

Il ruolo dell'udito nella comunicazione linguistica. Il caso della prosodia

Federico Albano Leoni

L'articolo tratta del ruolo dell'ascoltatore e dell'udito nella comunicazione parlata, con particolare attenzione all'intonazione. Dopo una breve presentazione del problema e una sommaria rassegna degli studi linguistici sull'argomento (§ 1), viene fornita una descrizione delle caratteristiche dell'apparato uditivo, della sinergia tra udito e voce e del ruolo degli schemi primitivi e degli schemi appresi nella decodifica del segnale linguistico (§§ 2-4). Nel § 5 viene discusso il ruolo dell'intonazione vista come un potente componente della comunicazione parlata, la cui importanza è universalmente riconosciuta, ma il cui studio e la cui descrizione appaiono ancora problematici a causa della sua intrinseca variabilità, della sua interconnessione con il continuum degli aspetti pragmatici e della mancanza di una tradizione di annotazione scritta. Nel § 6 vengono presentate alcune conclusioni e alcuni suggerimenti per un programma di lavoro sull'intonazione da condurre sul parlato spontaneo e tenendo conto delle verifiche percettive da parte dei parlanti*.

1. Premessa

È nota a tutti l'asimmetria tra il lessico del parlare, che è ricco, e il lessico del sentire linguistico, che è povero (De Mauro 1994b), al punto che per quest'ultimo spesso manca, come in italiano o in inglese, un verbo specifico. Uno dei motivi, forse il principale, di questa asimmetria risiede, a parere di molti (p. es. Ryalls 1996: 3-4), nella profonda differenza tra il parlare e l'udire: il primo è in gran parte esterno, visibile, percepibile e autopercepibile; il secondo è interiore, invisibile, sfuggente. Posso vedere chi parla (anche se non sento quello che dice), ma non posso vedere l'ascolto; posso osservare e percepire alcuni movimenti del mio apparato fonatorio (laringe, bocca, lingua, labbra), ma non posso vedere o percepire i movimenti dell'apparato uditivo (timpano, ossicini, perilinfia, membrana basilare ecc.) mio o dei miei interlocutori. Sembra dunque ragionevole supporre che un processo palese ed evidente sia più presente alla coscienza dei parlanti e arrivi quindi a una rappresentazione verbale più articolata di un processo meno palese ed evidente.

Questa asimmetria ha altre conseguenze, oltre a quella della diversa rappresentazione dei due processi sul piano lessicale. Qui ci interessa quella per cui i processi della comunicazione parlata, raffi-

gurati, ad esempio, nel circuito della *parole* di Saussure (1922: 27), sono stati studiati prevalentemente, a volte esclusivamente, osservando l'emittente o osservando il messaggio, cioè rispettivamente il parlante o il testo, come si può facilmente verificare osservando a caso manuali o opere di riferimento che riflettono il senso comune corrente in linguistica e in fonologia (come Akmajian *et al.* 1995; Simone 1995; Nespor 1993; Kenstowicz 1994; Goldsmith 1995; Laver 1991 ecc.)¹. È come se una parte non trascurabile della linguistica manifestasse un pensiero “primarily anchored to saying-without-listening” (Corradi Fiumara 1990: 3).

Naturalmente nessuno può pensare che i linguisti debbano occuparsi obbligatoriamente anche dell'udito o dell'ascolto o del ricevente, ma non si può non osservare una preferenza, domandarsi quali ne siano le ragioni e quali ne siano gli effetti sulle nostre conoscenze.

Un motivo di questa preferenza è certamente il fatto che la rappresentazione scritta della lingua, oggetto di studio obbligato in passato e prevalente ancora oggi, valorizza e stabilizza il prodotto, il testo, ma non induce a studiare la ricezione uditiva. Un secondo motivo, anche importante, è dato dalla natura interiore e invisibile della percezione, che richiede tecniche e metodi di osservazione, elicitazione e studio che la linguistica ritiene estranei ai suoi apparati.

Questa preferenza si manifesta in maniera molto chiara anche negli studi fonetici, dove per molto tempo è stato evidente il predominio della descrizione articolatoria (cioè del produttore) finalizzata alla descrizione e alla trascrizione del prodotto (cioè del testo). A partire dai primi decenni del Novecento è andato diffondendosi lo studio del versante acustico del messaggio (nato nella sua forma moderna con l'abate Rousselot alla fine dell'Ottocento e reso oggi facile dalla diffusione di strumenti economici e di uso relativamente semplice). Ma anche i risultati di queste analisi sono usati prevalentemente o per descrivere la struttura dei suoni in sé (cioè, ancora una volta, il prodotto), o per descrivere la relazione tra la struttura acustica e i gesti articolatori che l'hanno generata (cioè, ancora una volta, il produttore). Insomma, nella ricerca linguistica l'udito e il ricevente sono spesso trascurati perché in passato erano considerati argomenti propri dei filosofi e oggi sono considerati argomenti degli psicologi o di quanti si occupano della fisiologia della percezione².

In questo articolo non affronterò la storia di questo interessante problema³ ma mi limiterò a qualche osservazione preliminare, prima di passare all'argomento che vorrei trattare.

Tra la fine dell'Ottocento e i primi del Novecento sembrò che l'udito e il ricevente richiamassero l'attenzione dei linguisti. Saussure,

sia pure solo per cenni, aveva intuito la necessità di una svolta decisa. Ecco due passi molto istruttivi (sempre attuali i commenti di De Mauro alle note 113 e 114 dell'edizione italiana; ricordiamo che per Saussure l'aggettivo *acoustique* significa 'uditivo'):

Beaucoup de phonologistes s'attachent presque exclusivement à l'acte de phonation, c'est-à-dire à la production des sons par les organes (larynx, bouche, etc.), et négligent le côté acoustique. Cette méthode n'est pas correcte: non seulement l'impression produite sur l'oreille nous est donnée aussi directement que l'image motrice des organes, mais encore c'est elle qui est la base naturelle de toute théorie (Saussure 1922: 63).

La délimitation des sons de la chaîne parlée ne peut donc reposer que sur l'impression acoustique; mais pour leur description, il en va autrement. Elle ne saurait être faite que sur la base de l'acte articulaire, car les unités acoustiques prises dans leur propre chaîne sont inanalysables (*ibid.*: 65).

Ancora più esplicite sono alcune sue osservazioni contenute nel manoscritto di Harvard (Marchese ed. Saussure 1995: 98-100, 105, 113, e in particolare p. 98, da cui cito):

Un acte phonatoire est un fait (ou plutôt) un ensemble de faits physiologiques correspondant à un fait phonétique déterminé. Les fait phonétique nous étant a son tour donné par la sensation auditive, c'est d'après cette dernière seule que nous fixons les actes phonatoires.

Saussure aveva dunque assunto una posizione decisa a favore del primato degli aspetti uditivi e la sua intuizione aveva trovato conferma nella ricerca fonetica successiva, nei modelli sulle strutture foniche delle lingue, negli studi sui disturbi del linguaggio, nelle osservazioni sul fonosimbolismo (oltre che, naturalmente, negli studi di psicolinguistica). Rappresentanti illustri di questa linea erano stati, tra gli altri, fonetisti e linguisti come Gunnar Fant e Roman Jakobson. Mezzo secolo dopo Saussure, Jakobson & Halle (1956: 33-34) si erano espressi con grande chiarezza:

In order to decode the message, its receiver extracts the distinctive features from the perceptual data. The closer we are in our investigation to the destination of the message, the more accurately can we gauge the information conveyed by the sound chain. This determines the operational hierarchy of levels in their decreasing perti-

nence: perceptual, aural, acoustical and motor (the latter carrying no direct information to the receiver except for the sporadic help of lip-reading) [...] ⁴.

Questa linea acustico-uditiva non ebbe però successo tra i linguisti e subì una battuta d'arresto nel corso degli anni Sessanta, quando, per vari motivi, la ricerca fonetica tornò ad orientarsi verso gli aspetti articolatori. Questo cambiamento di direzione è ben mostrato da tre fatti emblematici.

Il primo fu l'affermarsi di una teoria della percezione nota come *Mothor Theory* (Liberman *et al.* 1963, 1985) che riconduce i processi della percezione linguistica alla riproduzione interiore dei gesti articolatori ⁵, recuperando così, forse senza saperlo, una tradizione filosofica e psicologica ottocentesca che risaliva a Maine de Biran (Albano Leoni & Dovetto 1996) ⁶. Naturalmente Liberman era lontanissimo dal ritenere che la sensazione uditiva fosse irrilevante ai fini della percezione linguistica e la sua posizione era molto più complessa, come appare dalla sua autobiografia scientifica, posta come introduzione a un volume del 1996 che raccoglie i suoi scritti più significativi (Liberman 1996: 1-44). Ma ugualmente essa ha facilitato, o rinforzato, la tendenza alla rimozione degli aspetti uditivi.

Il secondo è che la fonologia generativa, fin dal suo nascere (Chomsky & Halle 1968), ridimensionava drasticamente il ruolo dell'udito:

[...] what the hearer 'hears' is what is internally generated by rules. That is, he will 'hear' the phonetic shape determined by the postulated syntactic structure and the internalized rules (Chomsky & Halle 1968: 24)

Probabilmente in conseguenza di ciò, pur conservando l'impostazione binarista nella descrizione dei suoni, Chomsky e Halle cancellavano dalle matrici ogni tratto acustico-uditivo e generalizzavano quelli articolatori.

Il terzo è la pubblicazione del lavoro di Lenneberg (1967) che, pur dichiarando nell'introduzione che non avrebbe trattato dei meccanismi concreti della produzione e ricezione del linguaggio (sia perché troppo tecnici, sia perché irrilevanti ai fini della teoria che intendeva esporre), dedica invece uno spazio molto ampio alla anatomia e fisiologia dell'apparato fonatorio (cap. 2, II; 3, II-V) ma solo una veloce e marginale allusione all'udito (cap. 3, V, 1b-c), pur trattando dei fondamenti biologici del linguaggio. Ciò potrebbe sorprendere, ma se si considera che in

appendice al volume di Lenneberg compariva un saggio di Chomsky, il silenzio sulla percezione appare abbastanza ovvio.

Naturalmente questi autori non possono ignorare che l'udito è l'ingresso principale per il destinatario del messaggio, ma ritengono che esso non contribuisca agli aspetti profondi della comprensione o perché questi sono basati su schemi innati, o perché essi sono strettamente associati ai meccanismi neuromuscolari della produzione.

Si chiudeva così un circolo che, almeno in linguistica, non si è ancora riaperto.

L'ascolto e la percezione uditiva della lingua sono invece oggetto di studio e di discussione in un ambito disciplinare che risulta dalla felice commistione di psicoacustica, psicolinguistica, neurolinguistica e fonetica, come appare, p. es., dai lavori di Hess (1983), Moore (1988), Gernsbacher (1994), Goodman & Nusbaum (1994), Hardcastle & Laver (1997, specialmente nei contributi di Delgutte, di Moore e di McQueen & Cutler) e di Pickett (1999). Da qui prenderò spunto per fare qualche riflessione sul ruolo dell'udito e del ricevente dal punto di vista linguistico, senza presumere di dire cose nuove, ma solo sistemando e accostando notizie e dati che vengono da settori di studio diversi ⁷.

Cercherò quindi di ricordare non solo che l'udito svolge un ruolo importante, almeno da comprimario, nella comunicazione audioverbale, ma anche che in esso si incarna con particolare evidenza quello che Simone (1992: 47-48) ha chiamato il "principio del determinismo fisico", cioè l'insieme degli effetti che la natura psico-fisica dei nostri apparati di produzione e di ricezione provoca sulla struttura e l'uso delle lingue. In questo senso, dunque, questo lavoro vorrebbe essere anche un contributo indiretto alla discussione sulla iconicità linguistica, intesa nel senso di 'limitazione dell'arbitrarietà'. So bene che questa discussione, mutati i termini, è antica (sintesi in Simone 1992 e 1995, Gensini 1993), ma qui considererò la questione solo nella prospettiva odierna.

Cercherò infine di trarre qualche conclusione, partendo dall'udito e dalla prosodia, su alcuni aspetti generali della decodifica e della semiosi linguistiche. Il punto di vista che qui vorrei presentare può essere espresso con le parole di Rosen & Fourcin (1986: 373):

In evolution, the senses of vibration and hearing preceded speaking, and so it is inescapable that basic auditory constraints moulded the way speech developed. In short, trying to understand the auditory processes involved in the perception of speech is likely to lead not only to a better understanding of hearing, but also, of speech itself.

2. Perché l'udito è importante per i linguisti?

La voce e l'udito concorrono alla realizzazione della comunicazione parlata ma non sono del tutto alla pari. Se si osservano l'anatomia e la fisiologia degli apparati vocale e uditivo si vede che l'apparato vocale è costituito da organi (polmoni, bronchi e trachea, laringe, faringe, cavità orale, lingua, denti, labbra, cavità nasali) che hanno come funzioni primarie la respirazione e/o la nutrizione e per i quali il parlare è una funzione secondaria. L'apparato uditivo è costituito invece da organi (padiglione, timpano, ossicini, coclea, organo del Corti) che hanno come funzione primaria quella dell'udire (e del localizzare la sorgente sonora). L'ascolto è dunque coesenziale alla nostra biologia, il parlare no. Non è inoltre privo di rilievo il fatto che se non si sente non si parla (i sordomuti sono in realtà sordi): il parlare implica l'udire (ma l'udire, almeno in teoria, non implica il parlare) ⁸.

L'apparato uditivo è completo e funzionante prima della nascita: il feto sente e discrimina rumori e suoni interni ed esterni già al settimo mese della gestazione e il neonato sente e discrimina suoni, voci e ritmi con grande finezza (Mehler & Dupoux 1990 [1992]: 163-223; Hawkins 1999; molta bibliografia in Pennisi 1994: 268-75). L'apparato vocale, invece, si completa dopo la nascita, quando la laringe si abbassa e consente la formazione di una cavità faringale relativamente ampia, fondamentale per la produzione di vocali diverse. Inoltre esso diventa perfettamente operante solo dopo un tirocinio piuttosto lungo ⁹.

Queste osservazioni servono a dire che nella architettura degli strumenti della comunicazione verbale umana, l'udito ha una preminenza biologica e funzionale, anche se, naturalmente, tra le nostre capacità vocali e le nostre capacità uditive esiste una stretta correlazione neurologica, oltre che funzionale ¹⁰.

Tenendo presenti queste premesse, nel prossimo paragrafo cercherò di mostrare come si manifesta, in concreto, l'interazione tra voce e udito.

3. Il campo uditivo umano e le caratteristiche acustiche del segnale linguistico

La fig. 1 rappresenta il campo uditivo umano, cioè l'insieme dei suoni mediamente percepibili da un umano normoudente ed è ricavato dalle risposte di numerosi soggetti alla somministrazione di segnali acustici variati in frequenza e intensità.

La rappresentazione consiste in un diagramma in cui ogni punto è determinato da un valore in frequenza (in ascissa) e un valore in intensità (in ordinata). L'intervallo di frequenza va da circa 16 a circa 20.000 Hz, mentre l'intervallo di intensità va da -4 a 140 dB. Il limite superiore del campo uditivo è la soglia del dolore, che qui non ci interessa; il limite inferiore, cioè la soglia di udibilità, mostra invece il variare dell'intensità necessaria per percepire stimoli acustici alle diverse frequenze. Se lo si osserva con attenzione, si vede che la zona in cui il rapporto energia/frequenza è ottimale, cioè quella in cui il nostro orecchio

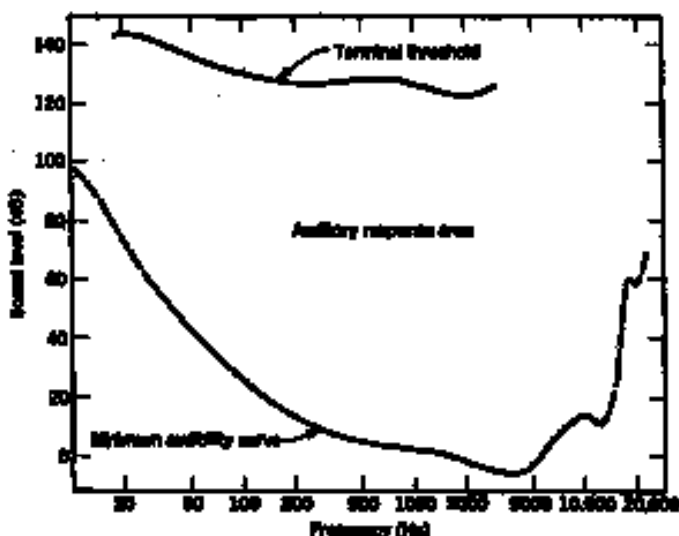


Fig. 1. Il campo uditivo umano (da Moore 1988: 48).

percepisce suoni anche di debole e debolissima intensità, va *grosso modo* dai 2000 ai 5000 Hz (dove vengono percepiti suoni al livello della soglia minima di energia); ma anche la fascia che va da 200 a 7000 Hz è molto sensibile perché per percepire questi suoni è sufficiente un'energia pari a circa 10 dB¹¹.

Il campo uditivo rappresenta dunque il primo, ovvio vincolo alla qualità acustica dei suoni linguistici i quali devono porsi necessariamente all'interno dei suoni udibili e preferibilmente all'interno di quella fascia di frequenza in cui risultino più facilmente percepibili.

A questa osservazione ne va aggiunta una seconda. Se il campo uditivo può essere considerato uno spazio fisico continuo, in cui ogni

punto è definito da due valori 'oggettivi' (intensità e frequenza), dal punto di vista delle capacità umane di discriminare tra suoni diversi per altezza e volume (capacità fondamentale anche per gli usi linguistici)¹², esso è articolato al suo interno in modo discontinuo, e le modalità della discriminazione si distribuiscono lungo scale 'soggettive', determinate, ancora una volta, dalla anatomia e dalla fisiologia dell'apparato uditivo, e in particolare dalla caratteristica della tonotopicità e dal fenomeno delle bande critiche (Zwicker 1961).

I tests psicoacustici mostrano inoltre che la capacità di discriminazione dell'orecchio decresce al crescere della frequenza del segnale. La figura 2 presenta in ascissa le frequenze del segnale e in ordinata la differenza in Hz necessaria perché due segnali vengano riconosciuti come diversi o perché venga riconosciuta una variazione di altezza: alle frequenze basse, fino a 250 Hz, il nostro orecchio è in grado di discriminare tra toni puri che differiscono in frequenza anche per una frazione di Hz. Al crescere della frequenza aumenta la differenza necessaria perché due suoni vengano percepiti come diversi. Considerazioni simili possono essere fatte per la discriminazione di differenze di intensità e di durata. Questo meccanismo ha il suo fondamento nella anatomia e nella neurofisiologia dell'orecchio interno.

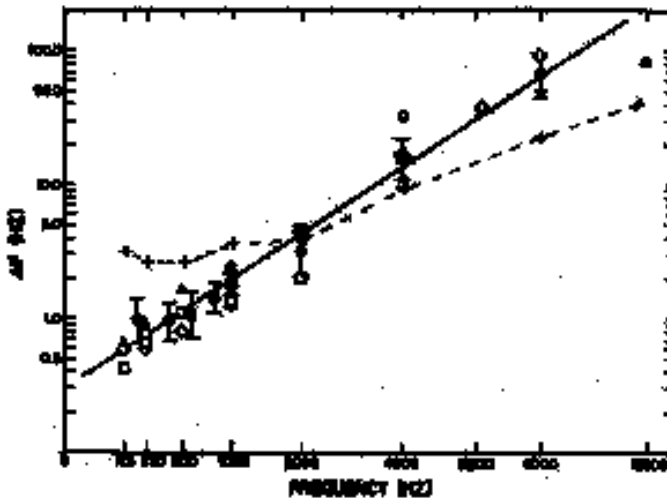


Fig. 2. Capacità di discriminazione dell'orecchio in rapporto alla frequenza del segnale (da Moore 1988: 160 con modifiche).

I dati che ho presentato finora sono basati sull'osservazione di toni puri. Essi forniscono indicazioni sulle caratteristiche fisiologiche

della percezione, dalle quali non si può prescindere, ma non si può dimenticare che i toni puri sono praticamente assenti dal segnale linguistico che è costituito da oscillazioni complesse quasiperiodiche e aperiodiche.

Lo studio della percezione e riconoscimento dei suoni linguistici è molto più complicato. Prescindendo qui dal problema generale della percezione primitiva e della percezione basata su schemi (v. *infra* p. 13), basterebbe ricordare, a proposito della variabilità osservata degli spettri dei suoni vocalici, la questione del timbro metameroico (Bregman 1990: 122)¹³, o la questione della percezione del *pitch*, ossia del correlato uditivo della frequenza fondamentale (che nei segnali linguistici è funzione delle variazioni di velocità del meccanismo laringeo). Qui mi soffermerò sul secondo problema.

La psicoacustica mostra che la percezione del *pitch* in un segnale periodico complesso avviene secondo modalità parzialmente diverse, a seconda che esso si trovi a frequenze basse o a frequenze alte. Qui considererò il primo caso perché esso è appunto quello generato dal meccanismo laringeo. Il *pitch* di bassa frequenza viene riconosciuto o perché è fisicamente presente una prima armonica di ampiezza percepibile, o perché viene ricostruito dall'orecchio attraverso un'analisi delle prime cinque armoniche. Questo secondo procedimento è importante perché la frequenza fondamentale della voce umana, per esempio di un maschio adulto, si aggira intorno ai 150 Hz, e può arrivare ai 400 Hz nella voce di un bambino¹⁴ e può essere di debole intensità. Se ora si torna al campo uditivo si vede che le frequenze al di sotto dei 500 Hz sono percepibili solo ad una intensità pari ad almeno 20 dB. Ma la terza, quarta e quinta armonica di un segnale complesso che abbia la prima armonica a 150 Hz si collocano tra 450 e 750 Hz, cioè in una fascia di frequenze al tempo stesso meglio percepibili e nelle quali l'orecchio ancora discrimina variazioni anche piccolissime. È questo il complesso meccanismo che consente il riconoscimento dell'intonazione, sul quale tornerò più avanti¹⁴. L'orecchio è dunque attrezzato per percepire e discriminare tanto toni puri periodici, quanto suoni complessi quasi-periodici, affetti da rumore e da piccole oscillazioni di frequenza (dette *jitter*) intorno a un valore centrale, nei quali alcune componenti possono essere mascherate o cancellate o troppo deboli ed è dunque in grado di sopperire alle imperfezioni dell'apparato fonatorio:

It is easy to show that, with respect to both measuring range and accuracy of the system, the ear outperforms the speech tract by far; the main constraints on the range of F0 and the accuracy of the

realization of a periodic signal in speech are due to the process of speech production (Hess 1983: 63)

Si deve quindi osservare che la pertinenza percettiva delle variazioni di uno stato fisico prodotte dalla fonazione (cioè la base materiale su cui si esercita la capacità linguistica di distinguere) non è determinata solo dalle variazioni in sé, ma anche dai meccanismi psicoacustici. Quindi per capire come funziona un messaggio verbale non è sufficiente studiare tutte le caratteristiche dei gesti articolatori che l'hanno prodotto, o tutte le caratteristiche spettrali che ne sono il risultato, ma è necessario valutare queste ultime sulla base della loro significanza percettiva.

No matter how systematically a phenomenon may be found to occur through a visual inspection of F0 curves, if it cannot be heard, it cannot play a part in communication ('t Hart *et al.* 1990: 25).

In altre parole, è lecito affermare che l'apparato uditivo condiziona due volte il meccanismo della comunicazione audioverbale: una prima volta, come abbiamo visto, determinando lo spazio fisico complessivo all'interno del quale la comunicazione avviene in modo ottimale; una seconda volta, determinando il livello minimo della percepibilità delle variazioni fisiche.

Il livello minimo della percepibilità delle variazioni fisiche non coincide però necessariamente con il livello minimo della loro pertinenza linguistica. Ciò significa, ancora riguardo al *pitch*, che la percezione di una variazione della melodia di una frase richiede una variazione della frequenza superiore alla soglia della discriminabilità meramente psicoacustica (Hess 1983: 78-79).

Se si osservano ora i tracciati spettroacustici ¹⁶ di un numero grande a piacere di enunciati linguistici naturali, si vede facilmente che l'energia non si distribuisce a caso nello spettro.

Se prendiamo le figure 3 e 4 come esempio rappresentativo della struttura acustica di enunciati umani, si vede che la banda di frequenza tra 200 e 5000 Hz è quella in cui si colloca la parte maggiore e più rilevante dei suoni linguistici prodotti dall'apparato fonatorio. Al di sopra di questo limite l'energia è in genere poca, limitata ad alcune fricative e, comunque, grazie alla ridondanza, non strettamente necessaria al riconoscimento dei suoni, come mostra, tra l'altro, la comunicazione telefonica analogica nella quale il canale fisico taglia le frequenze superiori a 2500/3000 Hz e inferiori a 500 Hz.

La correlazione tra prodotto articolatorio e capacità uditive è

dunque particolarmente evidente. L'energia acustica si trova, per così dire, là dove l'orecchio è in grado di riceverla meglio. La messa in relazione delle capacità uditive umane e della distribuzione dell'energia acustica nel parlato mostra una ottima sinergia tra produzione e ricezione.

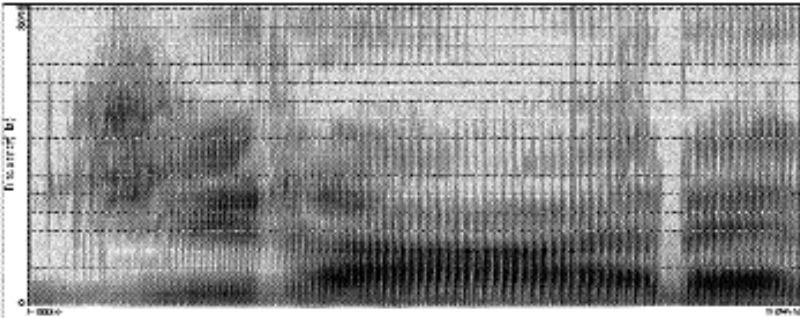


Fig. 3. Diagramma tempo/frequenza della parola italiana 'girare'.

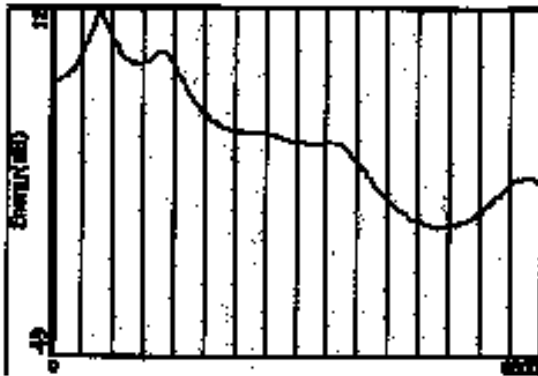


Fig. 4. Diagramma frequenza/energia del punto centrale della vocale [a] indicato dal puntatore nella fig. 3.

4. La percezione del parlato

Il problema oggi in discussione per quanto riguarda la percezione del messaggio fonico è se questa dipenda a) dalla attivazione di un modulo uditivo innato e generale, b) dalla attivazione di un modulo uditivo specifico innato preposto alla percezione di suoni linguistici già categorizzati, c) dalla attivazione di un modulo complesso che

porta a riconoscere i suoni a partire dai gesti articolatori necessari per produrli (che qui non considererò per quanto ho detto sopra alla nota 6).

Il problema, che pure avrebbe un risvolto filogenetico, viene in genere studiato osservando il comportamento infantile (a partire dalla fase perinatale), il comportamento di animali, il comportamento di adulti in contatto con lingue diverse dalla lingua madre. Il primo punto di vista è quello che oggi richiama la maggiore attenzione.

Mehler & Dupoux (1992: 215) dichiarano che “oggi prevale l’idea che il sistema di percezione della parola sia innestato su quello dei suoni acustici”. Questa posizione si basa sull’assunto di una priorità dell’apparato uditivo generale, a partire dal quale si sviluppa la capacità di percepire e riconoscere il parlato. Sembra infatti accertato che la capacità di categorizzare suoni (linguistici e non) da parte del feto e del neonato si costruisce su esperienze uditive prelinguistiche e quindi biologicamente determinate.

Miller & Eimas (1994: 43, 48) sono invece più cauti e ritengono che non si possa ancora decidere con certezza

“whether the mechanisms underlying phonetic perception form a system specialized for the processing of speech or whether processing is accomplished by the general auditory system. [Allo stato attuale] the evidence at hand does not provide a compelling case for choosing between an account of underlying mechanism based on specialized processing and an account based on the operation of the general auditory system”.

Nusbaum & Goodman (1994: 328) sembrano più vicini a posizioni ‘innatiste’ che riducono il ruolo dell’udito a quello di selezionare a posteriori i foni più utili partendo da un repertorio universale predefinito e precategorizzato:

The infant may be born with few a priori expectations about the way to distribute attention in analyzing the acoustic structure of speech. Experience of listening to speech directs this attention toward those properties that are most relevant for the phonetic and phonological processing of speech in the infant’s native language environment. This would be consistent with the data that suggest that infants are born with a universal phonetic inventory and that perceptual experience eliminates the nonnative contrasts from this inventory, maintaining the native contrasts alone”.

Pinker (1994: 256-57), che peraltro non è uno specialista e basa

le sue considerazioni su esperimenti di Mehler, Eimas e Jusczyk, non si pronuncia esplicitamente sulla esistenza di un inventario universale innato, ma dà un grande rilievo al tirocinio uditivo prenatale e, in particolare, alla percezione della melodia.

La questione è dunque se l'udito sia un ricettore passivo, che assegna le esperienze sensoriali alle categorie innate preformate e preesistenti, o se invece esso sia attivo e concorra a formare le categorie. Ma anche scegliendo questa seconda ipotesi, che per molti aspetti sembra più forte, la questione rimane complessa non appena si passi dall'osservazione del neonato a quella dell'adulto.

Gli studi di psicoacustica mostrano infatti che l'attività di decodifica del percepito uditivo è basata su due meccanismi: un meccanismo primitivo, innato, basato sulle capacità uditive in sé, e un meccanismo basato su schemi, appreso, idiolinguistico (Bregman 1990: *passim*). Il primo fornisce la base al secondo¹⁷, ma il secondo integra la base sensoriale. Su questo punto tornerò più avanti.

Quindi, malgrado differenze di posizione, la sollecitazione uditiva sembra essere il punto di partenza dello sviluppo delle capacità linguistiche (intese come capacità di codifica e decodifica audioverbali e non necessariamente come capacità cognitive generali) nel bambino.

5. Il caso esemplare dell'intonazione

Il senso di tutte queste osservazioni diventa particolarmente evidente quando esse vengono calate nell'analisi prosodica.

Fino ad anni recenti la linguistica aveva dedicato alla prosodia, cioè all'insieme dei fenomeni ritmici e intonativi delle lingue, un'attenzione minore di quella dedicata non solo alla fonetica e fonologia segmentali ma anche alla morfologia, alla sintassi e alla semantica. Con l'eccezione della scuola britannica, che da tempo aveva sviluppato una tradizione di studi prosodici su base uditiva (cito come esempio rappresentativo Cruttenden 1986), i lavori dedicati a questo settore erano pochi, come si può vedere osservando, ad esempio, le date di pubblicazione dei lavori citati da Bolinger (1989), che aumentano esponenzialmente a partire dagli anni '60; inoltre questi lavori, dedicati per lo più all'inglese, erano basati su osservazioni qualitative, spesso non sistematiche, per lo più concentrate sull'osservazione del rapporto tra alcune funzioni elementari sintattico-pragmatiche (come l'interrogazione) e alcuni profili intonativi. Solo da pochi anni gli studi di prosodia, e in particolare quelli dedicati all'intonazione sono aumentati drasticamente, grazie all'interesse per questi aspetti da

parte delle fonologie più recenti (autosegmentali, prosodiche)¹⁸, al punto che ora la situazione si è ribaltata. Oggi chi studia il parlato riconosce alla prosodia un ruolo determinante tanto nella ideazione ed esecuzione della stringa, quanto nella sua segmentazione e interpretazione da parte dell'ascoltatore.

La prosodia è un fenomeno difficile da studiare per molti motivi. Il primo è certamente il fatto che tutte le variabili fisiche che la determinano (il tempo e l'intensità per il ritmo, la frequenza per l'intonazione) si dispongono lungo un *continuum* non segmentato *a priori* (come è invece, grazie alla scrittura, per fonologia, morfologia, lessico e, in parte, sintassi)¹⁹.

Il secondo motivo è che l'intonazione, componente importantissimo della prosodia, non è analizzabile in una prima e seconda articolazione (anche Bertinetto 1981: 27): un segmento intonativo di prosodia ritagliato dal suo contesto è in sé totalmente privo di significato e di funzione, perché i valori che determinano la prosodia, e dunque anche l'intonazione, sono sempre e tutti radicalmente relativi, valutabili e interpretabili solo in rapporto a ciò che segue e a ciò che precede all'interno dell'intera unità prosodica considerata. Quindi, di un segmento vocalico in sé, del quale posso descrivere in modo 'oggettivo' l'altezza, la durata e l'intensità (oltre che il timbro), non posso dire se sia linguisticamente acuto o grave, lungo o breve, intenso o debole. Inoltre, allo stato attuale delle nostre conoscenze, abbiamo ancora difficoltà a segmentare i profili intonativi di molti enunciati naturali, e ad assegnare a una determinata porzione una determinata funzione o un determinato contenuto comunicativo: in molti casi è come se il profilo intonativo e il corrispondente contenuto semantico e pragmatico, pur perfettamente integrati nel sistema comunicativo, si muovessero in una dimensione olistica²⁰.

Un terzo motivo è la natura particolare della fortissima variabilità prosodica. La variabilità è naturalmente una caratteristica fondamentale di tutte le manifestazioni foniche (e non solo foniche) delle lingue. Ma, mentre la variabilità nella realizzazione dei segmenti che costituiscono la stringa è, in qualche modo e almeno in parte, predicibile e riconducibile a una qualche altra forma di variazione (diatopica, o diastratica, o diafasica che sia), per cui essa non concorre a determinare il contenuto semantico dell'enunciato²¹, la variazione prosodica, a parità di condizioni diatopiche, diafasiche e diastratiche, è sempre il riflesso di una sia pur lieve differenza nelle intenzioni comunicative del parlante: ad una variazione prosodica corrisponde sempre una variazione semantico-pragmatica dell'enunciato, perfettamente chiara a chi ascolta²². Ciò è detto molto chiaramente in 't Hart *et al.* (1990: 110-114):

Intonation features have no intrinsic meaning [...] We speculate that these choices [cioè quelle fra i vari contorni] are influenced by the attitudinal meaning that a speaker wants to add a literal meaning of his utterances. But the actual encoding of his attitudinal meaning into an individual pitch contour is evidently governed by so many pragmatic and situational factors that we are still looking for a manageable experimental paradigm in which to tackle this complicated issue.

Ma, infine, ciò che fa della prosodia qualche cosa di speciale tra gli strumenti della comunicazione audioverbale è che, mentre una stringa segmentale non può esistere se non dentro uno schema prosodico, uno schema prosodico può esistere senza contenere una sequenza segmentale (come quando mugoliamo a bocca chiusa una melodia [mmmmm]), o può esistere appoggiandosi a una sequenza segmentale artificiale e asemantica conservando una sua capacità comunicativa, come sapeva, p.es., Italo Calvino²³. Per usare termini antichi, ma molto efficaci, la *phoné* esiste senza la *diálektos*, ma la seconda non può esistere senza la prima (Laspia 1997: 59-69).

La prosodia, oltre a disporre di una sua grammatica (perfettamente nota a chi parla e a chi ascolta, ma ancora poco nota ai linguisti), che consente la trasmissione di un numero grandissimo di sensi a partire da una stessa sequenza segmentale, è anche in grado di dominare tanto la componente sintattica, grazie ai complessi processi della focalizzazione e della messa in rilievo, quanto la componente semantica (la prosodia può far sì che una doppia affermazione neghi)²⁴.

Lo studio della prosodia si colloca dunque in uno spazio delimitato da un lato dalla psicoacustica e dall'altro dalla rilevanza comunicativa (t Hart *et al.* 1990: 5).

Come si vede, questa componente di straordinaria importanza nella comunicazione orale: a) si realizza attraverso un meccanismo articolatorio estremamente economico, basato su una variazione di poche decine di Hz rispetto alla frequenza fondamentale propria di ciascuno e determinata dalla anatomia individuale, di una variazione che può anche essere di pochi millisecondi nella durata dei segmenti, e di piccole variazioni nella pressione dell'aria espiratoria a ridosso dei diversi ostacoli glottidali o supraglottidali; questo meccanismo a) è perfettamente congruente con le capacità uditive; b) riflette dinamiche fonatorie generali di ordine probabilmente biologico: la sua articolazione in unità tonali è legata ai cosiddetti 'gruppi espiratori'; le unità tonali si concludono sempre (tranne che nei casi di marcatezza) con una declinazione naturale di tutti gli indici (intensità e frequenza

tendono a zero e l'eloquio rallenta), e anche la marcatezza (cioè l'andamento terminale ascendente) non è che il rovesciamento di un andamento naturale e dunque anch'esso non 'arbitrario'.

La questione della segmentazione del discorso in unità tonali e della individuazione del punto focale all'interno di ciascuna unità tonale è forse l'aspetto determinante per individuare la centralità del ruolo dell'udito nella decodifica del parlato. Ciò è confermato anche dall'importanza della prosodia nell'apprendimento linguistico da parte dei bambini e nello sviluppo della loro capacità di segmentare il *continuum* della catena parlata (processo che avviene senza che il bambino abbia ancora alcun controllo consapevole sul meccanismo laringeo, e quando il suo controllo degli schemi prosodici della sua lingua è ancora ad uno stadio aurorale). Hawkins (1999: 195) presenta con grande chiarezza gli aspetti ontogenetici del problema:

The most global properties of the baby's to-be-native language are prosodic, and some of these may even be learned at or before birth. Rhythm and pitch patterns can be heard in utero, since they are what is left in the signal when the mother's voice reaches the uterus, low-pass filtered through her body. [...] their response patterns were the same regardless of whether the speech was heard unfiltered, or low-pass filtered so that only the prosodic information remained [...] Other evidence suggests that babies are sensitive to the language-specific prosodic properties that cue clause boundaries by about 6 months of age and to those marking phrasal boundaries by 9 months.

Quando gli esperimenti vengono condotti con materiale acustico più complesso (ma ancora artificiale) appare subito che le modalità di discriminazione di altezza sono molto complesse e tengono conto della durata del segnale e anche delle armoniche di frequenza più alta. Quando poi si passa ad osservare la percezione e la discriminazione dei movimenti del *pitch* nel parlato naturale (condizioni sperimentali delicatissime e ancora poco praticate), agli occhi dell'osservatore appare una specie di circolo vizioso: è certamente la percezione uditiva che mette in moto il processo di analisi e di decodifica, ma quasi contemporaneamente il segnale linguistico mette in moto attività cognitive, più alte, di analisi del contesto, del cotesto, delle aspettative dei parlanti ecc., che si sovrappongono alla percezione meramente uditiva e, per così dire, la guidano nella decodifica secondo schemi idiolinguistici. Ciò è vero non solo per il riconoscimento dei segmenti e delle sequenze di segmenti, ma è vero anche per i movimenti del *pitch*: questi, ferme restando le soglie fisiche di durata, altezza e

intensità, al di sotto delle quali non si ha percezione o non si ha discriminazione, possono venire trascurati nella decodifica linguistica anche quando siano psicoacusticamente discriminabili.

Questo circolo vizioso è oggi il vero problema teorico degli studi sull'intonazione e non è facile romperlo. Esso consiste, come abbiamo ricordato, nella compresenza e nella cooperazione dei meccanismi di percezione primitiva e di percezione basata su schemi, di cui parla Bregman (1990): le diverse posizioni degli studiosi dipendono in ultima analisi dal fatto che alcuni riducono drasticamente o annullano la percezione primitiva, mentre altri la prendono in considerazione, pur differendo nel dosaggio delle due componenti.

6. Alcune conclusioni

La linguistica può dare un contributo importante alla soluzione di questo nodo che non può essere considerato solo di pertinenza della psicoacustica. Non si deve dimenticare infatti che il segnale di cui ci si occupa è dotato, oltre che di proprietà prosodiche e pragmatiche, anche di proprietà semantiche, lessicali, sintattiche e morfologiche.

Il contributo della linguistica deve partire da alcune semplificazioni. Attualmente la principale difficoltà nello studio della prosodia dipende essenzialmente a) dalla sua intrinseca variabilità e b) da una interrelazione tra variazioni prosodiche e variazioni attitudinali del parlante che si dispongono ambedue lungo un *continuum* non facilmente discretizzabile (o comunque molto meno facilmente discretizzabile di quanto non sia il *continuum* segmentale sul piano fonologico, sul piano morfologico, sul piano sintattico e forse anche sul piano semantico, per il quale si dispone almeno di una unità operativa come la parola). Mettere in correlazione due variabili continue, relativamente poco note, e cercare di stabilire una gerarchia delle covariazioni è un'impresa disperata senza due operazioni preliminari ²⁵.

La prima è che è necessario assumere, come ipotesi di lavoro, che una delle due variabili possa essere considerata provvisoriamente la variabile indipendente. È ragionevole pensare che questa possa essere una qualche unità del parlato, come il turno nelle interazioni dialogiche, per la quale esistono, a partire dalla teoria degli atti linguistici di Austin e di Searle, le classificazioni elaborate dalla pragmatica: queste, per quanto controverse e, a volte, opinabili, si basano tuttavia, allo stato attuale, su un apparato più complesso e meno incerto di quelle elaborate dalla prosodia: il ruolo provvisorio di variabile indipendente può essere loro assegnato anche in base alla

considerazione che le unità pragmatiche possono essere classificate e descritte a prescindere dalla implementazione prosodica, ma il contrario non è vero.

La seconda è che ciascun insieme di enunciati raccolti sotto una determinata etichetta pragmatica deve essere ulteriormente analizzato per mezzo di tutti i sofisticati strumenti di analisi (semantica, sintattica, morfologica, funzionale) di cui dispone la linguistica: questa ulteriore analisi è necessaria perché la prosodia non può essere considerata solo in funzione della pragmatica ma deve essere considerata anche in funzione delle strutture linguistiche.

La paziente e graduale costruzione di queste griglie può fornire lo schema, relativamente robusto, rispetto al quale iniziare, su basi più certe, la classificazione delle regole prosodiche.

Questo programma di lavoro richiede ancora due corollari. Il primo è che ritengo una condizione ormai irrinunciabile che il materiale da analizzare sia costituito da enunciati naturali, estratti da raccolte sistematiche di parlato spontaneo (che nella grande maggioranza dei casi è parlato dialogico), con tutte le difficoltà che ciò può provocare: questo materiale verrebbe ad accompagnarsi a quello già disponibile sul parlato letto o recitato. Il secondo è che nei protocolli di analisi un posto preminente venga assegnato alle verifiche percettive condotte sui parlanti, intesi non come parlanti ideali senza corpo e senza storia, ma come parlanti concreti e definiti in base a caratteristiche esplicite. In questo i linguisti possono guardare ad altri specialisti che considerano importante che gli esperimenti siano espliciti e riproducibili e che i risultati siano verificabili. La convalida del ricevente è importante anche da un punto di vista teorico: nell'interscambio parlato la materia fonica diviene forma semiotica attraverso l'integrazione tra l'agire intenzionale del parlante e la convalida del ricevente. Senza quest'ultima si cadrebbe nel solipsismo.

Se gli schemi prosodici da utilizzare in questo confronto debbano essere rappresentati in forma di *pitch levels* o in forma di *pitch movements* è oggi oggetto di vivace discussione scientifica: il primo modello (la cui codifica formale più nota e diffusa è quella detta ToBI) ha dalla sua il prestigio indiscusso di buona parte della linguistica generativa più recente); il secondo ha dalla sua il prestigio della scuola olandese. Le opinioni sulle capacità esplicative dei due modelli possono essere diverse. Ma per evitare che questa disputa divenga una disputa teologica è necessario passare alle verifiche.

Indirizzo dell'autore: CIRASS - Università di Napoli Federico II, v. Porta di Massa 1, I-80133 Napoli - fealbano@unina.it

Note

* Ringrazio Pier Marco Bertinetto che ha letto questo articolo e mi ha suggerito numerosi miglioramenti.

¹ Naturalmente non è sempre così: il tema della fatica del capire è presente in molti lavori di De Mauro (p. es. 1994); in Pinker (1994), si leggono pagine molto efficaci sulla acustica dei suoni (155-163), sulla loro variabilità (173-175), sulle strategie di riconoscimento (175-180) e sul ruolo preminente assegnato all'udito nell'ontogenesi del linguaggio (256-57). Allo stesso modo si osserva che una parte della semiotica è giustamente interessata, *mutatis mutandis*, al ruolo del ricevente e della decodifica (p. es. Eco 1979). Diversa è la situazione quando si osservino lavori specialistici, come vedremo più avanti.

² Naturalmente ci sono eccezioni: per la situazione italiana, oltre a Uguzzoni (1990 a, b) si può vedere la rassegna in Albano Leoni (in stampa).

³ Questa storia sarebbe molto interessante e, per le vicende del Settecento e dell'Ottocento, è stata anche impostata (Gessinger 1994:485-631 e *passim*, Loi Corvetto 1992, 1995, 1998, Dovetto 1998, Albano Leoni & Dovetto 1997, Tani 2000: 104-105, 131-132). È ancora tutta da studiare, da questo punto di vista, la posizione di Paul, di Saussure, degli strutturalisti, della linguistica chomskiana e postchomskiana.

⁴ Lo stesso punto di vista è rappresentato ancora in Jakobson & Waugh (1979: *passim*), che costituisce una sorta di ricapitolazione del pensiero e degli studi di Jakobson sull'aspetto fonico delle lingue.

⁵ "the perception of speech is tightly linked to the feedback from the speaker's own articulatory movements" (Lieberman *et al.*, 1963: 4).

⁶ Bertil Malmberg (1967: 168) aveva visto con chiarezza i limiti di questa posizione: "Il est intéressant de constater qu'une nouvelle génération, élevée avec les résultats acoustiques modernes et trop jeune pour connaître l'histoire, plus ancienne, de la phonétique articulatoire et sa fallite avec Meyer, Russel etc., se sent maintenant tentée, devant cette même complexité des faits qui avait enlevé trente ans plus tôt à l'articulation son rôle comme base de classement primaire, de recourir à la physiologie et au sens musculaire pur expliquer des phénomènes qui, aux yeux de celui qui entreprend son analyse d'un point de vue phonologique, n'impliquent aucun problème". A queste osservazioni si potrebbe aggiungere che questa prospettiva articolatoria, almeno nella sua forma corrente, è problematica per almeno quattro motivi: il peso dato al ruolo del segmento configura un orientamento *bottom-up* molto rigido; il ruolo della prosodia, non riconducibile certo alle sensazioni articolatorie è azzerato; le esperienze uditive perinatali, che certamente precedono ogni capacità articolatoria, non vengono riconosciute; la presenza di deficit articolatori anche gravi non inficia la percezione linguistica. Perplessità esplicite sulla *Mothor Theory* sono anche in Jakobson & Waugh (1979:*passim*).

⁷ Il ruolo dell'udito è considerato in un recente libro italiano di psicolinguistica (Orsolini 2000) dedicato all'apprendimento della lingua nei bambini.

⁸ Naturalmente qui prescindo del tutto dai problemi e dalle controversie circa la rieducazione dei sordi (Pennisi 1994: 21-93), come anche da quelli sulla equipotenza di codici diversi dall'audioverbale.

⁹ Questa storia ontogenetica sembrerebbe riflettere, come si sa, quella filogenetica (Lieberman 1975: 153-165; Lieberman & Blumstein 1988: 205-213): la posizione della laringe nei primati, immediatamente al di sotto della lingua, corrisponde a quella del feto umano e in ambedue i tipi la cavità faringale è ridottissima. Ancora alla storia filogenetica appartiene il fatto che il diametro della laringe

umana rispetto a quello della trachea si è modificato a svantaggio della respirazione e a vantaggio della fonazione: il minor diametro della laringe favorisce il costituirsi di una pressione subglottidale, indispensabile per l'attivazione del meccanismo laringeo. Infine, non si può ignorare il fatto che mentre l'anatomia e la fisiologia dell'apparato uditivo umano sono in fondo molto simili a quelle di altri mammiferi, i rispettivi apparati fonatori sono invece molto diversi. Mi sembra che questi siano tutti indizi del fatto che nel processo di messa a punto degli strumenti per la comunicazione umana, l'interazione tra produzione e ricezione deve essere stata molto forte e che l'assegnazione all'udito di un ruolo subalterno non ha molto fondamento.

¹⁰ La mia incompetenza mi aiuta qui a tenermi alla larga dal problema, forse di impossibile soluzione, se il canale fonico-uditivo sia il canale necessario per lo sviluppo del linguaggio umano o se questo canale avrebbe potuto essere sostituito da altri (Cimatti 1998: 190-195).

¹¹ Si consideri, a titolo di esempio, che 20 db è la quantità di rumore di fondo presente di notte in una zona rurale.

¹² La capacità di discriminare in durata, anche molto importante, non figura nel campo uditivo, così come esso viene tradizionalmente rappresentato.

¹³ Il termine e il concetto, mutuati dagli studi sulla percezione visiva, indicano il fenomeno per cui due stimoli (visivo o uditivo) che esibiscono contenuti spettrali diversi vengono percepiti come uguali.

¹⁴ Prescindo qui tanto dai casi anomali quanto dai casi di voci educate, come quelle dei cantanti.

¹⁵ Sul concetto di 'pitch virtuale' cfr. Hess (1983: 73).

¹⁶ Questi diagrammi presentano il tempo in ascissa, la frequenza in ordinata e l'energia nel maggiore o minore annerimento del tracciato. Nei casi qui considerati la banda di frequenza va da 0 a 8000 Hz. Le linee orizzontali rappresentano un intervallo di 500 Hz.

¹⁷ "the learned schemas must obviously make use of primitive unlearned abilities (or else there would be nothing from which to build up the learned abilities)" (Bregman 1990: 602-603).

¹⁸ La bibliografia sull'argomento sta crescendo vertiginosamente. 'tHart e altri (1990: 1-6) danno una presentazione vivacemente efficace del problema; Bolinger (1989) è una introduzione insostituibile; Bertinetto (1981) e Bertinetto & Magno Caldognetto (1993), con bibliografia, forniscono il quadro della situazione italiana, per la quale una rassegna bibliografica aggiornata è in Albano Leoni (in stampa).

¹⁹ L'importanza della scrittura per la categorizzazione fonetica segmentale è intuiva mirabilmente da Leopardi (in Gensini 1988: 49-54).

²⁰ In una prospettiva semiologica potrà essere interessante osservare che in prosodia le due facce del segno sono ancora più profondamente integrate che in altri livelli (cfr. a questo proposito Lo Piparo 1991).

²¹ La parola *casa* può essere realizzata [ˈkaːsa], [ˈkaːza], [ˈhaːsa], [ˈkæːsa], [ˈkaːs] ecc. senza che questo alteri il suo contenuto e senza che ciò sia indizio di particolari intenzioni comunicative del parlante (se non, in qualche caso, di una scelta stilistica).

²² "the relation between intonational and attitudinal features may be one-to-many, possibly even many-to-many" ('t Hart *et al.* 1990: 111). Naturalmente questo vale per le variazioni pertinenti e intenzionali (anche se non è sempre facile distinguere tra queste e le variazioni casuali, in mancanza di una categorizzazione fine delle *attitudinal features*).

²³ "Ecchisietevòì, paladino di Francia?" viene sostituito da "Tàtta-tatàtài tàta-tàta-tatàta..." (Calvino 1959: 10).

²⁴ In italiano ad una affermazione come "ho mangiato una mela" si può risponde-

re “sì sì” con una intonazione particolare e si esprime in questo modo scetticismo, incredulità e il senso della risposta viene ad essere “non è vero, non ci credo”.

²⁵ Il programma che adombro in queste conclusioni nasce dalle riflessioni di un gruppo di lavoro sul parlato, attivo presso l'università di Napoli, e in particolare dalle discussioni con Renata Savy, Dalia Gamal e Claudia Crocco, che su questa strada stanno lavorando.

Summary

The paper deals with the role of the hearer in speech communication with special regard for intonation.

After a short presentation of the problem and of its recent history in linguistic studies (section 1), a description of the characteristics of hearing apparatus is presented and attention is paid to the synergy between speech production and speech perception and to the discussion about the role of primitive and inlearned auditory schemas (section 2-4).

In section 5 intonation is taken into account as major and powerful communication tool, whose importance and efficiency are universally recognised, but whose study and description appear still difficult because of some of its semantics and pragmatics properties and because of the lack of a notation tradition (as letters provide for segmentals strings).

Section 6 provides some conclusions and some suggestions for investigating intonation from a linguistic point of view.

Bibliografia

- AKMAJIAN, A., R.A. DEMERS, A.K. FARMER & R.M. HARNISH (1995⁴), *Linguistics. An Introduction to Language and Communication*, Cambridge (Mass.), The MIT Press; Ital. transl. il Mulino, 1996.
- ALBANO LEONI, Federico & Francesca M. DOVETTO, (1996), “From Maine de Biran to the ‘Motor Theory’: A Note in the History of Phonetics”, *Historiographia Linguistica* 23,3:347-364.
- ALBANO LEONI, F. (in stampa), “Fonetica e fonologia”, in Lavinio (in stampa).
- BERTINETTO, P.M. & E. MAGNO CALDOGNETTO (1993), “Ritmo e intonazione”, in Sobrero (1993: 141-192).
- BERTINETTO, P.M. (1981), *Le strutture prosodiche della lingua italiana. Accento, quantità, sillaba, giuntura, fondamenti metrici*, Firenze, Accademia della Crusca.
- BOLINGER, D. (1989), *Intonation and its uses*, London, Arnold.
- BREGMAN, A.S. (1990), *Auditory Scene Analysis. The Perceptual Organization of Sound*, Cambridge (Ma) / London, The MIT Press (paperback 1994).
- CALVINO, I. (1959³), *Il cavaliere inesistente*, Torino, Einaudi.
- CASULA, M.S. *et al.*, eds. (1998), *Linguistica e dialettologia. Studi in memoria di Luigi Rosiello*, Roma, Carocci
- CHOMSKY, N. & M. HALLE (1968), *The Sound Pattern of English*, New York, Harper and Row.

- CIMATTI, F. (1998), *Mente e linguaggio negli animali. Introduzione alla zoosemiotica cognitiva*, Roma, Carocci.
- CORRADI FIUMARA, G. (1990), *The Other Side of Language. A Philosophy of Listening*, London / New York, Routledge.
- CRUTTENDEN, A. (1986), *Intonation*, Cambridge, Cambridge University Press.
- DE MAURO, T. (1994), "Intelligenti pauca", in De Mauro (1994: 63-74).
- DE MAURO, T. (1994), *Capire le parole*, Roma / Bari, Laterza
- DELGUTTE, B. (1997), "Auditory Neural Processing of Speech", in Hardcastle & Laver (1997:507-538).
- DOVETTO, F.M. (1998), "Produzione e ricezione del linguaggio negli studi italiani della seconda metà del Settecento", *Lingua e Stile* 33:231-266.
- ECO, U. (1979), *Lector in fabula. La cooperazione interpretativa nei testi narrativi*, Milano, Bompiani.
- FORMIGARI, L. & D. GAMBARARA, eds. (1995), *Historical Roots of Linguistic Theories*, Amsterdam / Philadelphia, Benjamins.
- GENSINI, S. (1993[1994]), "Epicureismo e naturalismo nella filosofia del linguaggio fra umanesimo e illuminismo: prime linee di ricerca", *Annali della Facoltà di Magistero dell'Università di Cagliari*, n.s., 16:55-119.
- GENSINI, S., ed. (1988), Giacomo Leopardi, *La varietà delle lingue*, Firenze, La Nuova Italia.
- GERNSBACHER, M.A., ed. (1994), *Handbook of Psycholinguistics*, Academic Press, San Diego-London.
- GESSINGER, J. (1994), *Auge und Ohr. Studien zur Erforschung der Sprache am Menschen 1700-1850*, Berlin / New York, de Gruyter.
- GOLDSMITH, J.A., ed. (1995), *The Handbook of Phonological Theory*, Cambridge (Ma) / Oxford, Blackwell.
- GOODMAN, J.C. & H.C. NUSBAUM, eds. (1994), *The Development of Speech Perception*, Cambridge (Ma) / London, The MIT Press.
- HARDCASTLE, W.J. & J. LAVER, eds. (1997), *The Handbook of Phonetic Sciences*, Cambridge (Ma) / Oxford, Blackwell.
- HAWKINS, S. (1999), "Auditory Capacities and Phonological Development: Animal, Baby and Foreign Listeners", in Pickett (1999:183-198).
- HESS, W. (1983), *Pitch Determination of Speech Signals. Algorithms and Devices*, Berlin / Heidelberg / New York / Tokyo, Springer.
- JAKOBSON, R. & L.R. WAUGH (1979), *The Sound Shape of Language*, Ital. transl. il Saggiatore, 1984.
- JAKOBSON, R. & M. HALLE (1956), *Fundamentals of Language*, The Hague, Mouton.
- KENSTOWICZ, M. (1994), *Phonology in Generative Grammar*, Cambridge (Ma) / Oxford, Blackwell.
- LASPIA, P. (1997), *L'articolazione linguistica. Origini biologiche di una metafora*, Roma, Nuova Italia Scientifica..
- LAVER, J. (1991), *The Gift of Speech. Papers in the Analysis of Speech and Voice*, Edinburgh, Edinburgh University Press.
- LAVINIO, C., ed. (in stampa), *La linguistica italiana alle soglie del 2000 (1987-1998)*, Roma, Bulzoni.
- LENNEBERG, E.H. (1967), *Biological Foundations of Language*. With Appendices by Noam Chomsky and Otto Marx, New York, Wiley (repr. Malabar, Fla, Krieger, 1984)

- LIBERMAN, A.M. & I.G. MATTINGLY (1985), "The Motor Theory of Speech Perception Revised, *Cognition*, 21:1-36.
- LIBERMAN, A.M. (1996), *Speech: A Special Code*, Cambridge (Ma) / London, The Mit Press.
- LIBERMAN, A.M., F. S. COOPER, K. S. HARRIS & P.F. MACNEILAGE (1963), "A Motor Theory of Speech Perception", *Proceedings of the Speech Communication Seminar 2*: 1-12, Stockholm, Speech Transmission Laboratory, Royal Institute of Technology.
- LIEBERMAN, P. & S.E. BLUMSTEIN (1988), *Speech Physiology, Speech Perception and Acoustic Phonetics*, Cambridge/New York, Cambridge University Press.
- LIEBERMAN, P. (1975), *On the Origin of Language*, New York, Macmillan (rist. 1987), Ital. transl. Boringhieri, 1980.
- LO PIPARO, F. (1991), "Le signe linguistique est-il à deux faces? Saussure et la topologie", *CFS*, 45: 213-221.
- LOI CORVETTO, I. (1992), "Significati e usi di 'lettere' e 'son' nell'Encyclopédie", *Lingua e Stile* 27,3: 363-377.
- LOI CORVETTO, I. (1995), "Il processo della percezione dei suoni nell'Encyclopédie", *Lingua e Stile* 30,1:83-93.
- LOI CORVETTO, I. (1998), "Il processo uditivo in Étienne Bonnot de Condillac", in Casula *et al.* (1998:75-94).
- MALMBERG, B. (1967), "Réflexions sur les traits distinctifs et le classement des phonèmes, in *To Honor Roman Jakobson* (1967:1247-1251), The Hague, Mouton.
- MARCHESE, M.P., ed. (1995), Ferdinand de Saussure, *Phonétique. Il manoscritto di Harvard Houghton Library bMS Fr 266 (8)*, Padova, Unipress.
- MCQUEEN, J.M. & A. CUTLER (1997), "Cognitive Processes in Speech Perception", in Hardcastle & Laver (1997: 566-585).
- MEHLER, J. & E. DUPOUX (1992), *Appena nato. Che cosa vede, sente, capisce un bambino sin dai primi giorni di vita*, Milano, Mondadori, French transl. Odile Jacob, 1990.
- MILLER, J.L. & P.D. EIMAS (1994), "Observations on Speech Perception, Its Development, and the Search for a Mechanism", in Goodman & Nusbaum (1994: 37-55).
- MOORE, B. C. J., ed. (1986), *Frequency Selectivity in Hearing*, London, Academic Press
- MOORE, B.C.J. (1988³), *An Introduction to the Psychology of Hearing*, London, Academic Press.
- MOORE, B.C.J. (1997), "Aspects of Auditory Processing Related to Speech Perception", in Hardcastle & Laver (1997:539-565).
- NESPOR, M. (1993), *Fonologia*, Bologna, il Mulino.
- NUSMAN, H.C. & J.C. GOODMAN (1994), "Learning to Hear Speech as Spoken Language", in Goodman & Nusbaum (1994:299-338).
- ORSOLINI, M., ed. (2000), *Il suono delle parole. Percezione e conoscenza del linguaggio nei bambini*, Firenze, La Nuova Italia.
- PENNISI, A. (1994), *Le lingue mute*, Roma, NIS.
- PICKETT, J.M., ed. (1999), *The Acoustics of Speech Communication. Fundamentals, Speech perception Theory, and Technology*, Needham Heights (Ma), Allyn & Bacon.

- PINKER, S. (1994), *The Language Instinct*, Ital. transl. Mondadori, 1997.
- ROSEN, S. & A. FOURCIN (1986), "Frequency selectivity and the perception of speech", in Moore (1986:373-487).
- RYALLS, J. (1996), *A Basic Introduction to Speech Perception*, San Diego / London, Singular Publishing Group.
- SAUSSURE, F. de (1922 [1916]), *Cours de linguistique générale*, Paris, Payot, Ital. transl. Laterza, 1968.
- SIMONE, R. (1992), "Il corpo del linguaggio. Il paradigma dell'arbitrarietà e il paradigma della sostanza", in ID. (1992: 37-59).
- SIMONE, R. (1995⁶), *Fondamenti di linguistica*, Roma / Bari, Laterza.
- SIMONE, R. (1995a), "The language User in Saussure (and after)", in Formigari & Gambarara, (1995:233-249).
- SIMONE, R., ed. (1992), *Il sogno di Saussure*, Roma / Bari, Laterza
- SOBRERO, A.M., ed. (1993), *Introduzione all'italiano contemporaneo*, vol. I, *Le strutture*, Bari / Roma, Laterza.
- T HART, J., R. COLLIER & A. COHEN (1990), *A perceptual study of intonation. An experimental-phonetic approach to speech melody*, Cambridge, CUP.
- TANI, I. (2000), *L'albero della mente. Sensi, pensiero, linguaggio in Herder*, Roma, Carocci.
- UGUZZONI, A. (1990 a), "Uno sguardo al fenomeno della percezione categorica, *Lingua e stile* 25,1: 119-133.
- UGUZZONI, A. (1990 b), "Dagli studi sulla percezione fonica: dati, questioni, proposte", *Studi Italiani di Linguistica Teorica ed Applicata*, 19: 3-21.
- ZWICKER, E. (1961), "Subdivision of the Audible Frequency Range into Critical Bands (Frequenzgruppen)", *JASA* 33, 2: 248.