

MORRIS HALLE, *Fonetica*

Estratto da:

Enciclopedia, VI: *Famiglia-Ideologia*, Einaudi, Torino 1979.

- Foucault, M.
1961 *Histoire de la folie à l'âge classique, folie et déraison*, Pion, Paris (trad. it. Rizzoli, Milano 1963).
- Horkheimer, M., e Adorno, Th. W.
1947 *Dialektik der Aufklärung. Philosophische Fragmente*, Querido, Amsterdam (trad. it. Einaudi, Torino 1976²).
- Malthus, Th. R.
1798 *An Essay on the Principle of Population, as it Affects the Future Improvement of Society, with Remarks on the Speculations of Mr. Godwin, M. Condorcet and Others Writers*, Johnson, London (trad. it. Einaudi, Torino 1977).

La storia della follia è storia di un giudizio, quindi è storia dei valori, delle credenze dei sistemi di potere (cfr. **potere/autorità**) su cui si fonda il **gruppo** sociale. Le definizioni date della follia si scrivono quindi nel mondo magico (cfr. **magia**), religioso (cfr. **religione, divino**), rituale (cfr. **mito/rito**) di cui è impregnata la cultura (cfr. **cultura/culture**, e anche **natura/cultura**), che la ingloba nella totalità dei fenomeni umani naturali/sovranaturali.

La coppia follia/delirio scaturisce da un **conflitto** con la **ragione**, riconosciuto come contraddittorio e controproducente. Se la definizione di follia nasce da particolari forme di razionale (che riconoscono come irrazionale ciò che non è corretto in rapporto agli schemi dominanti; cfr. **razionale/irrazionale**), quando si parla di follia si tratterà non tanto di realtà ad essa riferibili in assoluto, quanto piuttosto di immagini della realtà, di forme di sapere per le quali determinati comportamenti (cfr. **comportamento e condizionamento**) non funzionali a un regime di realtà sono considerati anormali (cfr. **normale/anormale, censura, puro/impuro**).

I meccanismi di **esclusione/integrazione** che la definizione di follia sempre sottintende sono da riferirsi tanto ai sistemi di produzione (cfr. **produzione/distribuzione**) quanto ad una collusione, non di rado necessaria, fra **clinica** e ragione, fra cura e normalizzazione (cfr. **cura/normalizzazione, salute/malattia, farmaco/droga**) fondata su una concezione del **corpo** organico per la quale esso viene identificato, complice primario la **medicina/medicalizzazione**, nello stretto limite fisco del *côlux* (cfr. **soma/psiche**). Da un lato occorre dunque sottolineare che il sistema produttivo porta all'appropriazione della soggettività dell'uomo (cfr. **repressione, libertà**), dall'altro che l'uomo di cui si appropria il sistema produttivo risulta come sdoppiato fra corpo organico e corpo economico.

Se è vero che può essere la stessa **sistemática e classificazione** dei comportamenti a continuare a fare accettare il **mondo** com'è (la follia è invece un'esperienza tragica del mondo), vorrà dire che le modalità discriminanti (cfr. **classi**) che fondano la «garanzia sociale» della follia stanno in un sistema ben definito d'**interpretazione**, che la stessa storia della psicanalisi mostra essere prodotto dalla «cultura borghese» (cfr. **borghesi/borghesia**).

La follia e il delirio saranno allora da ricondurre alla totalità dei bisogni (cfr. **bisogno**) e desideri (cfr. **desiderio**) dell'uomo, alla qualità della sua **vita** (cfr. **marginalità, povertà**), alla inadeguatezza della sua fruizione dello spazio (cfr. **spazio sociale**), all'intolleranza della diversità che una società altamente industrializzata sa produrre. Ma l'interpretazione della follia non può che essere iscritta nella logica del **linguaggio** (cfr. **parola**) di chi la ascolta e la giudica e giudicandola determina i modi in cui deve esprimersi.

1. *Introduzione.*

La fonetica è la scienza dei suoni del linguaggio. Tradizionalmente, essa si è occupata del modo in cui tali suoni vengono prodotti dal meccanismo umano di fonazione (fonetica articolatoria), delle loro proprietà acustiche (fonetica acustica) e del modo in cui i suoni del linguaggio sono percepiti (psicoacustica). Per gran parte, l'oggetto della fonetica costituisce materia d'analisi anche per discipline altamente specialistiche, quali l'anatomia, l'acustica e la psicologia. Tuttavia, a dispetto di questo sconfinamento, la fonetica non è rimasta inglobata in nessuno di questi tre ambiti di ricerca, ma ha sempre mantenuto la propria autonomia disciplinare. La ragione di questa indipendenza sta tipicamente nel fatto che essa implica un aspetto che, di per sé, non emerge in modo specifico nelle tre discipline sopra menzionate: il fonetista deve infatti, in ogni momento, tener presente che egli non ha a che fare con suoni puri e semplici, ma piuttosto con suoni del linguaggio, ossia con suoni prodotti e percepiti da particolari organismi (gli esseri umani), i quali hanno accesso a uno speciale ambito di conoscenza che qui sarà detto 'conoscenza della lingua'. I suoni che non implicano in modo determinante questa conoscenza non riguardano direttamente la fonetica, anche quando essi, dal punto di vista fisico, non sono affatto distinguibili dai suoni del linguaggio, non più di quanto il suono [f] si distingua dal rumore che si produce soffiando su una candela. In effetti, la nozione stessa di suono linguistico, che costituisce l'entità fondamentale della fonetica, non è un dato fisico ma un concetto linguistico. I ruoli prodotti da animali o da macchine, come il mugugno delle vacche, i richiami degli uccelli, il frastuono di un treno in movimento, o il rimbombo di un fucile, possono suonare alle nostre orecchie come sequenze di suoni verbali. Siamo noi peraltro che — grazie alla nostra conoscenza della lingua — analizziamo in tal modo questi eventi acusticamente compatti. Una vacca che mugugisce non produce una sequenza di suoni linguistici più di quanto il vento, sospingendo le nuvole nel cielo, non crei vedute di montagne, volti o forme di mostri giganteschi. Un buon esempio circa il ruolo fondamentale svolto dalla conoscenza della lingua nella percezione del discorso verbale ci è dato dalla percezione delle parole. I parlanti privi di particolari conoscenze linguistiche, nel corso di una conversazione, sono mentalmente convinti di produrre e udire delle parole. Ma le parole non sono soltanto delle speciali sequenze di eventi acustici (suoni), bensì delle entità linguistiche, cioè dei fatti che sono noti unicamente a coloro che possiedono la padronanza di una particolare lingua, e non a coloro che ne sono privi. A chiarimento di ciò, si consideri l' enunciato seguente:

It's necessary to have { *a name*
an aim

Così come viene di norma pronunciato, quest'enunciato è ambiguo. Esso può essere inteso come una raccomandazione ad avere o un nome o uno scopo. La ragione di tale ambiguità sta nel fatto che, di solito, i confini di parola — quelle entità che appaiono come spazi bianchi nel nostro sistema di scrittura — non sono segnalati acusticamente quando parliamo. Di norma non si fanno pause tra le parole: ciascuno può facilmente convincersene leggendo ad alta voce questa stessa frase, sforzandosi nel contempo di eseguire una pausa dopo ciascuna parola. L'assenza di appropriati segnali acustici per i confini di parola nel linguaggio parlato non impedisce ai soggetti umani di percepire normalmente un dato enunciato come una successione di parole: a condizione, beninteso, che essi conoscano la lingua in cui l'enunciato viene realizzato. Ciò significa che il parlante comune percepirà, alla fine di questa frase, o *name* o *aim*, mentre solo soggetti particolarmente avvertiti (magari persone con qualche pratica di fonetica) reagiranno osservando che l'enunciato è ambiguo. (Si pensi ancora alle possibili ambiguità che potrebbero nascere, in italiano, di fronte a esempi come *preziosi orari* / *preziosi o rari*, *legare* / *le gare*, ecc.). Ora, è ovvio che una persona ignara d'inglese non potrebbe presentare nessuna di queste reazioni. Tutto ciò che tale persona potrebbe riferire di avere udito è la sequenza fonica [ɪns'ɛsəri'jə'leɪtərənçɪm] che peraltro conterrebbe probabilmente molti più dettagli di quanti sarebbe capace di udire la maggior parte dei soggetti che ignorano la lingua.

Fatti di questo tipo sono altrettanti esempi della funzione cruciale che assume nella percezione ciò che gli psicologi hanno chiamato *mental set* 'atteggiamento mentale' (in tedesco *Einstellung*). Quando si percepiscono certe combinazioni di linee su un foglio di carta come tridimensionali, o quando si vedono facce, montagne ed esseri mitologici nelle formazioni di nuvole nel cielo, ovvero si giudicano le dimensioni di oggetti familiari in modo diverso da quanto si fa per oggetti che non sono familiari, la reazione si spiega solo parzialmente in base al segnale fisico che raggiunge gli occhi. Ciò che si percepisce rappresenta, in parte, un contributo indipendente della mente; e ciò è vero non solo per una classe ristretta di fenomeni percettivi, come quelli descritti sopra, ma per tutti i fenomeni percettivi. L'inglese possiede un termine speciale — *illusion* 'illusione' — per caratterizzare certe situazioni in cui il ruolo di colui che percepisce risulta particolarmente evidente. Se tuttavia si considerano più dappresso i fenomeni percettivi, ci si rende ben presto conto che quel termine si applica a ciascuno di essi: ogni percezione è, in un certo senso, un'illusione. E ciò è specialmente vero nel caso della percezione del linguaggio, poiché, come si è appena visto, la percezione delle parole nel fluire del discorso implica in modo determinante la conoscenza della lingua da parte dell'ascoltatore. È proprio il ruolo centrale che la conoscenza della lingua svolge nella comprensione dei fenomeni fonetici ad istituire la fonetica come campo di ricerca autonomo, separato dall'acustica, dall'anatomia o dalla psicologia.

2. *Aspetti anatomici del linguaggio.*

È appena necessario dire che se s'insiste tanto sull'importanza della conoscenza posseduta dai parlanti e dagli ascoltatori, non s'intende peraltro minimizzare l'importanza dello stimolo fisico. Il suono è essenziale per la comprensione del linguaggio, e il modo in cui esso è prodotto dal meccanismo fonatorio e percepito dal sistema auditivo apre importanti prospettive su certe proprietà fondamentali del linguaggio. Occorre quindi passare a un esame dell'apparato anatomico utilizzato normalmente dagli uomini per parlare.

L'apparato anatomico da cui dipende la produzione del linguaggio fa parte della zona superiore del condotto respiratorio e digestivo dell'uomo (fig. 1). Il linguaggio viene di norma prodotto tramite un flusso d'aria che procede dai polmoni del locutore verso l'esterno, anche se non è impossibile produrre dei suoni perfettamente intelligibili invertendo la direzione di tale flusso d'aria. In casi normali, comunque, l'aria viene compressa nei polmoni e finisce verso l'esterno attraverso i due bronchi sino alla trachea. La trachea, nella parte superiore, termina in una complessa struttura di legamenti, muscoli e cartilagini — la laringe — la cui funzione fondamentale è quella di una valvola che chiude ermeticamente i polmoni dall'esterno. La parte inferiore di questa valvola è costituita da una coppia di protuberanze orizzontali separate da una fessura. Le protuberanze sono dette corde vocali, e la fessura tra di esse è nota come glottide. Le corde vocali controllano il grado di costrizione della glottide, che può variare dalla completa occlusione (come quando si trattiene il fiato cercando di spostare un ostacolo) sino alla massima apertura (come

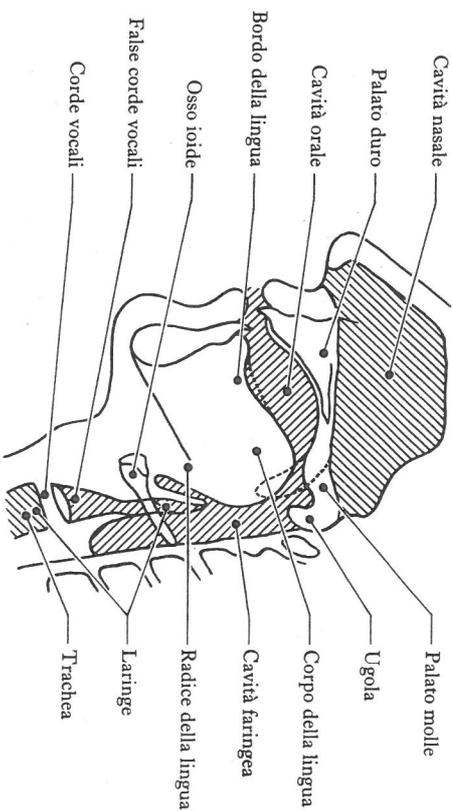


Figura 1.
Il condotto vocale umano.

nella respirazione profonda). Non sono questi estremi, comunque, a rivestire qui particolare interesse, bensì i gradi intermedi di costrizione della glottide, che si osservano, da una parte, quando si bisbiglia, e, dall'altra, quando si pronunziano vocali o suoni di tipo vocalico in modo normale, a piena voce. Nel bisbiglio, la glottide è ristretta in misura sufficiente a produrre una turbolenza perfettamente udibile. Nella normale pronunzia delle vocali e dei suoni di tipo vocalico la glottide viene invece ristretta in grado tale, e il flusso d'aria è regolato in maniera tale da mettere in vibrazione le corde vocali, che si aprono e chiudono molte volte al secondo, convertendo così il flusso espiratorio in piccoli sbuffi d'aria. Gli studiosi di fonetica chiamano «sonori» i suoni prodotti con questo tipo di vibrazione delle corde vocali. Ponendo la punta delle dita sulla laringe mentre si pronunziano vocali o altri suoni sonori si possono facilmente sentire queste vibrazioni delle corde vocali.

La rapidità di vibrazione delle corde vocali determina il livello melodico (comunemente detto 'altezza'; ingl. *pitch*) della voce: se cresce il numero delle vibrazioni, il livello melodico si innalza; se diminuisce, il livello cala. Ogni individuo presenta nel parlare una propria tipica gamma melodica: in media, essa è più bassa nei maschi adulti che nelle donne o nei bambini. Quando invece si passa al bisbiglio, o si pronunziano suoni *sordi* come [s, f, p, k], le corde vocali cessano di vibrare; e questo generalmente accade perché, divaricandosi, esse danno luogo a un'apertura glottidale di considerevole ampiezza.

L'uomo non è soltanto in grado di variare l'apertura della glottide, ma anche la posizione dell'intera laringe. Essa può infatti essere alzata o abbassata, variando così convenientemente il volume della cavità sopralaringale: quest'ultima si riduce quando la laringe viene sollevata, e aumenta quando essa viene abbassata. Inoltre, il sollevamento o l'abbassamento della glottide ha effetti concomitanti sul grado di tensione delle corde vocali: una laringe abbassata tende ad allentare le corde vocali, mentre una laringe sollevata tende ad irrigidirle. Entrambe queste variazioni (ossia nel volume della cavità e nell'elasticità delle corde vocali) vengono utilizzate nel parlare.

Lo spazio al di sopra della laringe consiste di tre cavità separate: la faringe, la bocca (o cavità orale) e le cavità nasali. La faringe è un tubo verticale, lungo circa dodici centimetri, provvisto di tre aperture: la prima, all'estremità inferiore, dà nella laringe; la seconda, a metà, dà nella cavità orale; la terza, all'estremità superiore, dà nelle cavità nasali. La parte nasale della faringe può essere separata dal resto della stessa ritraendo e sollevando il velo palatino, o parte molle del palato, in modo da portarlo a contatto con la parete posteriore della laringe. La forma della porzione mediana della faringe è controllata dalla lingua. A questo proposito è necessario precisare che esistono due distinti meccanismi, il primo dei quali controlla la posizione della radice della lingua, mentre il secondo controlla la posizione del corpo della lingua. La radice della lingua può essere spinta in avanti, come nella pronunzia delle vocali [r] o [ʀ], o può essere ritratta verso la parte posteriore della faringe, come nella pronunzia della vocale [ɹ] o delle consonanti faringalizzate dell'arabo.

Il corpo della lingua controlla non solo la forma della faringe ma anche quella della cavità orale. Il corpo della lingua può infatti essere mosso in avanti, come avviene nella pronuncia delle vocali [r] ed [l], o all'indietro verso la parte della faringe, come in [u] o [ɔ]. Differenze analoghe nella posizione orizzontale del corpo della lingua intervengono nella pronuncia delle consonanti; così, per esempio, il corpo della lingua è spinto in avanti quando si pronuncia la consonante iniziale nella parola *chiamo*, e all'indietro quando si pronuncia la consonante iniziale della parola *cane*. Oltre alla mobilità orizzontale, il corpo della lingua presenta anche una mobilità verticale, in su e in giù. Lo si può sollevare sino a toccare il palato, come nelle consonanti iniziali delle due parole appena citate, *chiamo* e *cane*, ma lo si può anche tirare alquanto verso il basso, come quando si pronuncia la vocale [a] o quando si cerca di mostrare le tonsille al dottore.

Il bordo (o lamina) della lingua viene usato per occludere o restringere la cavità orale in punti diversi. Per esempio, quando si pronuncia la consonante iniziale di *scato*, la cavità orale viene ristretta in una posizione più avanzata rispetto a quanto accade nella pronuncia della consonante iniziale di *giorno*. Il restringimento della cavità orale prodotto con il bordo della lingua può anche essere di entità più moderata, come per esempio avviene nella pronuncia della vocale contenuta nella parola inglese *bird*. Il bordo della lingua può essere sollevato in diversi modi. Da una parte, il sollevamento può avvenire come nella pronuncia della vocale di *bird*, in cui solo la punta della lingua è sollevata verso il palato. Ciò è in contrasto col modo in cui tale sollevamento si realizza nell'esecuzione della consonante iniziale di *giorno*, in cui l'intero bordo si muove verso il palato. Inoltre, è possibile controllare anche i soli margini laterali della lingua: quando si pronunziano suoni del tipo [l], i margini laterali sono abbassati e l'aria sfugge al di sotto dei canini e dei molari anziché al di sotto degli incisivi, come avviene invece per la maggior parte degli altri suoni.

Infine, è possibile sollevare il bordo della lingua in modo tale da farlo vibrare quando l'aria vi soffi contro. Ciò accade quando si producono certi tipi di suoni [r]. Queste r, dette appunto 'vibrate', sono un tratto dell'italiano letterario standard.

Il condotto vocale presenta due aperture verso l'esterno: le labbra e le narici. Le dimensioni e la forma dell'apertura labiale possono essere variate, e queste variazioni producono conseguenze acustiche significative. Se si prendono le labbra, di fatto si aumenta la lunghezza totale della cavità orale, ed è noto che la lunghezza di una cavità influisce direttamente sulle sue proprietà di risonanza. Anche le dimensioni dell'apertura influiscono sulle proprietà di risonanza della cavità (cfr. § 3). Ma è chiaro che l'apertura posta alle narici non può essere controllata dal parlante. Solo il sollevamento o l'abbassamento del velo palatino, o palato molle, è sotto il suo diretto controllo. Quando il velo è sollevato, tutta l'aria fluisce dai polmoni lungo la bocca senza minimamente passare attraverso la faringe nasale e le cavità nasali, le quali rimangono pertanto acusticamente inerti. D'altra parte, quando il velo pala-

tino è abbassato, il flusso d'aria proveniente dai polmoni viene diviso in due parti: una passa attraverso la bocca, l'altra attraverso la cavità nasale, nella quale si producono in tal modo dei particolari effetti acustici, che svolgono una specifica funzione nella fonazione.

Al centro del margine inferiore del velo palatino è posto un elemento a forma di coda, chiamato *ugola*; il quale, in certe lingue, viene messo in vibrazione nella produzione di certi suoni del tipo [r]: per esempio, quelli che si ritrovano nel francese *létérario*.

Con ciò si può dire completo il panorama delle strutture anatomiche che intervengono attivamente nella fonazione. Si noterà che la lingua è di gran lunga l'organo più importante, poiché essa controlla la forma delle principali cavità di risonanza, come la faringe e la bocca. Non c'è quindi da sorprendersi se in molte lingue una stessa parola ('lingua') indica tanto l'organo anatomico quanto quel particolare sistema grammaticale e lessicale mediante il quale gli uomini comunicano fra loro. Vale inoltre la pena di notare che i movimenti della mascella inferiore (o mandibola), così evidenti quando si osserva parlare una persona, non paiono avere una funzione centrale nella fonazione; non sembra esservi alcun suono linguistico nella cui produzione la posizione della mascella inferiore sia determinante. La ragione più probabile di ciò sta nel fatto che i movimenti della mandibola non possono venire controllati così rapidamente come quelli di certe altre strutture anatomiche che intervengono nella fonazione.

3. Alcune informazioni sull'acustica del linguaggio.

Poiché il linguaggio umano si manifesta come una forma sonora, per una sua piena comprensione è necessario comprendere alcune delle sue caratteristiche acustiche. Chiunque può facilmente osservare che un contenitore qualsiasi (per esempio una bottiglia, o una scatola, o un tubo), quando viene percosso, emette un suono. La natura del suono dipenderà principalmente da due fattori: la forma geometrica del contenitore (o cavità) e il modo in cui esso viene percosso o - più in generale - il modo in cui l'aria contenuta nella cavità viene messa in movimento, o eccitata. Per semplificare, si consideri dapprima il caso in cui una cavità (ad esempio un bicchiere) viene eccitata con un singolo colpo su di una parete laterale. Il suono che si produrrà in tal modo dipenderà soltanto dalla forma della cavità. Un bicchiere basso produrrà pertanto un suono più acuto di un bicchiere alto, e una simile differenza emergerà anche tra i suoni prodotti da bottiglie piccole di contro a quelle di bottiglie grandi. Da un punto di vista acustico, i suoni così prodotti contengono tutte le risonanze, o modi naturali di vibrazione, della cavità; e, come si è già detto, tali risonanze sono del tutto determinate dalla forma geometrica della cavità stessa.

Come esempio si consideri un tubo o cilindro di altezza a , chiuso ad un'estremità e aperto dall'altra. È stato dimostrato che una tale cavità presenta

un numero infinito di risonanze, o possibilità vibratorie naturali, e che queste sono date dalla formula

$$(1) \quad F = \frac{c}{4a} \times n$$

dove F è la frequenza delle singole risonanze, c = la velocità di propagazione del suono, a = l'altezza del tubo, n = un numero intero dispari, ossia, 1, 3, 5, ecc. Sia ora $c = 34\,500$ cm/sec (ossia la velocità di propagazione del suono al livello del mare, a una temperatura di 20°C), $a = 17,25$ cm (valore vicino alla lunghezza media del condotto vocale maschile misurato dalla glottide alle labbra) e $n = 1$: si otterrà così

$$(2) \quad F = \frac{34\,500}{4 \times 17,25} \times 1 = 500 \text{ Hz.}$$

che indica la più bassa risonanza di un tubo che potrebbe funzionare, in via di prima approssimazione, come simulazione del condotto vocale umano. Le risonanze più acute ottenibili per $n = 3, 5$, ecc. danno frequenze di 1500 Hz, 2500 Hz, ecc.

È possibile controllare la lingua e le labbra in modo tale che la sezione trasversale della cavità risulti abbastanza uniforme, ossia in modo che la forma della cavità stessa assomigli a quella di un tubo chiuso ad un'estremità. (Il fatto che il condotto vocale umano faccia una curva tra la faringe e la bocca determina, nella gamma di frequenze che qui interessa maggiormente, effetti del tutto secondari, che possono quindi essere trascurati). Il modo migliore per fare assumere al condotto vocale la forma di un tubo uniforme consiste nel pronunziare la vocale [o] senza protendere al tempo stesso le labbra, ossia facendo loro assumere la configurazione che esse presentano quando si pronunzia [e]. Il suono così prodotto, che è per inciso perfettamente documentato in alcune lingue (cfr., per esempio, l'ultima vocale nella parola inglese *China*), presenta risonanze a circa 500, 1500, 2500 Hz.

Un altro fatto acustico rilevante ai fini dello studio della fonazione è che anche un tubo che abbia entrambe le estremità chiuse produce delle risonanze. In particolare, un tubo di altezza a , in tali condizioni, risuonerà alle frequenze

$$(3) \quad F = \frac{c}{4a} \times n$$

dove F , c , a vanno interpretati allo stesso modo che nella (1), mentre n è zero o un numero intero pari: $n = 0, 2, 4, 6$, ecc. È facile stabilire in base a ciò che, se si assegnano gli stessi valori della (1) a c e a , il condotto vocale, una volta chiuso all'altezza delle labbra, risuonerà alle frequenze di 0, 1000, 2000, 3000 Hz, ecc. Non è certo difficile chiudere il condotto vocale all'altezza delle labbra — lo si fa quando si pronunzia un suono labiale come [b, p, m] —, e così, quando si pronunzia una vocale seguita da uno di questi suoni labiali

— per esempio, la sillaba [am] —, le risonanze del condotto vocale di fatto si spostano gradualmente più in basso, verso i valori 0, 1000, 2000 Hz, ecc., man mano che il condotto vocale stesso viene chiuso all'altezza delle labbra nella transizione dalla vocale alla consonante. È necessario notare, a questo punto, una conseguenza importante della (3). Mentre le risonanze di un tubo dipendono soltanto dalla sua lunghezza, e ciò traspare dall'inclusione di a nelle equazioni (1) e (3), la risonanza più bassa di un tubo chiuso ad entrambe le estremità è del tutto indipendente dalla sua lunghezza; essa è sempre a 0 Hz. Questo fatto avrà una certa rilevanza nell'esposizione che segue.

Si è osservato sopra che il suono emesso da una cavità è determinato da due fattori: la forma geometrica della cavità e il modo in cui essa viene eccitata. Fino a questo punto della discussione si è preso in considerazione soltanto il primo di essi. Quando si producono delle vocali, la fonte dell'eccitazione del condotto vocale è situata all'altezza della glottide, che rappresenta l'estremità chiusa. Nelle vocali normali, prodotte con piena sonorizzazione, l'eccitazione consiste di piccoli sbuffi d'aria emessi dalla glottide alla velocità di circa 100-120 sbuffi al secondo nei maschi adulti, e a una velocità doppia nelle femmine. L'effetto di ciascuno sbuffo d'aria è assai simile a quello di una singola percossa: esso eccita i modi naturali di vibrazione, o risonanze, della cavità; le quali, come si ricorderà, sono interamente determinate dalla forma geometrica della cavità stessa. La velocità degli sbuffi d'aria emessi dalla glottide può essere comodamente variata. Sono queste variazioni nella velocità di eccitazione che producono le inflessioni della voce. Queste ultime vengono utilizzate in molte lingue per particolari scopi comunicativi; per esempio, in italiano e in molte altre lingue le domande sono segnalate da una particolare intonazione ascendente, mentre le proposizioni dichiarative sono segnalate da una caduta del livello melodico. E, naturalmente, qualsiasi forma di canto implica tipicamente il controllo della velocità di eccitazione.

Si è notato sopra che nella percezione delle vocali viene fatta una distinzione sistematica tra la qualità (o timbro) della vocale — ossia, la proprietà del segnale che permette d'identificare una vocale come [a], un'altra come [i], ecc. — e il modo, ossia il livello melodico, con cui la vocale viene prodotta. Si è inoltre notato che la prima proprietà è controllata dalla forma geometrica della cavità sopralaringale, mentre la seconda è controllata dalla configurazione della glottide. Si è ora in grado di capire che questa distinzione percettiva ha anche una base acustica, dal momento che la qualità vocale è correlata con le risonanze presenti nel segnale, mentre il livello melodico è legato alla velocità di eccitazione, detta anche frequenza fondamentale.

Quando la colonna d'aria contenuta in un tubo viene eccitata introducendo a un'estremità uno sbuffo d'aria, ciò provoca la compressione dell'aria a quell'estremità del cilindro. Poiché l'aria è un mezzo elastico, tale compressione viaggia sino all'altra estremità, dove viene riflessa e ritorna indietro verso l'estremità opposta del cilindro, per essere ancora una volta riflessa, e così via. A proposito di questa propagazione in andata e ritorno della compressione attraverso il tubo vi sono due cose da osservare. A causa dell'elasticità del-

l'aria, ogni compressione è seguita da una rarefazione e, come le compressioni, anche le rarefazioni viaggiano su e giù lungo il cilindro. Poiché vi sono molte contrazioni e rarefazioni che viaggiano contemporaneamente in su e in giù per il cilindro, si può facilmente capire che a un certo punto una compressione che si sposta verso il basso può incontrare una rarefazione che viaggia nella direzione opposta. Il risultato di tale incontro è che la compressione annulla la rarefazione, e in quel punto l'aria non è né compressa né rarefatta. Il fatto curioso, a proposito del comportamento effettivo delle compressioni e rarefazioni entro un cilindro verticale come quello in esame è che si viene a produrre una configurazione perfettamente stabile di compressioni e rarefazioni, la quale è correlata in modo specifico con ciascuna delle frequenze di risonanza del cilindro; ma di ciò si parlerà dettagliatamente più avanti.

In fisica questo fenomeno è noto come «configurazione di onde stazionarie»; le zone lungo il cilindro in cui la compressione annulla la rarefazione si chiamano nodi di pressione, mentre le zone dove esse risultano indipendenti al massimo grado si chiamano «ventri di pressione». Come è chiaramente illustrato nella figura 2, ciascuna risonanza ha una propria configurazione spe-

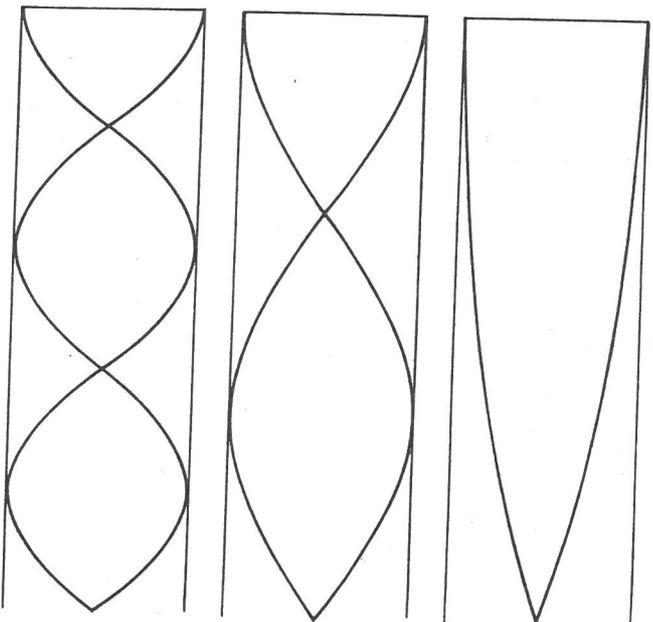


Figura 2.

Configurazioni di onde stazionarie delle tre risonanze più basse in un tubo a sezione trasversale uniforme chiuso a un'estremità.

cifica di nodi e di ventri, la collocazione dei quali è determinata dall'equazione (1).

La configurazione di onde stazionarie fornisce uno strumento assai utile per predire il complesso degli effetti acustici che si producono variando la forma geometrica del condotto vocale, da quella di un tubo con sezione trasversale uniforme a quella di un tubo con sezione trasversale variabile. Si ricorderà che la forma assunta dal condotto vocale quando si pronunzia una vocale come [a] è approssimativamente quella di un tubo con sezione trasversale uniforme, le cui risonanze sono poste a 500, 1500, 2500 Hz. Si deve a Chiba e Kajiyama (1941) quest'importante regola basata sull'esperienza pratica: Quando una parte di un tubo viene ristretta, ciò produrrà una diminuzione o un aumento nella frequenza di risonanza, a seconda che la costrizione sia prodotta vicino a un nodo o vicino a un ventre di pressione, associati con la configurazione di onde stazionarie della risonanza in questione. Data questa regola eminentemente pratica, si consideri che cosa ci si può aspettare quando il condotto vocale viene ristretto all'altezza delle labbra. Poiché tutte le risonanze hanno un ventre di pressione in quel punto, ci si dovrà aspettare che tutte le formanti (o risonanze), senza eccezione, vengano abbassate rispetto alle risonanze proprie della vocale neutra.

Si esaminino ora che cosa accade quando la costrizione viene spostata all'indietro lungo la cavità orale. È ovvio che la prima risonanza aumenterà di frequenza man mano che ci si muove all'indietro, poiché ci si sta allontanando dal ventre di pressione associato a questa risonanza per avvicinarsi a un nodo di pressione. Confrontando la tabella delle frequenze di risonanza, si vede che le vocali [a] e [æ], caratterizzate da costrizioni nel condotto vocale assai prossime alla glottide, presentano anche le più alte prime formanti.

Il comportamento della seconda risonanza sarà assai diverso. Man mano che la costrizione si allontana dalle labbra, la frequenza di risonanza aumenterà finché non si raggiunge il ventre di pressione situato a un terzo della distanza dalle labbra alla glottide. Oltre quel punto, la seconda risonanza diminuirà finché non si raggiunge il nodo di pressione situato a due terzi della lunghezza del cilindro, punto a partire dal quale la frequenza di risonanza aumenta di nuovo. Si può infatti constatare che la vocale [i], caratterizzata dalla massima costrizione nella parte anteriore della bocca (a circa un terzo della distanza dalle labbra alla glottide), presenta anche la seconda formante più alta. D'altra parte, la vocale [u], in cui la costrizione è molto più arretrata rispetto ad [i] (anche se non così arretrata come in [a]) possiede la più bassa seconda formante tra tutte le vocali.

Queste inferenze, circa il comportamento delle risonanze in rapporto alle configurazioni articolatorie delle diverse vocali, trovano conferma nelle frequenze di risonanza delle vocali:

i	e	æ	a	ɔ	o	u	
275	450	875	800	625	425	300	Risonanza 1
2300	2000	1800	1200	1000	950	850	Risonanza 2.

Si può, peraltro, far uso di queste stesse considerazioni per gettare qualche luce sul comportamento acustico delle consonanti. Dal punto di vista articolatorio, la differenza essenziale tra vocali e consonanti sta nel fatto che, nel produrre queste ultime, il condotto vocale dev'essere ristretto fino al punto di ostacolare o bloccare del tutto il libero flusso dell'aria. Quando il flusso d'aria viene ostacolato o bloccato, ciò equivale a dividere il condotto vocale in due cavità distinte: una cavità posteriore, chiusa ad entrambe le estremità (un'estremità essendo situata all'altezza della glottide e l'altra nel punto di costrizione), e una cavità anteriore, chiusa solo nel punto della costrizione. (Si trovano anche suoni linguistici che presentano due punti principali di co-ostrizione, o meglio occlusione: per esempio, i *clifs* dell'ottentoto. Questo tipo di articolazioni crea tre cavità distinte nel condotto vocale. Non si analizzeranno qui, tuttavia, le articolazioni di questo tipo, e ci si limiterà a considerare articolazioni con una sola costrizione nella cavità sopralaringale).

Si ricorderà che le risonanze di una cavità chiusa ad entrambe le estremità sono espresse dalla formula (3), che qui si riporta per comodità del lettore:

$$F_n = \frac{c}{4a} \times n$$

dove $n=0, 2, 4, 6$. Come si è già notato, la risonanza più bassa di una simile cavità sarà sempre a 0 Hz, indipendentemente dalla lunghezza del tubo. In altre parole, ogni qualvolta si produce un'occlusione in qualsiasi punto del condotto vocale, la risonanza più bassa tenderà verso 0 Hz.

Il comportamento delle formanti più alte può essere predetto sulla base del comportamento di un sistema di risonanza come quello della figura 3.

Dato un simile sistema, è ragionevolmente possibile supporre, in prima approssimazione, che le due cavità di risonanza siano del tutto indipendenti l'una dall'altra. La cavità posteriore avrà risonanze che sono date dalla formula (3), mentre la cavità anteriore avrà risonanze che sono date dalla formula (1), che qui si riporta per comodità del lettore:

$$F_a = \frac{c}{4a_a} \times n$$

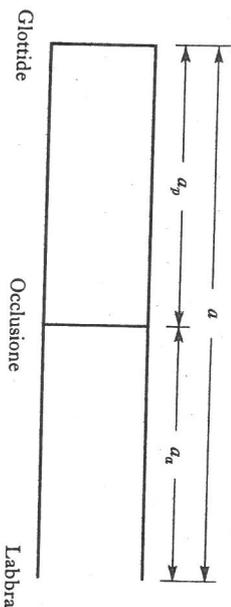


Figura 3. Sistema di risonanza atto a predire il comportamento delle formanti più alte.

dove $n=1, 3, 5, 7$. Si noti che la lunghezza della cavità anteriore a_a è uguale all'intera lunghezza del condotto vocale a (cui si assegnerà come prima il valore di 17,25 cm), meno la lunghezza della cavità posteriore a_p . La formula appena data può quindi essere riscritta come segue:

$$F_a = \frac{c}{4(17,25 - a_p)} \times n.$$

Data la formula qui sopra riportata, si possono tradurre i valori di F_p e F_a in funzione di a_p , ossia in funzione della distanza tra l'occlusione e l'estremità chiusa del tubo. Per $a_p=0$ si considera soltanto una cavità, vale a dire la cavità anteriore, che risuonerà a 500, 1500, 2500 Hz, ecc. Man mano che a_p aumenta, la cavità anteriore diminuirà di lunghezza, e quindi le sue frequenze di risonanza aumenteranno, come è mostrato dalla linea tratteggiata nella figura 4.

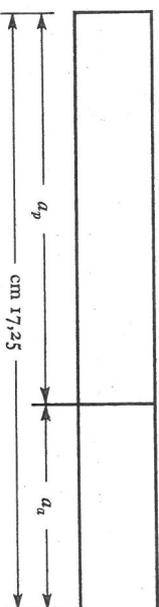
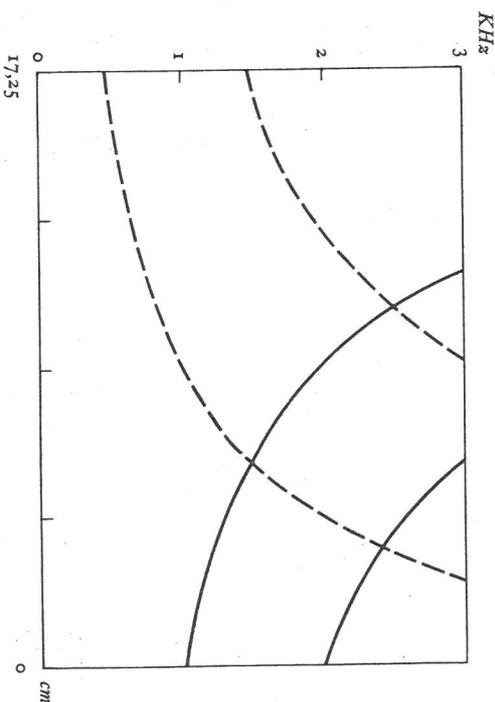


Figura 4. Le linee tratteggiate rappresentano le frequenze di risonanza della cavità anteriore aperta, in funzione della collocazione dell'occlusione; le linee continue, quelle della cavità posteriore chiusa. Si noti che quando l'occlusione è situata in modo tale che $a_p=2a_a$, le risonanze più basse delle due cavità coincidono in frequenza.

Si considerino ora le risonanze dovute alla cavità posteriore. Quando l'occlusione è all'altezza dell'estremità aperta del tubo, si avrà soltanto una cavità chiusa ad entrambe le estremità; pertanto, le sue risonanze, come stabilito sopra, saranno di 0, 1000, 2000 Hz, ecc. per un normale condotto vocale maschile. Man mano che l'occlusione si sposta all'indietro verso l'estremità chiusa del tubo, le risonanze (ad eccezione di quelle più basse che rimarranno sempre a 0 Hz) aumenteranno di frequenza, poiché la lunghezza del tubo viene progressivamente diminuita. Ciò è mostrato dalla linea continua nella figura 4.

In base alla figura 4 è possibile fare predizioni qualitative circa il comportamento delle risonanze durante l'esecuzione di una sequenza costituita da una vocale seguita da una consonante. A questo scopo, basta leggere le risonanze nella figura 4, nel punto particolare in cui è prodotta l'occlusione. Occorre tuttavia ricordare che, essendo una delle due cavità create dall'occlusione chiusa ad entrambe le estremità, la risonanza più bassa prodotta da questa configurazione sarà sempre a 0 Hz; ossia, indipendentemente dal punto in cui si produce l'occlusione, la risonanza più bassa della consonante sarà di

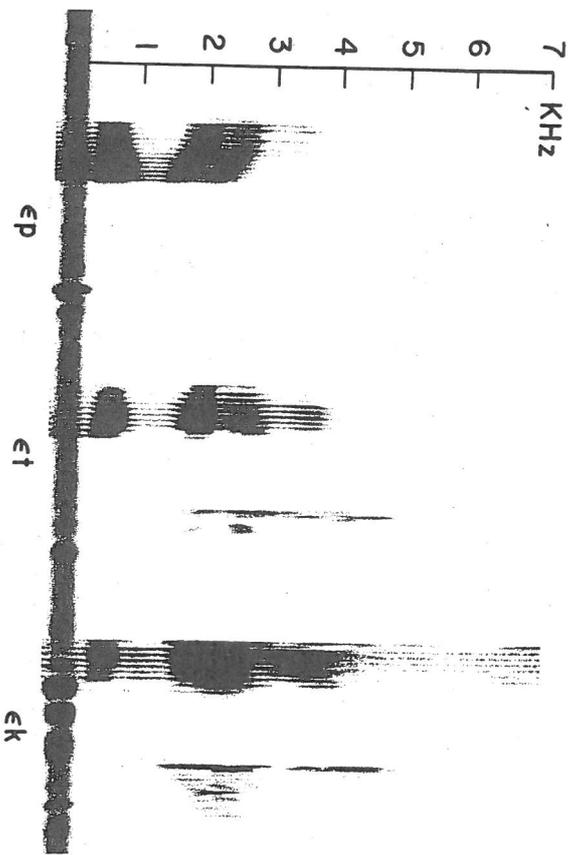


Figura 5.

Spettrogrammi delle sillabe /ep/, /et/, /ek/. Le barre orizzontali più marcate rappresentano le diverse risonanze (o formanti) della vocale. Esse si spostano verso il basso (cfr. /ep/); si avvicinano quando l'occlusione viene prodotta con un'occlusione all'altezza delle labbra /ek/; quando invece l'occlusione è prodotta all'altezza del velo palatino (cfr. la seconda risonanza resta invariata e la terza si sposta.

o Hz. Le rimanenti risonanze sono date nella figura 4. Se si produce l'occlusione all'altezza delle labbra (come avviene in suoni come [p, m]), si ottiene $R_2 = 1000$ e $R_3 = 2000$. Quando si produce l'occlusione in un punto situato a un quarto della distanza dalle labbra alla glottide - all'incirca come avviene in [ʃ] - si avrà $R_2 = 1333$ e $R_3 = 2000$. Si noterà che quando l'occlusione viene prodotta all'altezza delle labbra, la terza risonanza è dovuta alla cavità anteriore; quando invece l'occlusione è prodotta più verso l'interno della cavità, R_3 è dovuta alla cavità posteriore. Quando l'occlusione è prodotta in un punto situato a un terzo della distanza dalle labbra alla glottide - cioè, più o meno nel punto in cui si colloca l'occlusione nel produrre [k] o [g] - si otterrà $R_2 = R_3 = 1500$. Questo modello altamente idealizzato predice il comportamento reale delle risonanze in modo assai soddisfacente, come mostrato negli spettrogrammi della figura 5. Esso indica come, nella transizione dalla vocale [a] verso una qualsiasi consonante seguente, la risonanza più bassa diminuisca sempre di frequenza. Il comportamento delle altre risonanze dipende dalla posizione dell'occlusione - in [p] R_2 e R_3 si abbassano, in [t] sia R_2 sia R_3 si alzano, mentre in [k], R_2 si alza e R_3 si abbassa.

4. Sui tratti distintivi.

Si consideri ora che cosa accade quando si pronunziano vocali diverse, per esempio quando si pronunziano in successione le vocali [a] e [u]. Pronunziando [a], le labbra sono nettamente aperte e il corpo della lingua è spinto verso il basso e arretrato il più possibile, non diversamente da quanto avviene per la radice della lingua. D'altra parte, in [u] le labbra sono arrotondate in modo da formare un'apertura relativamente piccola, e il corpo della lingua è sollevato verso il palato, senza tuttavia arrivare propriamente a toccarlo, mentre la radice della lingua è spinta in avanti. Differenti configurazioni della cavità sopralaringale danno quindi luogo alla produzione di vocali diverse. In effetti, ciascuna vocale possiede la sua specifica configurazione della cavità sopralaringale.

È evidente che si è in grado di controllare le corde vocali indipendentemente dalla cavità posta al di sopra. Per esempio, con la lingua e le labbra è possibile assumere la configurazione tipica per la vocale [a], e poi variare la configurazione delle corde vocali producendo dapprima una [a] bisbigliata e quindi una [a] pienamente sonorizzata. Inoltre, in quest'ultimo caso possiamo anche variare il livello melodico della voce, dal livello più basso di cui si è capaci fino al più alto. Tuttavia, nessuna delle variazioni prodotte medianamente le corde vocali influisce sulla qualità intrinseca della vocale stessa: essa rimane sempre una [a]. Il segnale acustico continuo che colpisce il timpano è quindi analizzato percettivamente in base a componenti distinte: la qualità intrinseca della vocale (o timbro) e il modo della sua realizzazione (bisbiglio o sonorizzazione). Anche queste due componenti percettivamente distinte dello stimolo acustico continuo sono prodotte da strutture anatomiche distinte,

controllate in modo indipendente l'una dall'altra. Questo parallelismo tra il modo in cui il sistema uditivo analizza un suono linguistico e il meccanismo articolatorio che lo produce, è un fatto assai notevole, dato che biologicamente questi due sistemi sono affatto distinti. Sarebbe certo sorprendente se il suono prodotto da uno schiocco delle dita della mano destra suonasse sistematicamente diverso dal suono prodotto dallo schiocco delle dita della mano sinistra. Eppure, nel caso del linguaggio accade proprio qualcosa del genere: certi aspetti del segnale linguistico, che sono controllati da un dato meccanismo articolatorio, manifestano di solito una specifica proprietà acustica, e sono percepiti in modo altrettanto specifico del sistema uditivo. Darò altri due esempi.

Certi suoni vengono percepiti come dotati di un timbro nasale. In particolare, in una lingua come il francese, si distinguono le vocali nasali dalle altre vocali tramite un attributo psicoacustico chiaramente percepito, detto appunto «nasalità». In francese si trovano dunque vocali nasali in parole come

bon banc bain

e vocali non-nasali (o orali) in parole come

beau bas baie.

La qualità percettivamente uniforme della «nasalità», presente nei suoni della prima serie, viene prodotta in maniera altrettanto uniforme abbassando il velo palatino e permettendo all'aria di passare attraverso il naso. (Che i suoni nasali siano di fatto prodotti nel modo descritto si può facilmente verificare ponendo la punta di un dito sul naso mentre si pronunziano le parole francesi elencate sopra: se pronunziati correttamente, i suoni nasali saranno accompagnati da una lievissima sensazione di vibrazione, mentre nei suoni orali la vibrazione sarà assente).

Un secondo esempio circa la coincidenza tra attributi acustici e articolatori dei suoni linguistici ci è dato da vocali come [ɪ] e [ɛ], che hanno la seconda risonanza notevolmente più alta rispetto a vocali come [a] o [ɔ]. Alla differenza acustica si associa una chiara differenza nella posizione del corpo della lingua: esso è spinto in avanti nel caso di vocali con una seconda risonanza elevata, mentre è ritratto verso il fondo della cavità orale nel caso di vocali con una seconda risonanza bassa.

La principale deduzione che si può trarre da ciò che si è appena detto è che i suoni verbali non sono, per così dire, una sorta di atomi del linguaggio, compatti e non ulteriormente analizzabili, bensì dei complessi di attributi che appaiono radicati nel meccanismo articolatorio e in quello uditivo degli esseri umani: i suoni sono insomma dei «fasci di tratti distintivi». A prima vista, tutto ciò può sembrare nulla più che una pedante riformulazione delle osservazioni fatte sopra. Che si tratti di qualcosa di molto più significativo dovrebbe divenire evidente una volta considerati i fatti che si passerà ora a trattare.

Tutte le lingue sono soggette a regole speciali che guidano la composizione fonetica delle parole. Nelle descrizioni linguistiche tali regole sono discusse sotto varie rubriche, quali «fonologia», «morfologia», «morfofonemica», ecc.

Così, per esempio, il prefisso italiano *s* è pronunziato [s] se affisso a radici lessicali che iniziano con una consonante sorda (per esempio, *scorzare, sforzare, spremere*), ma è pronunziato [z] negli altri casi (per esempio: *smenitare, smolare, slogare, sbagliare, sdrucire, svegliare*). Ossia, il prefisso è pronunziato sordo quando è attaccato a una parola che inizia con una consonante anch'essa sorda; è invece pronunziato sonoro quando la consonante seguente è sonora. (L'inglese presenta una regola formalmente assai simile: cf. oltre). Nell'ebraico biblico, la congiunzione [w] 'e' compare nella forma [u] quando la parola seguente inizia con [p] [b] [f] [v] o [m], cioè con una consonante prodotta con una costrizione all'altezza delle labbra. In inglese vi sono tre suffissi plurali [z] [s] e [z], la cui distribuzione è determinata dal suono con cui termina la radice del sostantivo:

a) se il sostantivo termina con [s, z, ʃ, sʒ, tʃ, j], pl. → [ɪz];

b) se il sostantivo termina con [p, t, k, f, θ], pl. → [s];

c) altrimenti, pl. → [z].

Anche in questo caso abbiamo una classe di consonanti che condividono ovvi attributi fonetici. La serie in a) consiste delle consonanti «sibilanti e palatali», mentre la serie in b) comprende le restanti consonanti che sono prodotte senza vibrazione delle corde vocali.

Anche in questo caso siamo in presenza di una coincidenza degna di nota: non vi è alcuna ragione speciale per cui le regole del tipo appena considerato debbano essere sensibili a serie di suoni che hanno in comune specifici attributi fonetici. Non è logicamente inconcepibile che, anziché davanti alla serie [p, f, b, v, m], la variante della congiunzione ebraica possa apparire davanti a cinque altre consonanti qualsiasi utilizzate nella stessa lingua (per esempio, prima di [p, d, z, l, s]); o che il suffisso plurale inglese [ɪz] possa essere aggiunto a sostantivi che terminano con [m, f, t, d, g, r]. Tuttavia, non si sono mai riscontrate regole che implichino simili serie arbitrarie di suoni.

Ciò che si è scoperto, invece, è che tutte le regole in tutte le lingue riguardano classi di suoni che, come quelli illustrati nel paragrafo precedente, hanno in comune alcuni tratti distintivi. Questa scoperta suggerisce dunque che l'analisi in tratti distintivi non è soltanto un comodo procedimento utilizzato per trattare certi fatti concernenti l'articolazione e l'acustica dei suoni linguistici, ma rispecchia da vicino l'immagine che è implicita nella concezione della lingua posseduta dal comune parlante; il cui comportamento effettivo dovrà poi essere descritto dal linguista.

Per scorgere con maggior chiarezza il carattere non-arbitrario dell'analisi dei tratti distintivi qui proposta, può essere utile paragonarla a un'altra analisi di tipo relativamente arbitrario. Gli insegnanti di calligrafia (la cui scomparsa dalle nostre scuole sembra essere stata accompagnata da una singolare assenza di lamentele da parte degli scolari), erano soliti insegnare quando io ero ragazzo che le 26 lettere dell'alfabeto non sono affatto delle entità non ulteriormente scomponibili, ma consistono piuttosto in fasci (sequenze) di tratti di penna del tipo mostrato nella figura 6. Secondo questa teoria, si soste-

neva che i tratti di penna sono più facili da scrivere delle lettere; di conseguenza, agli allievi si faceva far pratica con le 'aste' prima di permettere loro di scrivere delle lettere complete, per non parlare poi delle parole. Che tale teoria sia o no corretta, non è questo il punto; ciò che importa è che gli insiemi di lettere che presentano gli stessi tratti di penna — cioè, le lettere su ciascuna delle righe e delle colonne nella figura 6 — non rivestono particolare interesse sul piano linguistico. Non esiste alcuna regola linguistica che coinvolga questi insiemi di lettere in quanto tali, né si dà alcun fenomeno grammaticale in cui questi insiemi manifestino un comportamento particolarmente rilevante. Ciò è in stridente contrasto con gli insiemi di suoni imperniati su comuni tratti distintivi, in cui è assai facile scoprire ogni sorta di regolarità linguistiche; le quali, per essere espresse, richiedono per l'appunto la costruzione di tali insiemi.

Vi sono, comunque, prove anche più dirette circa la realtà psicologica dell'analisi in tratti dei suoni linguistici. Tali prove rimandano alla regola del plurale inglese espressa sopra a p. 303. La formulazione ivi data non ricorre ai tratti distintivi: essa contiene soltanto degli elenchi di suoni linguistici. Ma se è vero che i suoni linguistici sono dei complessi di tratti distintivi, ci si aspetterebbe che regole come quella del plurale inglese siano formulate anche

l	l i t	l u z	l r	l j	l ɔ	l
e	e e l	e b	e h		e k p	e f
e	e e	e a d o		e g		e q
r	r x	r w w	r n m	r y		
j	j j					
ɔ	ɔ			ɔ z		

Figura 6.

Un'analisi in tratti distintivi (tratti di penna) delle lettere dell'alfabeto latino in scrittura corsiva.

sulla base di tratti, non sulla base di suoni linguistici non meglio analizzati. Questo vuol dire che invece della regola di p. 303, ci si dovrebbe aspettare:

- se il sostantivo termina con un suono che presenta il tratto «sibilante o palatale», pl. → [zɔ];
- se il sostantivo termina con un suono che presenta il tratto «sibilante», pl. → [s];
- altrimenti pl. → [zɔ].

È possibile ora eseguire un esperimento per stabilire se un qualunque parlante inglese ha interiorizzato la regola del plurale nella forma di p. 303 o nella forma presentata sopra. L'esperimento, che mi è stato proposto da Lise Menn, è assai semplice. Esso consiste nel domandare a dei parlanti inglesi, anche di buona cultura, ma privi di particolari conoscenze linguistiche, come formerebbero il plurale di nomi tedeschi terminanti con [x] o [ʃ], come in *Bach* e in *Milch*, rispettivamente. Queste parole non sono ignote ai parlanti colti; tutti conoscono il nome del compositore tedesco Bach, e alcuni ricordano ancora il nome del comandante Milch della Luftwaffe hitleriana. Come tutti i parlanti inglesi, i nostri soggetti sono evidentemente in grado di formare i plurali dei nomi propri, anche se non li hanno mai uditi prima. Si obietterà che né [x] né [ʃ] sono suoni dell'inglese standard: ma ciò non costituisce un ostacolo, o almeno non è sembrato che lo fosse nei pochi test che ho condotto.

Prima di rivelare il risultato del test, è utile domandarsi cosa ci si dovrebbe aspettare sulla base delle due formulazioni alternative. Poiché né [x] né [ʃ] sono suoni dell'inglese, non ci si aspetta certo di trovarli in alcun elenco di suoni in una regola dell'inglese. Quindi, se il parlante ha interiorizzato la regola del plurale nella forma di p. 303, la sola soluzione che gli viene offerta è la c), per cui egli dovrebbe rispondere rispettivamente [baxzɔ] e [milʃzɔ]. D'altra parte, se egli ha interiorizzato la regola del plurale nella seconda forma proposta, ci si dovrebbe aspettare che egli selezioni l'alternativa b), dato che [x] e [ʃ] non presentano gli attributi di «sibilante e palatale», ma possiedono l'attributo di «sordità»; si dovrebbe dunque presumere che egli dica [bax + s] e [milʃ + s].

Quando s'interrogano i parlanti inglesi, essi di solito forniscono quest'ultima risposta. Bisogna quindi concludere che i parlanti inglesi interiorizzano la regola del plurale nella seconda forma, eminentemente astratta, anziché nella prima forma, molto più concreta. E poiché si possono fare esperimenti analoghi in lingue diverse dall'inglese, ottenendo con ogni probabilità risultati analoghi, i quali dimostrano che tutti i parlanti interiorizzano le regole della propria lingua in una forma in cui i suoni linguistici sono analizzati nei loro tratti distintivi, si deve obbligatoriamente concludere che i comuni parlanti conoscono implicitamente l'articolazione in tratti distintivi tipica dei suoni della loro lingua. Sorge quindi subito la curiosità di sapere dove e quando è stata acquisita tale conoscenza. È chiaro che non è possibile indicare alcuna esperienza viva in cui i parlanti inglesi apprendano che [f, ʃ, p, t, k] hanno in comune il tratto della sordità; eppure essi agiscono come se avessero

Tabella 1.
Suoni dell'italiano standard e tratti distintivi che li caratterizzano.

	<i>Vino</i>	<i>Cesta</i>	<i>Testa</i>	<i>Uno</i>	<i>Fosco</i>	<i>Tosca</i>	<i>Nato</i>	<i>Pace</i>	<i>Bacio</i>	<i>Fatto</i>	<i>Via</i>	<i>Morte</i>	<i>Tanto</i>	<i>Do</i>	<i>Pezzo</i>	<i>Mezzo</i>	<i>Sono</i>	<i>Nota</i>	<i>Luna</i>	<i>Rosa</i>	<i>Ciao</i>	<i>Giù</i>	<i>Sciopero</i>	<i>Gnocchi</i>	<i>Maglio</i>	<i>Cane</i>	<i>Gatto</i>
Sillabico	+	+	+	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Sonante	+	+	+	+	+	+	+	-	-	-	-	+	-	-	-	-	-	+	-	+	-	-	+	+	-	-	-
Alto	+	-	-	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	+	-	+	+	+	+	+	+
Posteriore	-	-	-	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+
Basso	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
RLA	+	+	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Coronale	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+	+	+	+	+	+	+	-	-	-
Anteriore	-	-	-	-	-	-	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	-
Labiale	-	-	-	+	+	+	-	+	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Nasale	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	+	-	-	-	-	-	+	-	-	-
Laterale	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	+	-	-	-
Esplosivo	-	-	-	-	-	-	-	+	+	-	-	+	+	+	+	+	+	+	-	-	-	+	+	+	+	+	+
Corde vocali rigide	-	-	-	-	-	-	-	+	-	+	-	-	+	-	+	-	+	-	-	-	-	-	+	-	-	+	-
Stridulo	-	-	-	-	-	-	-	-	-	+	+	-	-	-	+	+	+	-	-	-	-	+	+	-	-	-	-

precisa conoscenza di questo fatto. Questo paradosso suggerisce che forse la domanda è posta in modo scorretto. È probabile che sia errato il presupposto stesso da cui nasce la domanda, ossia che si venga al mondo come *tabula rasa*, e che ogni nostra conoscenza debba essere necessariamente acquisita. Vorrei suggerire, invece, alla luce dei fatti appena considerati, che esiste una conoscenza che non può essere acquisita, né ha bisogno di esserlo. Un esempio di questo tipo di conoscenza è quella richiesta agli esseri umani per analizzare i suoni linguistici come complessi di tratti distintivi. Altri esempi sono la conoscenza che un salmone deve avere per poter ritornare al fiume in cui è stato generato, o quella che deve avere un'ape per eseguire correttamente il proprio ruolo nella vita dell'alveare, o un uccello per costruire il nido. Il possesso di questo tipo di conoscenza è ciò che distingue una specie da un'altra, per cui si può dire che la conoscenza implicita nell'analisi in tratti distintivi dei suoni linguistici fa parte del corredo genetico dell'*Homo sapiens*.

5. Correlati articolatori e acustici di alcuni tratti distintivi.

Scopo primario di quanto precede è stato quello di stabilire che, per gli esseri umani normali, il linguaggio è qualcosa di più che un insieme di eventi acustici prodotti da certi movimenti, eseguiti da opportune strutture anatomiche dell'uomo. Per soggetti umani normali, il linguaggio è anche e sempre una sequenza di parole, le quali a loro volta sono sequenze di suoni discreti, ciascuno dei quali è di per sé un complesso di tratti distintivi. È proprio questa concezione gerarchica del linguaggio che è stata posta alla base delle ricerche fonetiche, e che fa della fonetica una disciplina autonoma, distinta e separata dall'acustica e dall'anatomia. I suoni linguistici e i tratti distintivi costituiscono gli elementi di base della fonetica, così come gli elementi e i composti chimici sono gli elementi di base per la chimica, e la cellula e il gene per la biologia. In questo paragrafo descriverò alcuni tratti distintivi e ne illustrerò l'uso per la caratterizzazione di specifici suoni linguistici. Limito la mia discussione ai tratti e ai suoni del moderno italiano standard (tab. 1).

La descrizione che segue presenta quei correlati articolatori e acustici dei tratti distintivi che svolgono una funzione primaria nel sistema fonologico dell'italiano (tab. 1). Sebbene quest'area sia stata oggetto d'intensa ricerca per più di un secolo, vi sono ancora molte questioni cruciali a cui non è stata data una risposta soddisfacente. Inoltre, per ciò che riguarda alcuni punti, una soluzione soddisfacente richiederebbe una trattazione assai ampia di certi problemi, il che va ben al di là dei limiti di questo articolo. Quanto segue sarà quindi un tantino schematico e incompleto.

5.1. Sillabico/non-sillabico.

Dal punto di vista acustico, i suoni sillabici sono molto più prominenti dei suoni non-sillabici. I suoni sillabici costituiscono picchi di prominenza in

una sequenza di suoni contigui, i non-sillabici formano il resto. Per produrre suoni di maggiore intensità rispetto a quelli vicini, i parlanti dispongono di un certo numero di meccanismi: essi possono pronunziare il suono con maggior forza, e possono utilizzare una configurazione del condotto vocale che trasmetta in modo più efficiente l'energia acustica. Quindi, una vocale come [a], prodotta con la cavità orale che si allarga verso l'aria esterna, è un evento acusticamente molto più prominente del momentaneo scoppio di energia che caratterizza i suoni occlusivi labiali [p] o [b].

5.2. Sonante/ostruente.

I suoni ostruenti sono prodotti con costrizioni nel condotto vocale che imprigionano l'aria, producendo così un aumento della pressione interna, che viene a superare notevolmente quella dell'aria esterna. I suoni sonanti sono prodotti con il condotto vocale aperto in modo sufficiente a permettere che la pressione interna sia essenzialmente uguale a quella dell'aria esterna. Affinché la pressione all'interno della cavità aumenti, è necessario — ma non sufficiente — che il velo palatino sia sollevato, cosicché l'aria non possa sfuggire dal naso. Inoltre, per produrre il necessario aumento di pressione all'interno della cavità, si richiede una seconda costrizione in qualche punto della cavità stessa. Questa differenza nella produzione di tali due classi di suoni è correlata con la differenza acustica tra suoni aventi risonanze relativamente non smorzate, (sonanti) e suoni aventi risonanze molto smorzate (ostruenti). Nelle sonanti, quindi, l'energia acustica è concentrata entro zone relativamente ristrette nella parte più bassa dello spettro di frequenza, mentre nelle ostruenti l'energia acustica è diffusa su ampie bande dello spettro.

5.3. Alto/non-alto.

I suoni linguistici che possiedono l'attributo di «alto» sono prodotti sollevando il corpo della lingua verso il palato; i suoni classificati con «non-alti» sono prodotti senza che tale sollevamento avvenga. Il grado di costrizione prodotto da questo sollevamento del corpo della lingua è soggetto alle limitazioni imposte dal tratto sonante/ostruente. Quindi, nelle sonanti la costrizione prodotta dal sollevamento della lingua non può giungere ad ostruire il flusso d'aria fino al punto di produrre una forte pressione all'interno del condotto vocale, mentre per le ostruenti tali restrizioni non valgono.

I correlati acustici del tratto alto/non-alto sono differenti nelle sonanti e nelle ostruenti. Nelle sonanti, i suoni alti presentano una prima risonanza bassa al massimo grado (circa 300 Hz), mentre i suoni non-alti hanno la risonanza più bassa considerevolmente al di sopra di questo valore. Come si è visto sopra (§ 3), nelle ostruenti la risonanza più bassa si approssimerà sempre a 0 Hz, causa il prodursi, all'interno del condotto vocale, di una cavità chiusa ad entrambe le estremità. Pertanto, gli effetti acustici determinati dal prodursi di una costrizione contro il palato (cioè, nel terzo anteriore del condotto vo-

cale) non influiscono sulla risonanza più bassa. Una costrizione di questo tipo causerà piuttosto un avvicinamento delle frequenze della seconda e della terza risonanza, o della terza e della quarta. L'assenza di questo avvicinamento indica che l'ostruente in questione è non-alta.

5.4. Posteriore/non-posteriore.

Nei suoni linguistici che presentano il tratto «posteriore», il corpo della lingua è avvicinato alla parete della faringe; nei suoni che sono non-posteriori, il corpo della lingua è invece allontanato dalla parete della faringe. L'effetto acustico primario di questo tratto riguarda la frequenza della seconda risonanza: essa è bassa al massimo grado nei suoni posteriori, e alta al massimo grado nei suoni non-posteriori.

5.5. Basso/non-basso.

I suoni «bassi» sono prodotti ritraendo la radice della lingua il più possibile verso la parete della faringe. Questo movimento produce di norma, come effetto secondario, l'abbassamento della lingua; tuttavia, l'abbassamento del corpo della lingua non è automatico. Per esempio, si possono produrre le vocali alte [i] e [u] mentre si ritrae la radice della lingua: questi suoni compaiono in alcune lingue semitiche.

Nelle sonanti, la costrizione prodotta dalla ritrazione della radice della lingua aumenterà al massimo la frequenza della risonanza più bassa. Nelle ostruenti, la costrizione produrrà anche un abbassamento delle tre o quattro risonanze immediatamente adiacenti.

5.6. Radice della lingua avanzata (RLA).

Questo tratto, come indica il nome, si realizza avanzando la lingua il più possibile. Esso influisce in qualche misura sulla posizione del corpo della lingua, così come il tratto «basso»; tuttavia, i tratti che controllano la posizione del corpo della lingua (posteriore e alto) sono indipendenti dai tratti che controllano la posizione della radice della lingua (basso e RLA): ogni combinazione dei primi è compatibile con ogni combinazione di valori dei secondi. Basso e RLA, d'altra parte, non sono indipendenti. Un suono che sia [+basso] non può essere anche [+RLA]: per contro, le altre tre combinazioni di tratti sono possibili. Il tratto RLA svolge di solito il proprio ruolo soltanto tra le vocali: esso corrisponde alla vecchia distinzione di A. M. Bell tra suoni «stretti» e «larghi». Tuttavia, pare che esistano sistemi consonantici — l'indonesiano è un esempio spesso citato — in cui le consonanti sono marcate dal tratto RLA. Acusticamente, le vocali con radice della lingua avanzata presentano una configurazione di risonanza che si discosta nettamente da quella prodotta da una cavità orale che funzioni come un tubo di sezione trasversale uniforme chiuso a un'estremità (cfr. § 3); tale differenza è invece meno marcata nel caso delle vocali prodotte senza avanzamento della radice della lingua.

5.7. Coronale/non-coronale.

I suoni coronali vengono prodotti sollevando il bordo della lingua verso il palato. Dal punto di vista anatomico, ciò implica la contrazione dei cosiddetti muscoli longitudinali intrinseci della lingua, che si dividono in due insiemi: uno superiore ed uno inferiore. Quando si contrae soltanto l'insieme superiore, la costrizione viene formata con la punta della lingua (apice); quando sono contratti entrambi, la costrizione viene invece formata con una porzione più ampia del bordo della lingua. La coronalità è universale nei sistemi consonantici, mentre è relativamente rara tra le vocali. (La vocale inglese [ə] — come in *fur* o *err* — è un esempio di questo tipo di vocale relativamente insolita).

Acusticamente, i suoni coronali sono caratterizzati, nelle frequenze più alte dello spettro, da una energia maggiore rispetto ai loro corrispondenti non-coronali.

5.8. Anteriore/non-anteriore.

La corona alveolare, che costituisce la porzione anteriore del palato, serve come luogo di confine per la formazione di due classi di suoni. I suoni anteriori, come [n] [s] o [t], sono formati con una costrizione situata proprio davanti a questo confine, mentre i suoni non-anteriori, come [ŋ] o [ʃ], sono formati mediante costrizioni situate in posizione leggermente più arretrata.

Acusticamente, il correlato primario di questa distinzione è da mettere in relazione col fatto che, nei suoni anteriori, la cavità posta di fronte alla costrizione è più corta, e quindi possiede una frequenza di risonanza più elevata rispetto a quella dei corrispondenti suoni non-anteriori.

5.9. Labiale/non-labiale.

Come indica il termine, i suoni labiali sono formati con una costrizione all'altezza delle labbra, mentre quelli non-labiali sono formati senza tale costrizione. Fondamentalmente, vi sono due tipi di articolazione labiale: nel primo, quello che si ritrova per esempio nelle vocali [u] o [o], l'occlusione è accompagnata da una protrusione delle labbra; nel secondo, quello che si ritrova in consonanti come [p] o [m], l'occlusione non è accompagnata da alcuna protrusione delle labbra. In alcune lingue, la presenza o l'assenza di tale protrusione svolge un ruolo distintivo. Infatti nell'ewe, una lingua dell'Africa occidentale, vi sono due tipi di suoni [p], con o senza protrusione delle labbra; mentre lo svedese standard utilizza lo stesso meccanismo per distinguere due tipi di vocali labiali non-posteriori ([ɪ] e [ʏ]).

Dal punto di vista acustico, una costrizione labiale dà luogo a un abbassamento di tutte le risonanze della cavità.

5.10. Nasale/non-nasale.

I suoni nasali vengono prodotti abbassando il velo palatino, permettendo così all'aria di passare attraverso la faringe nasale e di uscire dal naso. Nei suoni non-nasali, il velo palatino rimane sollevato, e l'aria contenuta nelle cavità nasali non viene eccitata.

Acusticamente, questo abbassamento del velo palatino dà luogo al timbro «nasale», del quale — a quanto mi è dato conoscere — non si possiede a tutt'oggi un'adeguata caratterizzazione fisica.

5.11. Laterale/non-laterale.

I suoni laterali, il principale esempio dei quali è il suono [l], vengono prodotti abbassando uno o entrambi i lati della lingua, permettendo così all'aria di sfuggire da sotto canini e molari e impedendole di sfuggire da sotto gli incisivi.

Come per la nasalità, anche in questo caso manca un'adeguata caratterizzazione, in termini fisici, dei correlati acustici.

5.12. Esplosivo/continuo.

I suoni esplosivi vengono prodotti con un'occlusione totale nella cavità orale, che produce un'interruzione del flusso d'aria lungo la bocca. Nei suoni continui il flusso è invece ininterrotto. L'interruzione del flusso d'aria determina, ovviamente, un arresto nell'eccitazione delle risonanze dovute a tale flusso, e nella grande maggioranza dei casi ne deriva un'interruzione momentanea del suono emesso dal parlante. Quando il flusso d'aria riprende, il silenzio è rotto da un caratteristico e improvviso attacco (*onset*) del suono. Ma poiché il flusso d'aria non riprende alla fine di un enunciato, e poiché le esplosive possono facilmente ricorrere anche alla fine di un enunciato, l'improvvisa ripresa del suono che segue il silenzio non è necessaria per segnalare l'esplosività. D'altra parte, la mera cessazione della fonazione di per sé non segnala l'esplosività, poiché gli enunciati possono terminare con suoni continui, con la stessa facilità con cui terminano con suoni esplosivi.

Come si è osservato sopra, la cessazione del flusso d'aria lungo la bocca è prodotta dalla creazione di un'occlusione nella cavità orale. Ciò tuttavia, implica invariabilmente dei cambiamenti radicali nella forma geometrica della cavità, e quindi anche nelle risonanze che essa produce. Questi rapidi cambiamenti nelle frequenze di risonanza sono un importante indice di esplosività.

5.13. Corde vocali rigide / corde vocali non rigide.

L'irrigidimento delle corde vocali svolge un ruolo attivo, principalmente nelle ostruenti, quando serve a distinguere i suoni cosiddetti «sordi» (per esem-

pio, [s] [ʃ] [p] [k] dai corrispondenti suoni «sonori» (per esempio, [z] [ʒ] [b] [g]). Le ostruenti sonore, che vengono prodotte con le corde vocali in vibrazione, richiedono corde vocali relativamente allentate. Quando esse sono rigide, non possono vibrare, cosicchè vengono prodotte le ostruenti «sorde».

Nelle sonanti, gli effetti dell'irrigidimento delle corde vocali sono assai diversi. In questo caso, l'aumento della rigidità delle corde vocali non ne fa cessare la vibrazione, ma ne fa piuttosto aumentare la rapidità di vibrazione, che percettivamente corrisponde a un aumento del livello melodico. Poichè le variazioni nel livello melodico delle sonanti non svolgono un ruolo distintivo nelle parole italiane, ho contrassegnato tutte le sonanti come [—corde vocali rigide].

5.14. Stridulo/non-stridulo.

I suoni striduli sono caratterizzati, acusticamente, da una notevole rumorosità (turbolenza). Dal punto di vista articolatorio, ciò si ottiene dirigendo la corrente d'aria su di uno spigolo acuto con un angolo d'incidenza prossimo a 90°. Solo le ostruenti possono essere stridule; le sonanti sono non-stridule. (I suoni «sibilanti» e «palatali» cui si è fatto riferimento nella regola a p. 304 sono [+stridulo, +coronale].)

5.15. Affricate.

I suoni affricati non sono foneticamente distinguibili da sequenze del tipo «esplosiva+continua»; ciò nonostante, le affricate funzionano come suoni singoli piuttosto che come sequenze di suoni. Questo fatto pone una seria sfida al quadro teorico su cui si è fondata la discussione precedente, nella quale si è assunto che ogni cambiamento di valore di un dato tratto — nel caso delle affricate, da [+esplosivo] a [—esplosivo] — segnali la transizione da un suono al successivo.

Difficoltà di questo tipo, che obbligano ad ammettere una commutazione di tratti distintivi entro i confini di un singolo suono, hanno condotto a numerose e sostanziali proposte di modifica della teoria fonetica. Ma queste proposte — la più promettente delle quali sembra essere quella di John Goldsmith (1976) — non sono ancora suffragate da sufficienti prove sperimentali; non sembra perciò giustificata la loro inclusione in questa rassegna, visto che esse potrebbero subire radicali riformulazioni anche a brevissimo termine. [M. H.].

La fonetica viene definita come la **scienza** dei suoni del **linguaggio**. Essa si occupa della produzione e insieme della ricezione o **percezione** dei suoni, questi ultimi analizzati, a partire da un determinato *continuum*, nei loro elementi discreti (cfr. **continuo/discreto**) e proprietà costitutive. La fonetica, rappresentando materia d'analisi per più di una disciplina — medica (anatomia), fisica (acustica), psicologica —, in

quanto scienza, non può appunto considerare il suo oggetto — i suoni — se non in rapporto a una valutazione complessiva degli aspetti dei fenomeni articolatori. Bisogna però ricordare che i suoni di cui si occupa la fonetica sono quelli (e non altri: cfr. **suono/rumore**) prodotti e percepiti da particolari organismi (gli esseri umani) i quali hanno un preciso ambito di **conoscenza**: quella *del* e *attraverso* il linguaggio (cfr. **competenza/esecuzione, senso/significato**).

Gli oggetti fondamentali della fonetica non si limitano quindi ad essere i puri dati fisici di partenza, ma sono dei concetti (cfr. **concetto**). Tenuto conto della opposizione «classica» tra suoni della *parole* — fonetica — (cfr. **parola**) e suoni della *lingua* — fonologia — (cfr. **lingua/parola**), nell'articolo in questione trattati congiuntamente, andrà osservato che l'analisi dei suoni linguistici occupa, nell'ambito dello studio della **comunicazione** linguistica, un posto ben definito, essendone, insieme ad altri, uno dei livelli costitutivi (cfr. ad esempio **grammatica, semantica, lessico**; e, per altri aspetti, **atti linguistici, discorso, enunciazione; metrica**).

In questo articolo saranno presi in considerazione i principali significati che la parola 'forma' assume in rapporto al triplice sistema di opposizioni che essa stabilisce con i termini 'sfondo', 'materia', 'spirito'.

1. Forma/sfondo: la «Gestalt».

La psicologia associazionista può essere definita da tre postulati: 1) i fenomeni psicologici complessi hanno per principio dei fenomeni psichici semplici, da cui risultano per integrazione; 2) i fenomeni psicologici complessi sono i dati immediati; per arrivare ai fenomeni psichici semplici si rendono necessarie un'inferenza e una ricostruzione astratta; 3) questi fenomeni psichici semplici sono le impressioni che s'imprimono in noi come su una *tabula rasa*; è l'abitudine, fondata sull'associazione, che le organizza; non vi sono strutture psicologiche innate.

La psicologia della forma (*Gestaltpsychologie*) oppone a questi postulati i propri: 1) esistono fenomeni psichici complessi o forme irriducibili alla somma di fenomeni psichici elementari: sono questi, al contrario, che risultano dall'analisi di quelli; 2) se le forme sono date immediatamente, è necessario aver fiducia nell'esperienza immediata; nessun processo astratto di ricostruzione permette di ridurle a dati inconsci ma più primitivi; 3) il carattere originario delle forme tende a farne proprietà innate e non acquisite del sistema nervoso.

Anzitutto si definiranno la nozione di forma e le leggi principali che la governano. Poi se ne mostrerà l'uso al livello della percezione, dell'intelligenza e della stessa fisica. Infine si accennerà alla sua influenza sulla cibernetica e l'etologia.

La forma è una totalità, le cui parti non sono legate mediante una semplice relazione di giustapposizione e di contiguità, ma obbediscono a una legge intrinseca, che è la sola capace di determinare il loro significato nella totalità. Le leggi della forma sono quindi leggi di organizzazione: 1) la forma si distingue dallo sfondo, relativamente indifferenziato, come una figura costituente un'unità. Allo stesso modo si può riconoscere la comparsa e la ricomparsa di un tema in una sinfonia. Questa unità ammette tutti i gradi di complessità, di stabilità e di chiarezza. 2) Il campo percettivo, analogamente al campo elettromagnetico, è un sistema dinamico che tende verso una struttura. Questa struttura è precisamente quella dell'organizzazione di forme. 3) Le condizioni del campo e le relazioni formali tra i suoi elementi determinano l'emergenza delle forme. Le relazioni sono quelle di somiglianza, prossimità, simmetria, chiusura, continuità di direzione.

Wertheimer [1912] ha mostrato per primo, a proposito della percezione del movimento apparente, che non si può costruire questa apparenza mediante