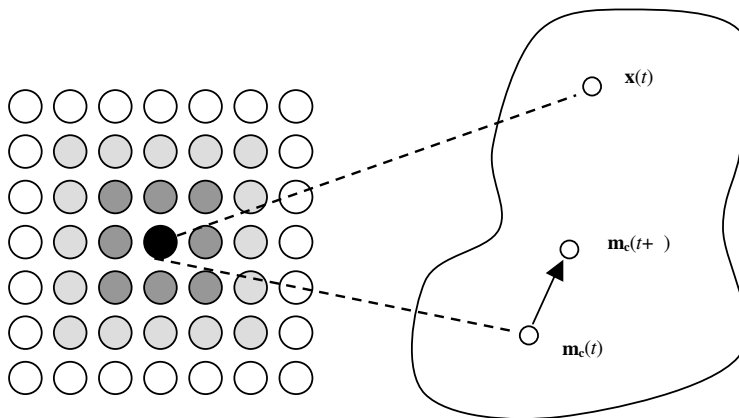


Fig. 2: Fase di adattamento a un dato $x(t)$, e di aggiornamento a $m_c(t)$.

La Fig. 2 mostra una fase di adattamento di una Self-Organizing Map durante il processo di adattamento. Come si vede, il nodo $m_c(t)$ si avvicina al dato $x(t)$ e si allontana dal nodo $m_c(t+)$. Il nodo $m_c(t)$ si avvicina al dato $x(t)$ e si allontana dal nodo $m_c(t+)$ solo se il dato $x(t)$ è diverso da $m_c(t)$. Durante il processo di adattamento, i nodi si muovono in base al dato $x(t)$ e si allineano con esso. Ma, se il dato $x(t)$ è uguale a $m_c(t)$, il nodo $m_c(t)$ non si muove.

3.3 Aspetti multipli di una SOM

La SOM può essere vista come un modello di apprendimento non-supervisionato. Come si vede bene a proposito delle sue caratteristiche:

- **Clustering.** Non è un algoritmo di clustering, ma si comporta come tale, dove i dati vengono raggruppati in cluster (o categorie) e così si può dire che si sta facendo un clustering.
- **Quantizzazione vettoriale.** La SOM funziona come un dispositivo di quantizzazione vettoriale. I nodi della SOM sono vettori di dati x , e si avvicinano a se stessi in base al dato x . La SOM funziona come un dispositivo di quantizzazione vettoriale.
- **Estrazioni delle caratteristiche.** La SOM estrae le caratteristiche dai dati, e si può dire che si sta facendo un'analisi delle caratteristiche.

La SOM non è solo un dispositivo per analisi statistiche. Una delle sue applicazioni è la riduzione della dimensionalità dei dati, cioè la riduzione della dimensionalità dello spazio dei dati n -dimensionale in uno spazio m -dimensionale, dove $m < n$.

in address, a tenore sa o condono zando pa S M d d ensone 20x 0
 n z a z z a a con res random zando e a ce ca de BM (best-matching unit) a
 d s anza t c dea.

Le f 3 e 4 o ano re o an z z a z on re a o s s e oca c (on co a a o
 re o, on co a a o se s on aneo, a ono a a o re o, a ono a a o se s on aneo).

Ad address, a tenore conciso, o h me om de a t a a d ten a a ten a o d
 o n na de re t a oca re ten me corpus (o a ten re p me one o a ten a re
 a con re o d se oca a a ten ane a ca re o t e d i re n . Le f 3 e 4
 os, aho, re a a d s a z z a z on e, so o re oca re con t a o re f re ten za
 'cadono' a 'n re no de nod /me on de a S M). Ecco pa a ana s de s a :

oca s da p pod s a a co a o o cc ano s az con t me re
 S Ms: n a re a o e, a d s os z on e s az a re, re a da 'a o n basso,
 ca ca continuum a co a o o an re o re cen a re o s re o re;

s re a co a s t o t an z z a n o re o t d s t o a co a : a
 s o a os z on e re o s az o ac s co d oca o a t co a re o t a s con o
 n na a re sen az on e re no da s a n o s az a re (a a o
 se s on aneo s o an z z a re o s re o a a a o re o, re a 'n re no
 de a a o se s on aneo oca s t o a ono s t o t an z z a re o de
 s s re a on co);

a t conosc b a d / /, a ca d a o ca re ecce ten za de s s re a t on co
 sano, o a t f resso me s o s o a re n o s az a re;

a os t o z z a z on e d /a/ a t a ca a re s ca d a o ca de s s re a on co
 sano re d os a a da a t a o re t c hanza s az a re a a c asse de re
 oca os re o .

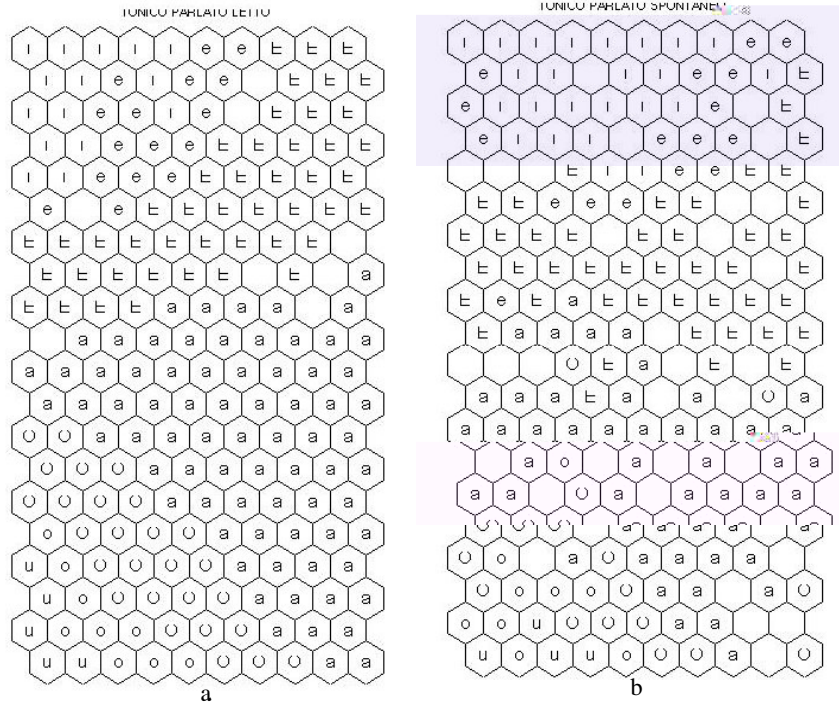


Fig. 3: an zzaz one de s se a ~~M~~ PARLA ~~L~~ (a) de S/S ~~M~~ MA ~~M~~ PARLA ~~S~~ M/S ~~M~~ A ~~M~~ (b). ~~de~~ 'addes_t a ~~re~~n_o ~~re~~s_aa ~~re~~ a_a pa S M 20X 0.

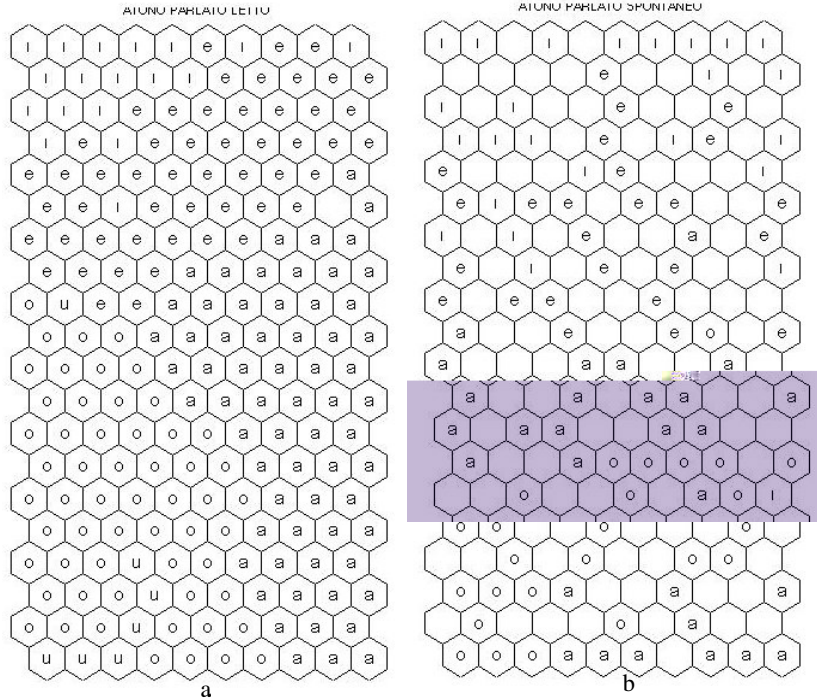
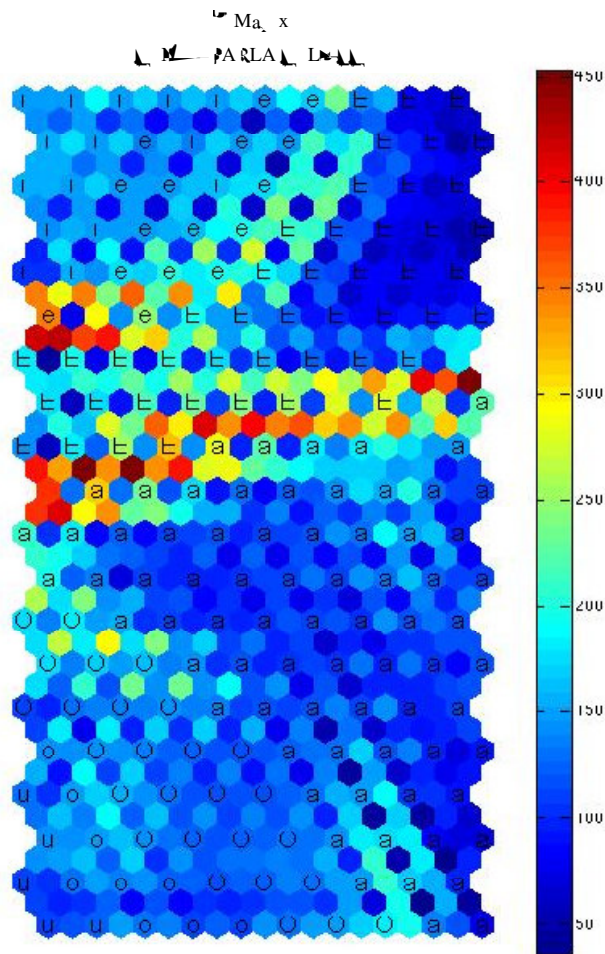


Fig. 4. (a) *an zzaz one de s se a A M PARLA L D* (a) *re de s se a A M PARLA L S M S P M A M* (b). *Pe t 'adde s t a re n o e s a a re a a pa S t M 20 X 0.*

5.2. *Economia descrittiva (less is more)*

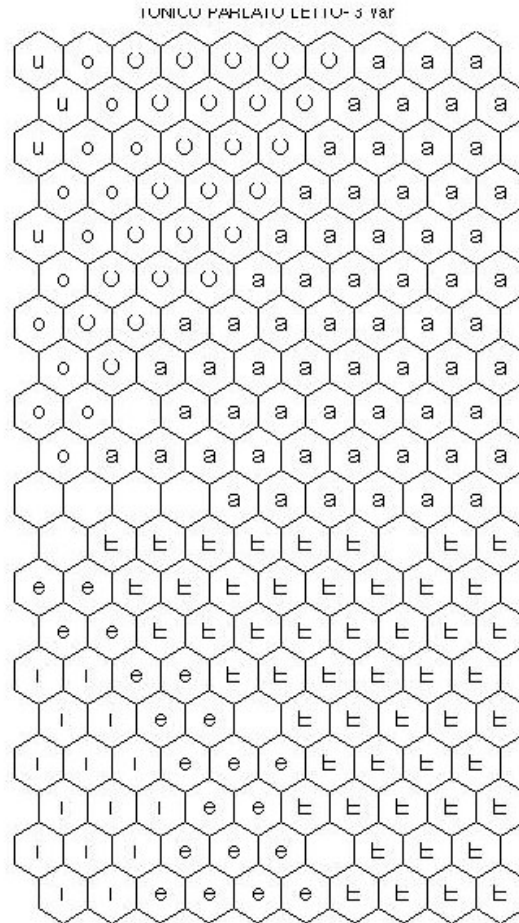
Le o an zzaz on f na , res b re da a S M re a o corpora oca c, anno re resso 'nd d az one d o o t e e e, de re o re, ben d s n e ca re o re oca t e re (co re s re a re na s o, a bon a de re o an zzaz on t a a n acco do a t a d f re nza d s , re c s , re a t co a t uz onano re o' d s o a co a). na s ccess a aha s cons de re a a t a ab , sa re me a a resen az one de t oca , a a ono resse re re men re '(a o) o an zzaz one de da o resso me re a re c ass oca t e re. Pe t o t d s az o, c a o a o a t t s a so o de A M PARLA L D , res t endo re conc s on de a hos t a nda me an t e a t e s an t corpora.



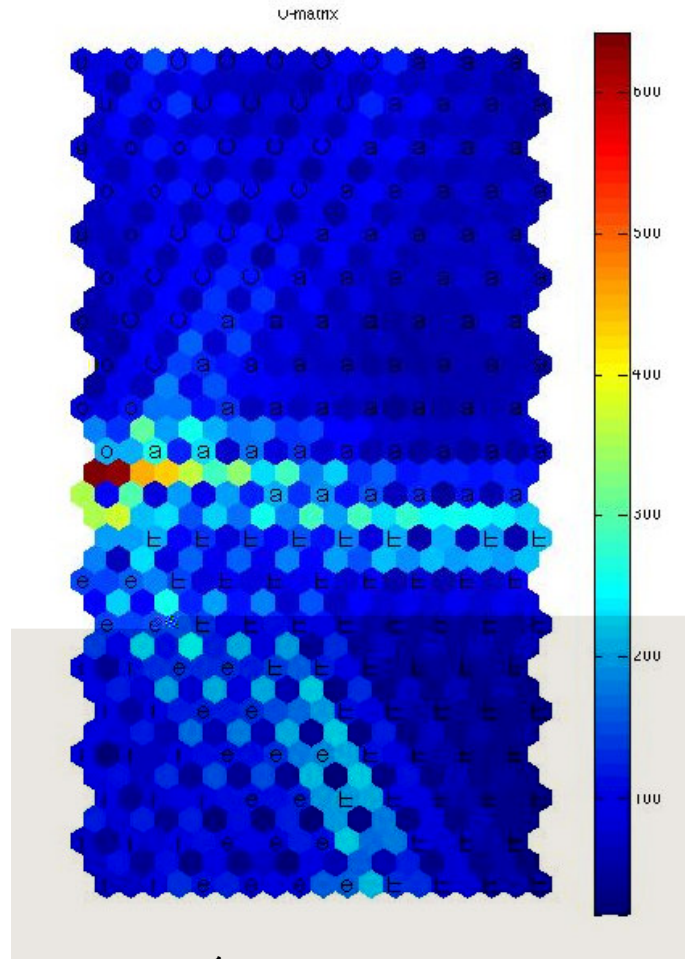
La visualizzazione Max di un corpus di parole, mostrando la distribuzione delle lettere. La scala a destra indica il numero di occorrenze di ciascuna lettera, con valori da 50 a 450. Le lettere 'e', 'a', 'o' e 'u' sono le più frequenti.

La visualizzazione Max di un corpus di parole, mostrando la distribuzione delle lettere. La scala a destra indica il numero di occorrenze di ciascuna lettera, con valori da 50 a 450. Le lettere 'e', 'a', 'o' e 'u' sono le più frequenti.

¹ Sci & Sironi (2000).



τ τ a τ an zzaz one de s τ a τ PARLA τ τ zzando τ (τ , τ τ τ
 τ) a ab τ a a τ sen az one de τ oca co.



Il grafico a U-maxx mostra la distribuzione delle distanze tra i punti del corpus. La scala a destra indica i valori delle distanze, che variano da 0 (blu scuro) a 600 (rosso scuro). La distribuzione è prevalentemente blu, con una banda di colori più caldi (giallo, arancione, rosso) che indica zone di maggiore vicinanza tra i punti, localizzate nella parte superiore del grafico.

In questo caso, i punti sono ordinati in base alla loro distanza da un punto di riferimento. La scala a destra indica i valori delle distanze, che variano da 0 (blu scuro) a 600 (rosso scuro). La distribuzione è prevalentemente blu, con una banda di colori più caldi (giallo, arancione, rosso) che indica zone di maggiore vicinanza tra i punti, localizzate nella parte superiore del grafico.

Johnson, K. (2000), *On the acquisition of the phonetic space*, *Journal of Phonetics*, 28, 227-254.

Kohonen, T. (2001), *Self-Organizing Map*, Springer, Berlin.

Lapinskas, R. (1988), *Phonetic adaptation in speech synthesis*, *JASA*, 94, 445-454.

Lisciani, A. & Sprosser, P. (2000), *Phonetic adaptation in speech synthesis: a data analysis*, in *Proceedings of International Neural Network Conference*, London, UK, pp. 305-308.