

ISABELLA CHIARI

Coppie minime e misura del rendimento funzionale

L'approccio quantitativo non è certo l'unico, né forse il più utile e perspicuo per lo studio della fonologia, e non è certamente il più praticato, benché, con la sola eccezione del dominio lessicale, quello fonologico sia uno dei territori della linguistica meglio rappresentato con strumenti statistici. È indubbio però che tra i diversi modi di rappresentare i sistemi fonologici quello statistico permette di osservare e distinguere finemente i gradi con cui si manifestano determinati rapporti tra sistema, norme linguistiche ed usi¹.

L'utilità degli strumenti quantitativi è però subordinata al tipo di teoria, rappresentazione, descrizione linguistica che si adopera come modello. Dal punto di vista storico, come spesso emerge in questo volume, l'uso e la elaborazione di tecniche quantitative è legato ad approcci strutturalisti e/o funzionalisti ai fenomeni linguistici, sviluppati con una certa fortuna dagli anni Trenta del Novecento fino agli anni Sessanta circa.

Sul piano della fonologia in particolar modo – ma spesso anche per altri ambiti della linguistica – i lavori quantitativi più significativi maturano con maggiore continuità nella Scuola di Praga o nella tradizione strutturale sovietica. Oltre al semplice computo delle frequenze assolute e relative dei fonemi e delle combinazioni di fonemi, utili strumenti descrittivi ed euristici, compare negli anni Trenta anche una nozione complessa e un po' impressionistica che attira l'attenzione di numerosi linguisti europei e americani negli anni a seguire: il rendimento funzionale.

L'espressione "rendimento funzionale" (soprattutto il sostantivo, variamente detto *Belastung*, *load*, *yield*, *burdening*, *carico*, *peso*, *resa*) richiama direttamente i metodi quantitativi – più per la misurazione di

¹ Una discussione critica in questa direzione è fornita anche in questo volume dal contributo di Albano Leoni e Clemente, pp. 27-44.

grandezze fisiche a dire il vero – in totale accordo con una visione dinamica, aperta, integrata e stratificata del linguaggio come quella praghese – e simile alle modellizzazioni statistiche dei processi biologici, più che in un senso puramente matematico.

Questo contributo vuole essere una discussione della nozione teorica di rendimento funzionale: in particolare saranno esaminati alcuni nuclei centrali per la determinazione quantitativa del rendimento funzionale e della nozione correlata di coppia minima. Sono esaminate alcune misure di quantificazione del rendimento funzionale applicate dagli anni Sessanta in poi ad alcune lingue moderne mettendone in luce gli aspetti salienti e critici dal punto di vista sia teorico sia metodologico. È inoltre proposta una nuova formulazione quantitativa applicabile a diversi domini linguistici. Nel §1 è brevemente presentato il significato teorico della nozione di rendimento funzionale (FL)² nel paradigma della Scuola di Praga; il §2 introduce alcune misure prodotte dagli anni Sessanta in poi per la misura del FL, mentre nel §3 si discutono alcuni problemi interpretativi relativi alla nozione di FL e di coppia minima; il §4 invece descrive i requisiti teorici necessari per l'elaborazione di diversi tipi di rendimento funzionale e ne suggerisce una formulazione matematica integrata.

1. Coppie minime e rendimento funzionale

Che cos'è il rendimento funzionale (FL)³ (*rendement fonctionnel, functional/semantic load/burdening/yield, funktionelle Belastung, функциональная нагрузка*)?

Il termine *rendimento funzionale* risulta definito per esempio nel *Projet de terminologie phonologique standardisée* (CLP, 1931: 313) come “the degree of utilization of a phonological opposition for distinguishing different meanings of words in a given language”. In questa definizione il peso o rendimento è associato direttamente all'uso di opposizioni fonologiche ed è inoltre associato alla nozione di coppia

² Si è preferito indicare come abbreviazione del termine rendimento funzionale FL (*functional load*) piuttosto che RF, poiché quest'ultima è usata sistematicamente altrove per designare la ridondanza fonematica.

³ In italiano anche detto *carico, peso* o *portata funzionale*.

minima, seppure non detto esplicitamente, poiché connesso con la possibilità di una coppia (opposizione fonologica) di distinguere significati diversi.

Probabilmente il primo a utilizzare questo termine (*funktionelle Belastung*) fu Mathesius (1931), ma la nozione ha incominciato a circolare soprattutto in seguito alla menzione che Trubetzkoy ne fa nei *Gründzuge der Phonologie* (1939). Colui che tuttavia ha dato il maggiore contributo, soprattutto nell'interpretazione diacronica, alla nozione è stato André Martinet nell'*Economie des changement phonétiques* (1955).

Pur non trattando in modo specifico del rendimento funzionale Trubetzkoy dà per scontato che si tratti di uno degli aspetti centrali per la descrizione delle regole statistiche che governano la combinazione dei fonemi:

La statistica ha in linguistica un significato duplice. Da una parte deve mostrare con quale frequenza un elemento fonologico della lingua studiata (fonema, nesso fonemico, tipo di parola o di morfema) appare nel discorso, dall'altra quale è il carico o aggravio funzionale di questo elemento o di una data opposizione fonologica. Per il primo scopo bisogna esaminare statisticamente di testi continuati, per il secondo di vocabolari (Trubetzkoy, 1939: 288)

Trubetzkoy inoltre chiarifica in cosa consista la propria idea su cosa sia il rendimento funzionale asserendo che:

a) dal punto di vista statistico il rendimento funzionale si studia osservando dei lessici (in contrapposizione ai testi “correnti” necessari per l'elaborazione statistica della frequenza delle unità fonologiche) o dizionari;

b) il rendimento funzionale è il “grado di utilizzazione distintiva delle singole opposizioni fonologiche” (*ivi*, 300), dunque il tasso di uso distintivo delle opposizioni.

Sul primo punto Trubetzkoy è lapidario. Mentre per lo studio delle caratteristiche dei fonemi e delle loro combinazioni auspica e sottolinea incessantemente il ricorso a testi rappresentativi del discorso⁴, lo

⁴ L'utilità di dati testuali (e non sistemici come quelli derivabili per certi versi dai lessici) per esempio è rimarcata da Trubetzkoy per determinare i diversi gradi d'uso nel discorso di raggruppamenti di fonemi, al fine di osservare la discrepanza tra valori derivanti dalle possibilità teoriche di combinazione dei fonemi e effettive realizzazio-

studio del rendimento funzionale deve procedere a livello di sistema e dunque prendere in considerazione lessici (*Wörterbücher*, bisogna dedurre nel senso di repertori di forme astratte – non calate nei loro contesti, non flesse, ecc.). È come se in qualche modo la nozione di funzione per il linguista russo si calasse dall'astratto al concreto, rimanendo tuttavia su un piano sistemico.

Ma questa rappresentazione sistemica del rendimento funzionale è indicativa davvero di quello che lo stesso Trubetskoy definisce come FL ossia il tasso d'uso distintivo delle opposizioni? Il termine *uso* [*Ausnützung*] spiazza il lettore di oggi perché solitamente associato al piano concreto e testuale della lingua. Ci si chiede dunque se la posizione “sistemica” del rendimento funzionale non sia quantomeno discutibile.

L'idea della distintività delle opposizioni richiama inoltre un'altra nozione altrettanto complessa e discutibile, ossia quella di coppia minima, anch'essa come si vedrà in bilico tra piano sistemico (di lessico, direbbe Trubetskoy) e piano testuale.

Dal punto di vista storiografico, una delle prime attrattive che ha spinto all'approfondimento della nozione di rendimento funzionale è il suo potenziale per la spiegazione e interpretazione del mutamento fonetico, in piena consonanza con il suo trattamento da parte di Martinet. Martinet riprende una osservazione *en passant* di Jakobson che suggerisce come un basso rendimento funzionale possa indicare una direzione diacronica verso la perdita di una data opposizione fonologica (cfr. Jakobson, 1931: 259). Martinet infatti più volte insiste nell'*Economia dei mutamenti fonetici* (1955) sul rapporto tra FL e possibilità per una coppia di opposizioni di perdersi, unendo le classi delle realizzazioni relative a diversi fonemi nel momento in cui l'opposizione non contribuisca più a definire numerose coppie di parole quasi-omofone. Tale considerazione, apparentemente intuitiva, ha per molto tempo costituito l'unica funzione, l'unica indicazione, l'unica attribuzione del rendimento funzionale⁵.

ni di tali possibilità nel vocabolario. Si tratta di quella che oggi chiameremmo *ridondanza fonologica (distribuzionale) combinatoria* (cfr. Chiari, 2002: 207-51).

⁵ Un esplicito contributo di Martinet a questo proposito riguarda la nozione di equilibrio dei sistemi fonologici, collegata direttamente all'elaborazione della nozione di economia linguistica. Non si entrerà qui in dettaglio nel tema, molto rilevante per la teoria linguistica del mutamento fonetico, e punto critico che è costato a Martinet

Martinet dedica qualche pagina alla discussione del possibile ruolo del rendimento funzionale in diacronia, riconoscendone immediatamente la difficile definibilità e l'assenza di accordo sul ruolo della nozione. Vi riconosce un uso "semplice e ingenuo" per cui il FL coinciderebbe con il numero delle coppie minime presenti nel lessico di una lingua relativamente a due fonemi in opposizione. A questa accezione Martinet oppone le prime utili critiche che riguardano soprattutto l'eventuale ruolo della coppia minima come sorgente di fraintendimento, ruolo giocato solo a determinate condizioni sia di frequenza dei due membri della coppia, sia di appartenenza grammaticale (che determina la coesione del termine in questione con il suo cotesto), ma su questo si tornerà in seguito.

Inoltre anche Martinet come Trubetskoy sottolinea in modo chiaro come il rendimento funzionale debba essere collegato al livello sistemico e dunque essere computato (mediante le coppie minime) in un lessico, inteso come "lista di parole" (*ivi*, 49). Martinet è forse sedotto dall'idea di tenere in conto le frequenze dei fonemi, tanto da suggerire che i fonemi frequenti hanno maggiore probabilità di svolgere funzioni distintive, tuttavia conclude che per la valutazione del FL:

La frequenza nel lessico è spesso preferibile alla frequenza reale nei testi o nel discorso, poiché non è affatto eccezionale che un fonema come quello francese /œ/, che fa parte del paio della differenza minima /ẽ/-/œ/ con una resa funzionale debolissima, compaia con una certa frequenza nei testi parlati o scritti perché si trova in una parola di grande frequenza con l'articolo indeterminativo *un* (Martinet, 1955: 48).

Un'osservazione quantomeno ambivalente nelle interpretazioni. Lo stesso fatto linguistico che riguarda *un* infatti ci potrebbe spingere proprio a tenere in considerazione il fattore testuale nella valutazione del rendimento funzionale, piuttosto che muoverci a escluderlo. Se una coppia è mantenuta da singole parole o insiemi di parole frequenti intuitivamente non ne aumenta anche il peso funzionale? Questo è

l'accusa di fornire una rappresentazione teleologica dei fatti linguistici. Per una prima panoramica legata alla problematicità della nozione di economia linguistica si vedano Chiari (2002: 77-91), e già Peeters (1994; 1997) in cui si mette in evidenza la parentela del cosiddetto teleologismo martinettiano con la teleologia della Scuola di Praga, legame che progressivamente scompare dai lavori successivi di Martinet.

chiaramente uno dei punti cruciali dell'interpretazione del FL che sarà ripreso in seguito.

In modo parallelo e completamente indipendente Zipf, seppure in maniera più abbozzata, aveva proposto qualche anno prima l'idea che le soglie di frequenza potessero giocare un ruolo significativo nell'evoluzione dei sistemi fonologici (cfr. soprattutto Zipf, 1935). Le implicazioni filosofico-biologiche del lavoro anche successivo di Zipf (1949) e alcune sue ingenuità linguistiche e matematiche hanno tuttavia fortemente limitato la diffusione del suo lavoro, fino alla ripresa di alcuni aspetti da parte di Pierre Guiraud (in particolare, 1959)⁶. Similmente a Guiraud, sempre con dati latini, anche Albert Maniet (1993), più recentemente, partendo proprio dalle formulazioni martinetiane, tenta una valutazione del rilievo diacronico della nozione di rendimento funzionale, giungendo a risultati generalmente negativi⁷.

Mańczak (1970) propone addirittura di evitare il tentativo di chiarire la nozione di rendimento funzionale e di sostituirla con la nozione di frequenza d'uso, a suo parere più chiara⁸.

⁶ La caduta delle nasali finali e la sonorizzazione della occlusive intervocaliche in latino sono per esempio attribuite da Guiraud a fenomeni di riequilibrio del sistema fonologico in conseguenza di scarti anomali dalla frequenza media. Si tratta solo di un'esemplificazione di come è stata usata la correlazione tra frequenza e uso delle unità fonologiche per proporre interpretazioni diacroniche. Nel caso specifico, l'interpretazione di Guiraud non sembra particolarmente plausibile, né pare fondata metodologicamente, poiché lo scarto dalla frequenza relativa media è calcolato su un insieme di una ventina di lingue, la cui scelta non ci sembra sufficientemente motivata, e poiché in generale il confronto tra sistemi fonologici è particolarmente problematico.

⁷ Il tema del valore diacronico del rendimento funzionale è qui solamente accennato. Si segnala unicamente che uno dei problemi più gravi che derivano dall'applicazione della nozione di rendimento funzionale per spiegare i mutamenti diacronici risiede nel renderlo "la" condizione per il mutamento e non solamente "una" delle condizioni (cfr. Vachek, 1969; Maniet, 1993).

⁸ Che la frequenza assoluta o relativa delle unità fonologiche nella lingua e nei testi siano fondamentali per la rappresentazione del livello fonologico delle lingue è fatto assodato. Che la nozione di frequenza d'uso sia chiara e univoca tuttavia non è altrettanto condiviso, né risulta scontata la metodologia per quantificare tale frequenza d'uso. In effetti la determinazione del tasso d'uso delle unità ci riporta alle stesse questioni di relazione, opposizione e distinzione di livelli poste dalla nozione di rendimento funzionale. Peraltro dirimere la questione del computo della frequenza è comunque un prerequisito per qualsiasi discussione sulla costituzione funzionale dei si-

Nell'ottica appena delineata non si fa accenno al modo di valutare in maniera quantitativa la nozione, e solitamente, anche nei testi di tipo manualistico, si asserisce quanto basti la presenza sistemica anche di una sola coppia minima per dover asserire che una data opposizione fonologica è effettivamente presente e usata in una lingua. Questa è la posizione espressa per esempio da Toegeby (1951: 67). Sembra necessario tuttavia cercare un modo di valutare più finemente le diverse gradazioni che vanno da un alto a un basso rendimento funzionale, capire quali sono i fattori linguistici che entrano in gioco nella sua determinazione e anche esaminare come questa nozione possa mostrare lo stretto e a volte indeterminato legame tra asse paradigmatico e asse sintagmatico di una lingua.

2. Alcune proposte per la misura del rendimento funzionale

In una prospettiva funzionale la nozione di FL, se congiunta a quella di coppia minima, può fornire un quadro articolato delle caratteristiche di un sistema fonologico e del suo uso. La proficuità della nozione però è soprattutto dipendente dalla possibilità di associare ai termini generici *alto e basso rendimento funzionale* un qualche corrispettivo quantitativo. Il compito è spinoso e multiforme tanto da spingere a John Lyons ad affermare che “the precise quantification of functional load is complicated, if not made absolutely impossible” (Lyons, 1968: 82).

I primi scritti in cui si tenta di elaborare una nozione più precisa e matematicamente fondata della nozione di rendimento funzionale risalgono ai primi anni Sessanta. Cronologicamente e teoricamente l'avvio alla discussione è dato dal lavoro di Hockett, che nel *Manual of Phonology* (1955) inizia a riflettere sulla nozione fino a proporre nel 1966 una forma di quantificazione (cfr. Hockett, 1967).

Dal punto di vista del modello usato per l'elaborazione della misura del rendimento funzionale si distinguono i metodi ispirati alla teoria dell'informazione come il primo dal punto di vista cronologico di Ho-

stemi fonologici, sia che si sposti la questione sulla nozione di uso, sia che la sia conservi in termini di rendimento funzionale.

ckett (1955, 1967), Kučera (1963) e Wang (1967) e metodi puramente statistici come quelli proposti da Rischel (1962) o King (1967)⁹.

Hockett correla la nozione di rendimento funzionale con quella di *perdita di entropia* del sistema. Essenzialmente la proposta di misura di Hockett (1955; 1967) si basa sull'idea che il rendimento funzionale di una coppia di fonemi $\{a, b\}$ potrebbe essere pensato come una sottrazione tra l'entropia H del sistema quando a e b sono distinti e l'entropia H^* che il sistema avrebbe se i fonemi a e b non si distinguessero più rientrando in una unica classe, e lasciando il resto del sistema immutato. Dunque:

$$[1] \quad FL^{Hockett} = (H - H^*) / H$$

Poiché la misura di H ed H^* possono essere piuttosto elaborate, Hockett propone di sostituirle con le corrispettive approssimazioni di secondo ordine¹⁰. Concepito in tal modo il FL può anche essere interpretato come “the extent of homonymy that is caused by the obliteration of contrasts” (Wang, 1967: 49).

Come per Hockett, anche per Wang il FL è essenzialmente computato in relazione (corrispondenza) con l'entropia del sistema fonologico, dunque sommando i contrasti tra tutte le opposizioni si deve ottenere il valore dell'entropia H . Wang (1967: 43-44) definisce, sulla scorta di Hockett, alcune condizioni matematiche che devono essere rappresentate dalla formula per la misura del FL come conseguenza di alcune osservazioni teoriche sulla nozione. In particolare Wang osserva che la misura del FL deve essere zero se il sistema fonologico (teorico) contiene un solo fonema; una opposizione di fonemi in distribuzione complementare deve avere $FL = 0$, dato che ovviamente l'opposizione non è mai usata per tenere distinte delle coppie minime.

In modo simile Kučera (1963: 209-217) adopera il quadro della teoria dell'informazione per giungere a una quantificazione del FL, u-

⁹ Per la lingua italiana l'unica indicazione, di natura puramente critica, senza proposte o computi operativi tuttavia, è quella di Francescato (1962) in particolare sul ruolo del basso FL in diacronia per alcuni casi della lingua italiana (/e:/:ε/ e /o:/:ɔ/).

¹⁰ Per una esemplificazione chiara di cosa significa una approssimazione (di primo ordine, secondo ordine, ecc.) si vedano Shannon e Weaver (1949) e Miller e Friedman (1957).

sando, in maniera interessante, come “configurazione fonemica” per il FL la sillaba fonologica e come unità di analisi i tratti distintivi.

In una prospettiva molto simile, usando però non la nozione di entropia, bensì il computo delle coppie minime, Rischel (1962) propone due diverse misure di FL, l’una per il lessico l’altra per il testo. Come si misura il FL testuale? Per Rischel basta integrare i dati sulle frequenze testuali dei membri della coppia minima. Dunque il FL di una coppia oppositiva di fonemi $\{x, y\}$ è dato dalla somma di tutti i prodotti delle frequenze di occorrenza delle parole in coppia minima. Se $f_i(P^x)$ è la frequenza di occorrenza testuale della prima parola della coppia che contiene il fonema x , $f_i(P^y)$ è la frequenza di occorrenza della seconda parola che contiene il fonema y , allora n è il numero di diverse coppie minime registrate nel lessico (dunque per Rischel è il FL del lessico)

Il FL testuale secondo Rischel corrisponde a [2] per esteso e in sintesi a [3]¹¹:

$$[2] \quad FL_{testuale}^{Rischel} = f_1(P^x)f_1(P^y) + f_2(P^x)f_2(P^y) + \dots + f_n(P^x)f_n(P^y)$$

$$[3] \quad FL_{testuale}^{Rischel} = \sum_{i=1}^n f_i^x f_i^y$$

Per Rischel quindi il FL del lessico è semplicemente misurato mediante il computo di tutte le coppie minime della coppia oppositiva, il FL testuale, aggiungendovi le indicazioni sulla frequenza delle parole della coppia.

In aperta opposizione Robert D. King (1967) critica aspramente l’utilità della nozione di rendimento funzionale come fattore del cambiamento diacronico, sulla base di uno studio sperimentale su campioni fonologici di diverse epoche, giungendo alla conclusione che non ci sono dati empirici validi per considerare la nozione di FL come indicatore di qualsivoglia direzione nel mutamento fonologico. C’è da dire tuttavia che, come prontamente registrano Hockett (1967: 320) e Vachek (1969: 64), sul piano della validità dei risultati, l’analisi di King

¹¹ Per chiarezza espositiva la notazione usata da Rischel (1962) è stata modificata rispetto all’originale, non ovviamente nella formulazione matematica, ma nella scelta dei simboli.

presenta una carenza tutt'altro che trascurabile, ossia il fatto che i campioni di fonemi usati per registrare ciascun stato di lingua sono insufficienti dal punto di vista statistico ammontando a soli 20.000 fonemi¹².

King, metodologicamente, rifiuta l'approccio informativo di Hockett, sostituendolo con un metodo statistico-strutturale. La proposta originale di King è particolarmente interessante poiché integra nel FL anche dati riguardanti la frequenza relativa dei fonemi analizzati, tuttavia si basa solamente su co-testi immediati (un fonema prima e uno dopo quello analizzato), non tiene conto delle coppie minime, né della posizione del fonema.

La formula semplificata di King (1967: 4) è:

$$[4] \quad FL^{King} = L(x_i, x_j) = (1/N^2) \sum_{k=1}^m f_{ik} f_{jk}$$

in cui L è il rendimento funzionale della coppia di fonemi (x_i e x_j), in cui N è il numero di fonemi dell'inventario e f_{ik} è la frequenza di occorrenza di x_i un dato co-testo (da 1 a m sono i co-testi possibili). La formula di King così com'è può essere applicata sia a lessici sia a input testuali, con ovvie differenze nell'interpretazione.

3. Coppie minime e produttività: problemi di interpretazione e di trattamento

Cos'è che fa problema nella possibile identificazione e definizione del FL? Perché è così difficile coglierlo in maniera non banale e utilmente critica? In realtà come si è visto nei paragrafi precedenti sono diversi i nodi teorici e metodologici da sciogliere per giungere a una formulazione matematica utile e densa dal punto di vista teorico. In particola-

¹² Effettivamente soprattutto per considerazioni quali la valutazione del FL, che coinvolge dunque l'analisi di sequenze fonologiche (i fonemi nel loro ambiente fonologico contestuale), il campione di King è davvero molto piccolo. Dato che la lunghezza media delle parole inglesi è di circa 4,5 fonemi, tale campione testuale raccoglie solamente circa 4.500 parole *tokens*, pochissime per valutare la posizionalità e la presenza di contesti fonologici comparabili. Questo dato tuttavia inficia solamente la validità dei risultati sulla non rilevanza del FL nell'evoluzione diacronica, non la proposta di quantificazione di King.

re, non è tanto il concetto di *rendimento* a far problema – semmai a quel proposito si toccheranno solamente questioni metodologiche a posteriori – quanto il suo attributo *funzionale*. Cosa significa inoltre dare una misura o una rappresentazione quantitativa della funzione che gioca un fonema in un sistema fonologico?

In primo luogo, è necessario cogliere da una lato la *relazionalità* della nozione di funzione, ossia il gioco di forza che si stabilisce tra una unità funzionale e le altre del sistema, e tra l'unità e quelle che la precedono e la seguono. In secondo luogo è utile capire in che modo la funzione si manifesta nel sistema e in modo molto diverso nei testi concretamente prodotti.

Se l'individuazione delle coppie minime è, se non altro, uno degli indicatori del rendimento funzionale, un passo appena antecedente alla determinazione di quest'ultimo – metodologicamente e teoricamente – deve essere appunto l'esplorazione della nozione di coppia minima sia nelle sue insidie interpretative sia nella sua quantificazione in una data lingua. Ma la nozione di coppia minima non presenta poche difficoltà dal punto di vista sia teorico sia applicativo.

Innanzitutto è necessario distinguere le vere e proprie coppie minime dalle coppie dette contrastive. Anche le coppie contrastive – coppie di parole distinte per un fonema nella stessa posizione, ma mediante più di un tratto distintivo (/pane/~/rane/) – sono rilevanti per la determinazione del peso funzionale di una distinzione fonologica, per quanto in maniera un po' diversa rispetto alle coppie minime. È necessario dunque tenere in conto entrambe le possibili situazioni (la coppia minima essendo un caso particolare di coppia contrastiva). Ma questo ci interessa meno, poiché è evidente che si dovranno prendere in considerazione tutti i tipi di opposizione e dunque non strettamente le coppie minime in senso tecnico.

Un secondo punto cruciale riguarda il luogo, per così dire, in cui andare a reperire le coppie minime: il lessico potenziale rappresentato per esempio da un dizionario oppure un corpus di testi concreti? Il tema ricalca da vicino la questione già emersa per il FL: le coppie minime sono coppie paradigmatiche o piuttosto potenziali coppie sintagmatiche?

La differenza non è per nulla da poco. Se si guarda un lessico (un *Wörterbuch* come dice Trubetskoy) cosa si deve osservare: le forme di

citazione, le forme testuali? e analizzare dunque non solamente i lemmi registrati, ma anche tutte le possibili flessioni e coniugazioni di ciascun lessema?¹³ E in quest'ultimo caso ogni coppia individuata avrebbe lo stesso peso sia che le coppie siano formate da parole molto frequenti, sia che si formino con parole rare, o settoriali poco diffuse, sia che siano formate da coppie con frequenza disomogenea?

Ma il dato di frequenza non conta forse qualcosa? Le nozioni di coppia minima e di rendimento funzionale non hanno solamente un valore strutturale astratto, ma implicano presumibilmente anche un certo significato nella competenza implicita dei parlanti¹⁴. Competenza di cui gli utenti di una lingua si servono nei quotidiani compiti di discriminazione e individuazione fonologica fine. Se la nozione di coppia minima detiene anche una funzione di questo tipo sarà allora necessario combinare se non altro a un'analisi sul lessico un'analisi su corpora testuali che permettano di mettere in evidenza la effettiva produttività delle coppie minime.

In questa prospettiva oltre alla semplice analisi delle coppie minime testuali sarà opportuno analizzare e distinguere il ruolo delle parole messe in coppia. Dal punto di vista della discriminabilità e della possibile confusione degli elementi della coppia sarebbe utile osservare anche la categoria grammaticale e la classe flessiva o la coniugazione dei membri, come è stato opportunamente osservato da Avram (1964).

Potremmo dunque distinguere *coppie minime assolute* (coppie in cui le parole appartengono alla stessa categoria e alla stessa classe come /kane/~/pane/) e *coppie minime relative* (come /karta/~/parta/,

¹³ Non si giustifica la scelta della sola forma di citazione, poiché scelta per pura convenzione. Altra questione è eventualmente il problema di quale forma abbiano le parole nel lessico mentale dell'individuo.

¹⁴ Questa è per esempio la funzione che al FL viene ascritta da Lyons (1968: 82) che afferma: "The importance of functional load is obvious. Misunderstanding will tend to occur if the speakers of a language do not consistently maintain those contrasts which serve to distinguish utterances which differ in meaning". Maggiore è il rendimento funzionale, maggiore per Lyons sarà la possibilità per un parlante di includere l'opposizione nei suoi *speech habits*. Forse, benché si riconosca una qualche funzione a questo livello al FL, è necessario altresì riconoscere che nell'atto della ricezione sono molti altri i fattori determinanti di incomprensione e/o *malentendu*.

ma anche solo /pane/~/rane/, che pur essendo entrambi nomi, sono l'uno al singolare l'altro al plurale)¹⁵.

Ma a questo si aggiunge un'indeterminabile ulteriore difficoltà che deriva dal fatto che il co-testo spesso anche da solo separa i membri di una coppia minima, anche assoluta. Quante volte potremmo confondere le due frasi *qui occorrerebbe controllare i dati* con *qui occorrerebbe controllare i dadi*?

In modo del tutto analogo anche la definizione di rendimento funzionale tradizionale, come si è visto sopra anche per Trubetskoy e Martinet, sembra doversi calare e rilevare nel lessico potenziale, come terreno di reperimento delle coppie minime¹⁶. La questione risulta particolarmente confusa tanto che le posizioni di chi si è occupato di FL dopo Trubetskoy incominciano a divergere dalla posizione originaria.

Rischel (1962) per esempio sottolinea solamente come il FL possa essere computato sia in testi sia in lessici in modo del tutto analogo; Herdan (1958: 10) esplicitamente minimizza la differenza quantitativa tra dati di provenienza lessicale e testuale, asserendo invece che: "what is remarkable is the close similarity between the phoneme distribution for dictionary material and for running texts", che ci permette di valutare a suo parere il rendimento funzionale di testi e di estenderne le conclusioni al lessico¹⁷; Palau Marti (1968: 30) sostiene

¹⁵ Anche la posizione dei fonemi nelle coppie minime, il loro presentarsi in più luoghi e il loro essere eventualmente neutralizzati fornisce indicazioni che dovrebbero essere tenute in considerazione per la valutazione quantitativa del rendimento funzionale (cfr. Avram, 1964: 44).

¹⁶ Per quanto riguarda invece il computo statistico delle coppie minime bisogna notare come un computo completo non sia stato eseguito per nessuna lingua. In tutto il volume bibliografico di Köhler (1995) non risultano saggi o volumi che esplicitamente trattino questo aspetto. Le osservazioni sulle coppie minime sono solamente indiziarie e casuali, usate per esemplificare le opposizioni fonologiche. L'impresa oggi non pare particolarmente difficile con le strumentazioni informatiche, posto che si convenga su un determinato inventario fonologico.

¹⁷ Posto anche che i dati in possesso di Herdan davvero permettano tale conclusione, bisogna dire che non sono certo estensibili a tutte le lingue. Herdan infatti osserva le frequenze raggruppate per modo e, separatamente, per luogo di articolazione (non dei singoli fonemi) in un dizionario e in alcuni testi inglesi, constatandone la relativa stabilità e basandosi su un noto lavoro di Krámský. Se anche l'operazione fosse corretta metodologicamente, cosa quantomeno dubbia, bisognerebbe concludere che tale stabilità è puramente casuale, poiché in moltissime altre lingue (specialmente lin-

che possa bastare aggiungere alle forme di citazione del dizionario tutte le forme flesse relative; Avram (1964) propone invece di combinare e integrare i due dati.

Axhmanova (1960: 315) rappresenta la posizione opposta e asserisce che: «La différenciation réelle des énoncés ne se base aucunement sur les oppositions phonologiques minimales, mais sur des oppositions nombreuses et variées. Du point de vue phonologique, les énoncés réels différents sont très éloignés l'un de l'autre». Sempre a favore di una posizione eminentemente « testualista » è anche Wang (1967) che in particolare suggerisce di usare i dati sul FL testuale per trarre inferenze sul valore delle opposizioni nella lingua.

Dal punto di vista terminologico è quindi necessario differenziare le nozioni. Una prima proposta è dunque di considerare il *rendimento funzionale* come il prodotto tra due diverse misure: una misura della produttività funzionale sistemica e una misura del peso funzionale a livello testuale. Sarà chiamata dunque *produttività funzionale sistemica* la misura delle opposizioni fonologiche in un sistema fonologico indipendentemente dalla frequenza reale di occorrenza nei testi. Sarà chiamato invece *peso funzionale* il dato che tiene conto della presenza delle opposizioni nei testi effettivamente prodotti (a livello fonologico e non fonetico).

4. Considerazioni metodologiche per la misura del rendimento funzionale

Dal punto di vista puramente tecnico si riassumeranno di seguito alcune considerazioni generali utili per comprendere la formulazione quantitativa del rendimento funzionale che sarà illustrata in seguito. La proposta quantitativa presentata può essere considerata come una variazione articolata delle proposte tecniche di Rischel (1962) e King (1967). Molte delle nozioni sono espresse per variabili e non in termini specificatamente matematici. I dettagli dell'analisi devono infatti adattarsi alle caratteristiche del singolo sistema fonologico. La proposta è dunque più teorico-metodologica che statistico-matematica.

gue con morfologia flessiva particolarmente ricca) le cose stanno in modo molto diverso.

Il rendimento funzionale può essere considerato come un indice globale, che possa essere anche scorporato nelle sue componenti minime, della distintività funzionale e semantica di una coppia di fonemi o di fasci di tratti all'interno di un sistema fonologico dato. Una rappresentazione generale della proposta di analisi del FL fornita è data nella Figura 1.

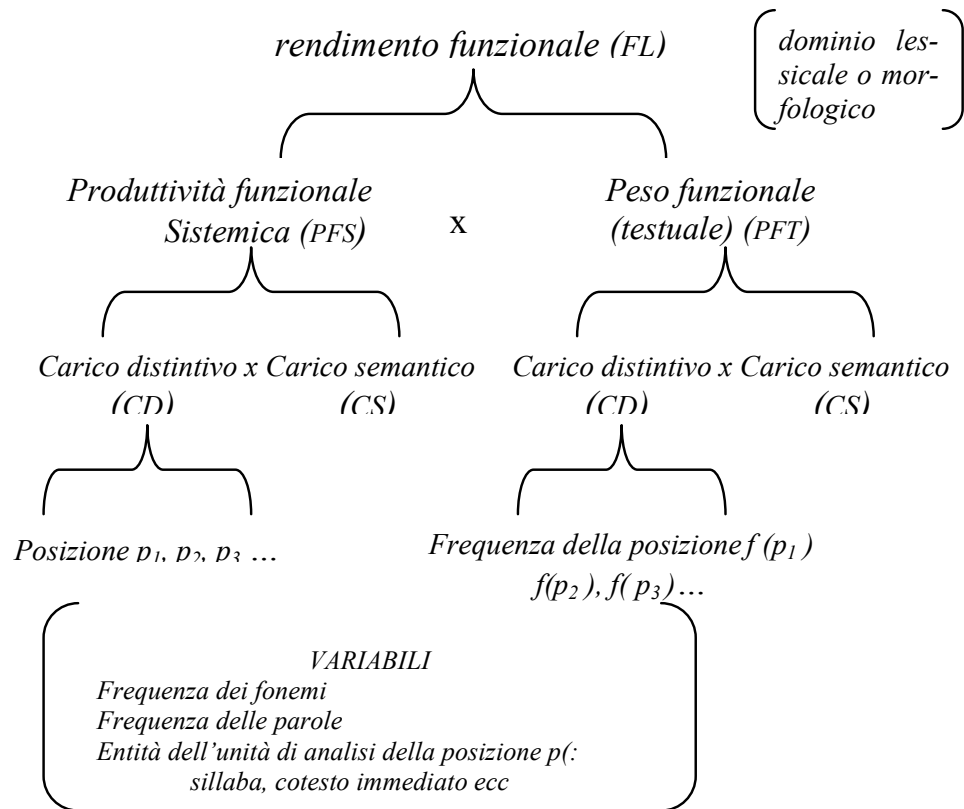


Figura 1 - Schema delle componenti del FL

Le componenti da tenere in considerazione dovranno essere almeno:

- 0) *Unità di analisi*: opposizioni fonologiche, tratti fonologici
 - a) *Dominio*: lessicale o morfologico
 - b) *Tipologia del peso*:
 - carico distintivo (per co-testi)
 - carico semantico (per coppie minime)
 - c) *Prospettiva del peso*:
 1. sistemica o lessicale (*produttività funzionale sistemica*)
 2. testuale (*peso funzionale*)

Il problema dell'unità di analisi (punto 0) in realtà più che una questione di tipo metodologico richiede semplicemente la specificazione del tipo di unità fonologiche messe in correlazione e confrontate. Se si osserva un sistema fonologico dal punto di vista delle dinamiche di forza distintiva assunte dai vari elementi infatti possiamo voler osservare di volta in volta le singole opposizioni fonologiche, oppure il valore e la produttività distintiva di singoli tratti (come sonorità/sordità, per esempio). Il contributo maggiore su questo punto è forse quello di Surendran e Niyogi (2004) in cui si tenta l'applicazione della nozione di FL non solamente coppie di fonemi in opposizione, ma anche oltre ai tratti distintivi, e soprattutto aspetti soprasegmentali quali i toni e l'accento, idea questa già lanciata da altri (cfr. Greenberg, 1959; Hockett, 1967).

Al punto a) si definisce la necessità di individuare il DOMINIO cui è sensato applicare l'analisi del FL. In genere, anche in considerazione del ruolo cruciale giocato dai fenomeni di neutralizzazione, la scelta è tra il dominio lessicale, in cui i fenomeni saranno classificati a seconda delle opposizioni create entro la parola, e il dominio morfologico, in cui le opposizioni sono osservate per la posizione che occupano entro il morfo. Vi sono lingue in cui è più utile scegliere un dominio lessicale (per esempio l'inglese), altre per cui è più utile quello morfologico (tipicamente il tedesco).

Al punto b) invece si determina la prima distinzione significativa dal punto di vista teorico e metodologico, ossia la distinzione della capacità che ha una opposizione di distinguere significanti. Da una parte il rendimento funzionale deve rappresentare la reale capacità distintiva di una coppia di opposizioni fonologiche nel creare parole con significati diversi (dunque coppie minime), tenendo conto anche eventualmente di variabili di tipo grammaticale. Tale capacità si dirà *carico semantico* (CS) dell'opposizione. Dall'altra il rendimento funzionale deve rappresentare anche la capacità distintiva funzionale tout court dell'opposizione, tenendo questa volta in conto non la parola intera ma i co-testi dell'unità presa in esame (la posizione dunque del fonema, poniamo, nella catena fonologica). Chiameremo questa capacità *carico distintivo* (CD) dell'opposizione.

Il rendimento funzionale, sia in termini di produttività funzionale sistemica sia di peso funzionale testuale, dovrà dunque combinare entrambi i fattori, quello puramente posizionale (in qualche modo anche potenziale) e il più rilevante aspetto della distintività semantica.

La nozione appena introdotta di *carico distintivo* dell'opposizione pone inoltre alcuni problemi di carattere metodologico non indifferenti: quanto grande deve essere il co-testo fonologico analizzato (1 fonema prima e 1 dopo, posizione e ruolo nella sillaba, ecc.), che chiameremo *U* (con *U* che va da 1 a *m*)?¹⁸ *U* (unità di co-testo fonologico) deve essere lo stesso ambiente per entrambi i fonemi da opporre¹⁹. Per quanto riguarda l'ampiezza del campione di fonemi da considerare per la misura del rendimento funzionale è necessario premettere che essa dipende dall'unità di analisi presa in considerazione (*U*). Se si considerassero solamente i co-testi brevi (*U*) basterebbero relativamente pochi fonemi, tuttavia, dovendo includere nel computo anche il carico semantico, che si serve delle coppie minime, si rende necessario prendere in considerazione un campione molto più ampio²⁰.

Da tali decisioni cruciali discende il senso che si può dare a carico funzionale e inoltre l'ulteriore problema della rappresentazione del carico distintivo di una opposizione in posizioni diverse a livello di parola o di morfema. Le diverse posizioni individuabili entro *U* saranno dette $p_1, p_2 \dots p_n$, in cui *n* è il numero delle posizioni considerate.

Dal punto di vista metodologico bisognerà rispondere alle domande: che tipo di analisi per posizione è necessario fare? Come si possono combinare i dati delle diverse posizioni per generare un indice globale del carico distintivo che renda bene quantitativamente il ruolo delle neutralizzazioni e delle diverse frequenze per posizione? Oltre alla posizionalità *tout court* è inoltre necessario tenere in conto della differenza nelle capacità discriminative di determinati fonemi soprattutto in relazione alla posizione tonica o atona.

La produttività funzionale sistemica e il peso funzionale testuale si configurano entrambi come il prodotto tra carico distintivo e carico

¹⁸ *m* sarà dunque sempre inferiore a $(N+1)^k$, se *N* è il numero dei fonemi, 1 indica il confine dell'unità (della parola, del morfo a seconda della *U* che si è scelta), e *k* indica la lunghezza in numero di fonemi di *U*.

¹⁹ Per *stesso ambiente* si userà la definizione di Fischer-Jørgensen (1956: 146): "Same environments' must mean environments which have been proved to be functionally identical, or which we, at the given stage of the analysis, considers functionally identical, so that they cannot be made responsible for the difference of content entailed by the replacement".

²⁰ King (1967) suggerisce per esempio una base di circa 28.830 parole per indagare un sistema a 30 unità circa, ma su questo occorrerebbe riflettere con maggior cautela.

semantico. Il carico distintivo (CD) ha tra le sue variabili il numero delle posizioni analizzate (p), uguale per tutto il sistema e dunque costante per la lingua analizzata, e si configura come il prodotto dei valori ottenuti per ogni p . Un p uguale a zero equivale a una posizione neutralizzata (ridotta per non produrre un CD uguale anch'esso a zero a un valore bassissimo convenzionale). Nella valutazione del FL una opposizione che ha una posizione p neutralizzata deve a livello testuale avere un corrispettivo che permetta di valutare quanti casi testuali presentano la posizione neutralizzata, in modo da abbassare il FL complessivo; il carico semantico CS obbligatoriamente più basso nella quantificazione grezza deve essere moltiplicato per una costante (a) che lo faccia pesare più del CD.

La distinzione della posizione in relazione ad altre unità serve infatti a tenere in conto, nella valutazione del FL, anche della possibilità che determinate opposizioni siano neutralizzate in certi contesti. Per ogni posizione infatti è necessario valutare il carico distintivo per poi combinare per ogni coppia i dati raccolti in ogni posizione possibile.

Dunque per determinare il carico distintivo di una coppia si dovrà determinare per ogni co-testo individuato entro U , di una coppia di fonemi x e y , la somma dei prodotti delle frequenze con cui ciascuno dei due fonemi compare in tutte le posizioni p (da 1 a n):

$$[5] \quad CD_{U_j} = f(p_{1x} + a)f(p_{1y} + a) + f(p_{2x} + a)f(p_{2y} + a) + \dots + f(p_{nx})f(p_{ny})$$

Oppure sinteticamente:

$$[6] \quad CD_{U_j} = \sum_{k=1}^n f(p_{kx} + a)f(p_{ky} + a) = f_{U_j,xy}$$

Il carico distintivo descritto in [6] andrà successivamente sommato a tutti i carichi distintivi delle altre unità di cotesto (U_j) che variano da un minimo di $j=1$ a un massimo di $j=m$. E dunque il carico distintivo della coppia di fonemi sarà uguale a:

$$[7] \quad CD_{U_{xy}} = \sum_{j=1}^m f_{U_j,xy}$$

La formula in [7] varrà ovviamente anche per calcolare il CS, con la particolarità che la dimensione di U equivarrà alla parola e potrà non essere presa in considerazione la posizionalità.

CD e CS devono essere computati su due diversi “corpora”:

- a) per ottenere la *produttività funzionale sistemica*, si partirà da un dizionario con forme flesse (per le lingue a ricca flessione

come l'italiano) o in altri casi su un dizionario delle radici morfologiche;

- b) per ottenere il *peso funzionale testuale* si partirà invece da un corpus vero e proprio costituito da testi trascritti fonologicamente sufficientemente estesi.

Il prodotto di CD e CS per la *produttività funzionale sistemica* e il prodotto di CD e CS per il *peso funzionale testuale* saranno allora da integrare, per produrre l'effettivo rendimento funzionale (FL) complessivo per l'opposizione fonologica analizzata.

Una strutturazione come quella proposta per il FL dovrebbe produrre le seguenti conseguenze generali:

1) Più una opposizione fonologica è ripartita, ossia rappresentata in diverse posizioni (quindi più è indipendente dal contesto fonologico) più il rendimento funzionale è alto; e dunque più sono le posizioni di neutralizzazione, minore è il rendimento funzionale dell'opposizione (cfr. Avram, 1964: 44)²¹;

2) Se un'opposizione con una media produttività funzionale sistemica ha invece peso funzionale testuale molto basso, il rendimento funzionale sarà basso; così come una opposizione con media o bassa produttività funzionale sistemica e con peso funzionale testuale alto avrà un FL alto: di conseguenza il peso funzionale è più rilevante della produttività sistemica per la determinazione del FL;

3) La misura del rendimento funzionale deve anche tenere in conto della frequenza relativa dei due membri della coppia fonologica: se questi sono entrambi frequenti anche in altre posizioni infatti il rendimento deve essere considerato maggiore, se entrambi rari, il rendimento deve essere considerato basso, se la frequenza di uno solo è alta allora il rendimento funzionale deve essere considerato più basso di quello di una coppia in cui entrambi i fonemi sono frequenti in tutte le posizioni (di questo tiene conto anche la formula proposta da King, 1967).

²¹ La ripartizione delle posizioni entra nella valutazione del carico distintivo (CD) e indirettamente anche del carico semantico (CS).

Conclusioni

La relazione – eventuale – tra rendimento funzionale e capacità distintiva va certamente indagata solamente tenendo conto del legame tra fonologia e realizzazioni fonetiche, dunque ponendo come centrale il delicato tema dell'interfaccia fonetica/fonologia. La valutazione del rendimento funzionale serve dunque per dare un'indicazione sull'interrelazione tra asse paradigmatico e sintagmatico per le sequenze fonologiche, ma non esaurisce certamente la questione dell'uso fonologico, né tanto meno quella del mutamento fonetico.

La valutazione del ruolo che il rendimento funzionale svolge nei processi di costruzione e ricostruzione del significante e di comprensione necessita della considerazione congiunta di un grandissimo numero di altri fattori (direttamente linguistici, paralinguistici e perfino extralinguistici) che, in momenti e modi diversi, intervengono a rendere plastico e flessibile sia il percorso di produzione sia quello di ricezione linguistica²². È chiaro che la fonotassi di una lingua, rappresentata anche attraverso una descrizione statistica delle combinazioni di fonemi e del rendimento delle diverse opposizioni, non è l'unico mezzo che conferisce prevedibilità al segnale e che facilita l'individuazione del senso di un enunciato. Né forse è il principale strumento di ricostruzione.

Certo è che la fonotassi svolge un ruolo, talora centrale, nella competenza linguistica che viene usata quando, di fronte al problema di identificare il significante di un messaggio, si provano diverse strategie risolutorie. Talora invece, forse anche più frequentemente, il ruolo giocato dalla fonotassi (e dalle competenze statistiche possedute dai parlanti nativi) può essere minimo o nullo.

Se la comprensione metaforicamente può essere vista come un'operazione (manuale) per aprire una scatola chiusa, si può pensare alla fonotassi (come peraltro a quasi tutti gli altri fattori che fungono da appiglio per la comprensione linguistica) come a uno strumento disponibile nella nostra cassetta degli attrezzi linguistica, parte di un coltellino svizzero che contiene altri dispositivi da questo inseparabili. Lo strumento fonotattico è accessibile ma non obbligatoriamente adot-

²² Come suggerito in Albano Leoni – Clemente, in questo volume, pp. 27-44.

tato come principale per dischiudere uno spiraglio nella scatola. Possiamo usarlo oppure no, in genere lo teniamo sempre in mano mentre operiamo. I modi di aprire le scatole diverse – fatte di diversi materiali e chiuse con diverse serrature – sono molteplici, tutti necessitano l'utilizzo di più di uno strumento alla volta.