

## STATISTICHE DI FREQUENZA FONDAMENTALE PER UNO STESSO LOCUTORE IN DIVERSE CONDIZIONI DI PRODUZIONE

Antonio Romano

Centre de Dialectologie - Université Stendhal - Grenoble (France)

### SOMMARIO

Questo studio, basato sulle produzioni linguistiche di un solo individuo, si prefigge di valutare quanto spesso un dato valore di  $F_0$  ricorre nei movimenti melodici realizzati dalla voce di questo.

Sono state prese in considerazione le produzioni in italiano e in dialetto di un locutore salentino, in condizioni di riproduzione di contesto o di parlato spontaneo. Le circa 20000 misure di frequenza fondamentale ottenute (per complessivi 180 s netti di vibrazione delle corde vocali) sono state analizzate statisticamente e sottoposte a una prima verifica di normalità gaussiana. Le distribuzioni di valori sono state confrontate in base al tipo di produzione (parlato di laboratorio vs spontaneo, frasi affermative vs interrogative, parlato spontaneo dialogico vs narrativo etc.).

Le distribuzioni ottenute presentano visibili differenze e permettono di concludere, in prima analisi, in favore di un diverso uso dei registri di fonazione in funzione delle diverse necessità comunicative: la normalità gaussiana potrebbe essere verificata significativamente solo in un numero limitato di casi.

### INTRODUZIONE

Le caratteristiche della distribuzione statistica dei valori di frequenza fondamentale sono determinanti quando si vogliono effettuare dei test di significatività nel confronto di diversi insiemi di valori. In particolare, si è soliti assumere implicitamente che i dati in analisi seguano una distribuzione di tipo gaussiano (come discusso in [4], ciò accade anche nel caso dell'analisi statistica di insiemi di dati relativi alle formanti delle vocali). Per  $F_0$ , le pubblicazioni disponibili, relative alla verifica di tale ipotesi per diverse lingue, sembrano discordare. D'altra parte, l'esistenza di bersagli melodici perseguiti nel corso della produzione di messaggi, è tradizionalmente ricondotta a un tipo di analisi detta per *livelli*. Anche se uno speciale statuto spetta naturalmente alle *configurazioni* melodiche che caratterizzano il parlato - e quindi anche alla forma dei contorni di  $F_0$  che si sviluppano tra un livello e l'altro -, è ragionevole supporre che tali valori di riferimento "guidino" lo sviluppo melodico delle frasi, facendolo deviare da una aleatorietà che renderebbe gaussiana la distribuzione della variabile  $F_0$  (v. la maggior parte dei fenomeni "fisici" e le condizioni di applicabilità del teorema del limite centrale [2]). Se tali livelli esistono, anche solo per produzioni all'interno di un registro limitato, essi potrebbero emergere da un'indagine statistica.

## LA CURVA DI DISTRIBUZIONE DI $F_0$

L'analisi della curva di distribuzione di  $F_0$  è largamente diffusa in logopedia e foniatra dove è utilizzata nella diagnosi di voci patologiche e nello studio della psicomotricità della parola [5]. Tuttavia, anche l'approccio qui descritto non è nuovo: A. Prisse d'Avennes (1972) [8], nell'intento di ritrovare l'esistenza di tre *livelli* nel svolgimento melodico della frase affermativa in francese, affronta la delicata questione basandosi sull'osservazione dei tracciati della funzione di ripartizione, senza peraltro giungere a nessuna chiara conclusione in questo senso. Servendosi del test della retta di Henry, esegue la verifica di normalità dei valori di frequenza laringea di 10 locutori (150 s di registrazioni ciascuno relativi a frasi isolate di parlato di laboratorio) per poi concludere ugualmente in favore dell'esistenza di tre livelli.

Nel loro studio statistico di  $F_0$ , Boë et al. (1975) [3] ritrovano la normalità di tale distribuzione, per 20 locutori francesi, sulla base dell'accumulo dei valori di  $F_0$  superiori a certe soglie. I dati considerati per tale analisi sono ottenuti dall'estrazione di  $F_0$  da produzioni, di durata superiore ai 60 s, relative a corpus letti e frasi di laboratorio. Per ciascun locutore vengono analizzate più di 5000 misure.

Risultati contrastanti si ritrovano invece in [6]. In questo studio, il più preciso tra quelli a nostra disposizione, l'autore - oltre a dare una definizione univoca per la tonalità media e l'estensione del registro - analizza le produzioni di 10 locutori polacchi, sulla base dell'assunto che l'osservazione di 50 s di produzione naturale e non marcata emotivamente sono sufficienti per determinare le principali proprietà statistiche di  $F_0$ . Per ottenere ciò, considera però soltanto la lettura di un testo (brano di un giornale): inoltre i suoi risultati sono basati su misure di  $F_0$  non puntuali, ma determinate mediante calcolo di medie a breve termine eseguite negli intervalli di sonorità (l'autore definisce tra l'altro il *tasso di vocalizzazione* e ne stima l'intervallo di variazione tra il 40 e il 70% della durata complessiva della produzione). La dispersione dei valori di  $F_0$  è rappresentata su un istogramma con classi di 5 Hz. Come parametri significativi sono utilizzati la media, la deviazione standard (anche usata per determinare una stima dell'estensione del registro, delimitato al 95% da  $\pm 2\sigma$ ), l'asimmetria e la curtosi secondo Walker. Delle voci considerate, solo una presenta un'asimmetria non significativa. Per tutte le altre distribuzioni, che si allontanano invece significativamente dalla normalità, vengono precisate le caratteristiche di un locutore la cui  $F_0$  si distribuisce insistentemente su frequenze più alte della media e quelle di un locutore dalla distribuzione "bimodale", situazione a cui ci riportano alcuni dei dati da noi raccolti, che sembrano condurre piuttosto sulla pista di una sovrapposizione di gaussiane a media diversa, come lasciano pensare anche i dati di Hammarberg et al. (1980, v. p. 446) [5].

## ANALISI DEI DATI

Sono state prese in considerazione le produzioni in italiano e in dialetto di un solo locutore, la cui voce non risulta connotata da particolari tratti di differenziazione dalla media dei parlanti (e, quindi, ancor meno patologica), in condizioni di riproduzione di contesto o di parlato spontaneo. Il locutore, un salentino meridionale, della prov. di Lecce, di 28 anni al momento delle registrazioni (eseguite negli anni 1996 e 1997), presenta un idioletto particolarmente rappresentativo della varietà linguistica da lui parlata ma senza eccessiva presenza di tratti segmentali segnalanti su scala microregionale la provenienza del suo italiano regionale.

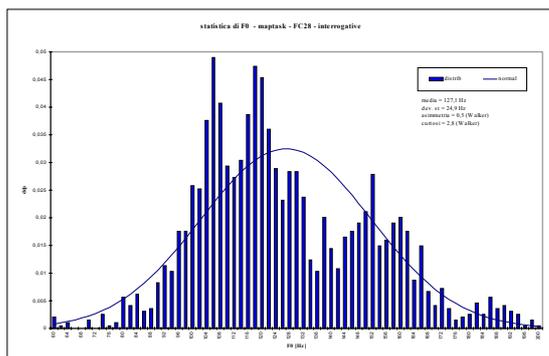


Fig. 1. Istogramma di distribuzione di 1943 valori di  $F_0$  misurati per frasi interrogative prodotte nel corso di una sessione di *map task*. Ad eccezione di alcuni valori erratici, i dati si ripartiscono all'interno di un registro di circa 130 Hz (da 70 a 200 Hz, con media a 127 Hz, dev. st. a 25 Hz e mediana a 122 Hz). Si può notare come la distribuzione sia notevolmente eccentrica rispetto alla densità di probabilità teorica. Delle concentrazioni di valori si presentano attorno a frequenze particolarmente usate (ad es.: 186, 160, **150**, 138, **117**, **107** Hz). Nell'intervallo 96-136 Hz si concentra il 62% dei valori contro il 28% nell'intervallo di stessa estensione 137-176 Hz. Dato l'elevato valore di asimmetria, la variabile, in questo caso, risulta non gaussiana.

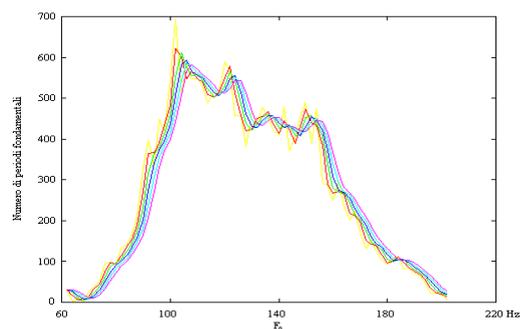


Fig. 2. Curva di distribuzione dei 19897 valori di  $F_0$  misurati nelle diverse condizioni di produzione. I valori si ripartiscono in un registro di circa 130 Hz (da 70 a 200 Hz, con media a 127 Hz, dev. st. a 25 Hz). Si può notare come la distribuzione sia sensibilmente diversa da una curva di densità di probabilità gaussiana, con una presenza accentuata di valori più ricorrenti di altri (intorno a circa **105** e **125** Hz, ma anche 138 e 150 Hz). Le diverse curve sovrapposte presentano varie fasi di *smoothing* della curva ottenuta con passo di 2 Hz, nel tentativo di ovviare, per esempio, agli effetti di riversamento nelle diverse classi (determinate dal passo prescelto) all'approssimarsi dei valori misurati alle soglie delle classi stesse, e ad altri possibili errori di approssimazione delle misure.

Le registrazioni sono state eseguite, mediante l'impiego di un *Sony Walkman Professional* e di un microfono *Sony ECM 909*, su cassette *Sony CDit*. Il materiale audio, numerizzato a 16kHz e su 16 bit, è stato in seguito sottoposto a procedure che hanno permesso di ottenere delle misure di  $F_0$  ciclo per ciclo. All'applicazione delle *routine* di estrazione, di tracciamento delle curve e di analisi statistica dei dati, appositamente sviluppate in ambiente Matlab™, ha fatto seguito una procedura di organizzazione dei dati in un foglio elettronico. I valori misurati per diversi tipi di produzione sono stati tenuti separati. In particolare il corpus si compone di 6 sezioni, tre delle quali relative a parlato di laboratorio (frasi a struttura fissa), una relativa a parlato stimolato a vocalizzazione continua, e infine due di parlato spontaneo tratto da un dialogo (*map task*, cfr. [1]) e da un monologo (basato sul racconto in [7]).

Sono state valutate media e deviazione standard tenendo distinte, ove possibile, le frasi affermative da quelle interrogative. Per ogni sessione sono stati conteggiati i cicli vocali e misurati tempo totale di parola e tempo di vibrazione delle corde vocali, ottenendo una stima del tasso di vocalizzazione. Una prima verifica di normalità è stata effettuata per ogni campione, rappresentando i dati in scala assoluta e analizzandone la distribuzione con passo lineare di 2 Hz (v. Fig. 1). Gli stessi dati, rappresentati su scala logaritmica, sono stati rianalizzati con passo di 1 semitono. Dalla sovrapposizione degli istogrammi ottenuti per ciascun campione sono state ricavate 11 curve di cui è stato valutato lo scostamento rispetto alla densità di probabilità teorica associata alle variabili corrispondenti, osservandone asimmetria e curtosi (test di Walker). Non sono stati eseguiti i test di normalità basati sulla retta di Henry o sul criterio di Lilliefors: la loro applicazione resta interessante soltanto nei 5 casi (su 11) in cui i test di Walker hanno dato valori positivi. In tutti gli altri casi, in cui lo scostamento dalla distribuzione teorica è evidente, la sua significatività è avvalorata da tali test.

## RISULTATI

Una netta suddivisione del materiale raccolto può essere ottenuta in base al tasso di vocalizzazione, molto più alto nel caso di produzioni spontanee e, naturalmente, anche nel caso delle frasi a vocalizzazione ininterrotta. Queste ultime hanno presentato i due principali casi di simmetria e curtosi "gaussiane" (oltre alle frasi affermative di parlato di laboratorio, con distribuzione significativamente diversa da quella delle interrogative), mostrando inoltre media e dev. st. dello stesso ordine di grandezza di quelle del parlato spontaneo, caratterizzato da asimmetria negativa.

Il risultato però forse più interessante riguarda l'esistenza di diverse concentrazioni di valori attorno a delle frequenze particolarmente usate. Tuttavia, se anche un livello tonale alto costituisce un livello di riferimento per il locutore, esso lo sarebbe soltanto come bersaglio raramente raggiunto (*target*), la curva melodica permanendo più spesso su valori disposti lungo la traiettoria necessaria per raggiungerlo. La ricerca di livelli diviene definitivamente improbabile nel caso della curva di distribuzione cumulativa riportata in Fig. 2, nella quale confluiscono certamente registri diversi. Anche tale curva presenta però distinte concentrazioni, oltre che una notevole eccentricità, che ne riducono le caratteristiche di normalità (cfr. [5]).

## CONCLUSIONI

Dall'analisi statistica effettuata sui valori di  $F_0$  misurati per produzioni diverse è emerso che le distribuzioni di  $F_0$  possono presentare visibili differenze, evidentemente determinate in funzione delle diverse necessità comunicative. Tali curve permettono di concludere, in prima analisi, in favore della presenza di diverse concentrazioni in corrispondenza di valori attorno ai quali la voce si soffermerebbe particolarmente: la normalità gaussiana risulterebbe quindi verificata significativamente solo in un numero limitato di casi. Tuttavia, se è lecito supporre l'esistenza di bersagli melodici (raggiunti o no) nel corso della produzione di messaggi, è anche ragionevole pensare che i valori di riferimento non siano assoluti ma siano piuttosto condizionati dal registro di produzione definito dagli usi e dagli stati d'animo del locutore. Indipendentemente dalla caratterizzazione di un abito linguistico o di semplici disposizioni fonatorie idiosincriche del soggetto osservato, qualcosa in questo senso è comunque emerso, accentuando la possibilità di una distribuzione non gaussiana dei valori di  $F_0$ .

## RIFERIMENTI BIBLIOGRAFICI

- [1] Anderson A.H. et al., *The HCRC Map Task Corpus*. Language and Speech, 34, 1991, 351-366.
- [2] Benedetto S. & Biglieri E., *Teoria delle probabilità e variabili casuali*. Torino, Bollati Boringhieri, 1980.
- [3] Boë L.J., Contini M. & Rakotofiringa H., *Étude statistique de la fréquence laryngienne. Application à l'analyse et à la synthèse des faits prosodiques du français*. Phonetica, 32, 1975, 1-23.
- [4] Cutugno F. & Falcone M., *Sulla distribuzione statistica dei valori formantici delle vocali*. Atti del XXIV Convegno Nazionale dell'A.I.A., 1996, 163-168.
- [5] Hammarberg B., Fritzell B., Gauffin J., Sundberg J. & Wedin L., *Perceptual and acoustic correlates of abnormal voice qualities*. Acta Otolaryngol, 90, 1980, 441-451.
- [6] Jassem W., *Pitch and compass of the speaking voice*. J. of the Int. Phonetic Ass., 1, 1971, 59-68.
- [7] Mayer M., *Frog, where are you?* New York, Pied Piper Book, 1969 (rist. 1980).
- [8] Prisse d'Avennes A., *Essai d'étude statistique de la détermination des niveaux intonatifs de la phrase énonciative française*. Bull. Inst. Phon. Grenoble, 1, 1972, 67-75.