

LAURA BAFILE

L'individuazione dei primitivi della Teoria degli Elementi: la questione dell'economia

1. Introduzione

Secondo la teoria segmentale nota come Teoria degli Elementi, i suoni linguistici sono dei composti di “atomi”, cioè combinazioni di suoni elementari. Tale concezione si basa sull'osservazione, intuitiva e tradizionale, che le vocali medie sono composte da vocali alte e vocali basse. I primitivi segmentali così concepiti, a differenza dei tratti binari della teoria fonologica tradizionale, sono privativi, cioè possono soltanto essere presenti o non presenti in un segmento, e sono autonomamente interpretabili dalla fonetica, proprietà per la quale ciascun elemento, se isolato, corrisponde direttamente a un certo suono, e in combinazione con altri elementi apporta le proprie caratteristiche al composto. Questa concezione, che aveva già trovato diverse formulazioni nella fonologia generativa successiva a Chomsky – Halle (1968) (cf. Anderson – Jones 1974; Schane 1984; Anderson – Ewen 1987) ha raggiunto nella Teoria degli Elementi una teorizzazione compiuta e coerente, inserendosi nel più ampio quadro teorico della Fonologia della Reggenza (*Government Phonology*) (cf. Kaye – Lowenstamm – Vergnaud 1990). In questo contesto, la Teoria degli Elementi condivide l'obiettivo di un'adeguatezza esplicativa della rappresentazione fonologica e della costruzione di un modello della competenza e dell'acquisizione restrittivo e non arbitrario, offrendo una soluzione ai problemi derivanti da una illimitata potenza derivazionale delle regole della fonologia generativa standard (cf. ad esempio Hooper 1976; Kiparsky 1982; Kaye 1989; Harris 1994).

La formulazione della Teoria degli Elementi che possiamo definire “classica” è quella di Harris – Lindsey (1995). Gli sviluppi dei decenni a seguire hanno portato a revisioni rilevanti e i modelli più recenti differiscono dalle proposte originarie in molti aspetti significativi (per una rassegna delle successive revisioni della Teoria degli Elementi si veda Backley 2012). Un'affermazione generalmente condivisa è che il numero di elementi previsto da Harris – Lindsey (1995) debba essere diminuito, e la discussione ampia che ne è scaturita si focalizza su come ottenere questo risultato e quali tagli operare all'interno dell'inventario originale. In questo articolo sono presentate alcune considerazioni su que-

sto tema, con lo scopo principale di spostare la discussione dal “come” ridurre l’inventario degli elementi al “se” tale risultato sia veramente un obiettivo della teoria fonologica.

Il paragrafo 2 è una sintetica presentazione del concetto di elemento, il paragrafo 3 illustra i criteri generali alla base dell’individuazione degli elementi, mentre nel paragrafo 4 si discute la questione della dimensione dell’inventario, proponendo una diversa prospettiva, secondo la quale la riduzione del numero di elementi non va nella direzione di una migliore capacità esplicativa della teoria fonologica. Segue un breve paragrafo di conclusioni.

2. Proprietà fondamentali degli elementi

Secondo la fonologia generativa classica (Chomsky – Halle 1968) le unità subsegmentali sono tratti binari, e dato un tratto T si individuano due insiemi, i segmenti [+T] e i segmenti [-T]. Questa definizione può riferirsi a due situazioni diverse, entrambe osservabili nelle lingue naturali. Nella prima, T è una proprietà dei segmenti individuabile in quanto fonologicamente attiva o comunque pertinente nei processi fonologici. Questa condizione è chiaramente esemplificata dai processi di assimilazione dovuti alla diffusione di un tratto, come i fenomeni di nasalizzazione o labializzazione. In casi come questi, il tratto T, che corrisponde rispettivamente a [nasale] e [arrotondato], è il tratto pertinente che può essere solo presente o assente, producendo quindi solo opposizioni privative. Dato un tratto T privativo, non esiste una proprietà [-T] fonologicamente attiva, cioè, per restare all’esempio, non si osserva la diffusione di una proprietà [non nasale] o [non arrotondato]. In altri casi, invece, entrambi i valori di T corrispondono a proprietà segmentali attive in processi fonologici e individuano una classe naturale di suoni. È il caso, ad esempio, del tratto [±sonorante], che per ciascuno dei due valori individua una classe di consonanti accomunate da caratteristiche di distribuzione o da processi fonologici; ad esempio, i segmenti [+sonorante] possono essere l’unica coda possibile in molte lingue, mentre [-sonorante] sono le consonanti soggette in molte lingue a fenomeni di lenizione o desonorizzazione.

Gli elementi, al contrario, sono monovalenti, e quindi esclusivamente privativi. Come tali, essi formano i segmenti grazie all’informazione fonologicamente pertinente che contengono, e non esistono elementi corrispondenti al valore negativo di un tratto binario. In un modello a elementi, ad esempio, l’armonia vocalica consistente nella propagazione dell’arrotondamento labiale è attribuita al ruolo dell’elemento |U| presente nella [u] e

nelle altre vocali labiali. D'altra parte, un contrasto non privativo come la distinzione tra segmenti sonoranti e ostruenti, che non può essere rappresentato dal cambiamento di valore di un singolo tratto, è determinato da più complesse interazioni all'interno dei composti segmentali. Nel modello classico, gli elementi che entrano in gioco nella composizione delle consonanti ostruenti, oltre a eventuali elementi di luogo, sono le unità di modo $|\text{ʔ}|$ e $|\text{h}|$. Il primo è l'elemento di occlusione, il secondo è il "rumore" che corrisponde al rilascio turbolento e alla componente spettrografica aperiodica. In combinazione, questi sono due ingredienti delle occlusive vere e proprie, dotate di occlusione e fase di rilascio udibile. Da solo, $|\text{h}|$ corrisponde alla componente di modo delle consonanti fricative, mentre $|\text{ʔ}|$ senza $|\text{h}|$ forma le consonanti occlusive prive di rilascio udibile. Questa rappresentazione per mezzo di unità privative permette di caratterizzare in modo molto diretto i fenomeni di lenizione, che sono interpretati come impoverimento del contenuto segmentale. La lenizione può trasformare le consonanti ostruenti per effetto di spirantizzazione (perdita di $|\text{ʔ}|$), glottalizzazione (perdita di $|\text{h}|$), debuccalizzazione (perdita degli elementi di luogo), vocalizzazione (perdita degli elementi di modo).

L'altra caratteristica degli elementi è che, al contrario dei tratti binari della teoria standard, essi sono singolarmente pronunciabili; ad esempio, mentre il tratto [+basso] non individua di per sé un suono, l'elemento $|\text{A}|$ ha una propria interpretazione fonetica e forma da solo la vocale [a]. Nel quadro teorico della Fonologia della Reggenza e della Teoria degli Elementi, la relazione tra categoria fonologica e forma fonetica è definita in base a un modello che nega la priorità della dimensione articolatoria assunta nella fonologia generativa standard. Ciò che è in comune tra parlanti e ascoltatori è il segnale acustico, e pertanto quelli che la grammatica utilizza sono simboli uditivi (cf. Kaye – Lowenstamm – Vergnaud 1985; Lindsey – Harris 1990; Harris – Lindsey 1995; Backley 2011). Va precisato però che il segnale acustico è costituito solo in parte da contenuto linguistico: il parlato, infatti, veicola anche informazione non linguistica, consistente, ad esempio, in segnali per l'identificazione di gruppo o per l'espressione di emozioni, o determinati da altri fattori individuali e contingenti (Kaye 1989; 2005; Traunmüller 1994). All'interno del segnale acustico, gli elementi corrispondono a schemi uditivi, «acoustic signatures» nei termini di Kaye (2005, 285), cioè astrazioni mentalizzate di caratteristiche acustiche, che vengono cercate e individuate nel flusso sonoro dagli utenti della lingua e dai bambini in fase di acquisizione linguistica.

Sebbene dotati di un'interpretazione fonetica, in quanto categorie di livello fonologico gli elementi corrispondono a forme fonetiche non definite rigidamente. Ad esempio, in una lingua con una sola vocale bassa centrale, l'elemento $|\text{A}|$ può essere realizzato da un

suono come [a], oppure [ɑ] o [ɛ]. Un elemento può corrispondere in effetti a una certa varietà di interpretazioni fonetiche, a seconda della lingua, del contesto o anche del singolo parlante, che possono dare forma alquanto diversa a una stessa informazione fonologica.

Il concetto di primitivi segmentali come unità privative e singolarmente pronunciabili, fin dalle prime teorizzazioni, è connesso con la nozione di dipendenza, secondo cui gli elementi all'interno di un segmento possono avere pesi diversi, contribuendo in modo ineguale alla sua composizione (cf. Anderson – Jones 1974; Schane 1984; Kaye – Lowenstamm – Vergnaud 1985; Harris – Lindsey 1995). Anche questa concezione trova una prima, chiara evidenza nelle distinzioni di altezza vocalica in cui, ad esempio, la fusione degli elementi |A| e |I| può produrre più di un segmento, come una vocale [æ] in cui predomina |A|, e una vocale [e] in cui predomina |I|. Sulle relazioni di dipendenza segmentale sono stati proposti approcci diversi. L'ipotesi più restrittiva è quella formulata nel modello classico della Teoria degli Elementi (Harris – Lindsey 1995), secondo cui all'interno di un composto solo un elemento può essere la testa, cioè dominare sugli altri; pertanto, dati due elementi, ad esempio |A| e |I|, sono possibili i composti, |A I| e |A I| (la sottolineatura identifica la testa). Secondo interpretazioni più permissive della dipendenza (cf. Backley – Nasukawa 2009; Backley 2011; 2012) un composto può essere privo di testa o contenere più di una testa: ad esempio in Backley (2011) la consonante aspirata dell'inglese [p^h] è rappresentata dal composto |U ? H|. Inoltre, nella concezione meno restrittiva dei rapporti di dipendenza, dato un solo elemento, possono esistere due espressioni segmentali diverse, una in cui l'elemento è testa e una in cui l'elemento non è testa. Ad esempio, nella rappresentazione di Backley (2011) relativa all'inglese, la vocale piena [i:] in *green* 'verde' corrisponde a |I|, mentre la vocale debole [ɪ] in *because* 'perché' corrisponde a |I|. È evidente che tale concezione nega il valore intrinsecamente relazionale del concetto di dipendenza, permettendone applicazione potenzialmente arbitrarie. Inoltre, l'utilizzo meno vincolato della nozione di testa permette, a parità di numero di elementi, un numero maggiore di composti diversi. Torneremo su questo punto al paragrafo 4.

3. L'individuazione dei primitivi segmentali

In un modello in cui le unità subsegmentali sono privative, i segmenti sono costituiti solo dalle unità fonologicamente pertinenti. Ogni segmento ne contiene perciò un numero variabile e comunque piccolo (nella maggior parte dei casi da uno a quattro). Ciò comporta

che l'insieme dei primitivi della Teoria degli Elementi sia più piccolo di quello dei modelli basati su tratti binari. Harris – Lindsey (1995) proponevano un sistema con dieci elementi; da allora, la dimensione dell'insieme si è andata progressivamente riducendo, perché quasi tutti i modelli sviluppati successivamente hanno proposto l'eliminazione di elementi.

Harris – Lindsey (1995) individuano cinque elementi di luogo, |A|, |I|, |U|, |R|, |@|. |R| è l'elemento di coronalità e |@| è l'elemento neutro, cioè l'elemento privo di specificazione di luogo, interpretato foneticamente come *schwa* e, nelle consonanti, come componente di velarità. Gli elementi di modo, già menzionati, sono due, |ʔ| (occlusione) e |h| (“rumore”). L'attività laringale è rappresentata da |L| (energia periodica di bassa frequenza, corrispondente a tono basso nelle vocali e a sonorità nelle consonanti), e |H| (energia di alta frequenza, tono alto nelle vocali e aspirazione nelle consonanti). Infine |N| corrisponde alla nasalità.

Le revisioni di questo modello classico, che hanno quasi senza eccezioni portato a una riduzione del numero di elementi, sono intervenute su tre aspetti del sistema.

Il primo aspetto è quello che riguarda la dimensione luogo di articolazione (*place*). Rispetto a questa questione, già Harris e Lindsey riconoscevano che |R| e |@| sono categorie problematiche: la problematicità dell'elemento neutro è quella di essere una “non proprietà”, cioè di essere privo di proprietà fonologicamente pertinenti. Per quanto riguarda la coronalità, da un lato le caratteristiche acustiche corrispondenti a |R| sono sfuggenti, dall'altro il comportamento fonologico delle consonanti coronali è notoriamente particolare in molte lingue. Le coronali mostrano diffusamente un comportamento “passivo”, nel senso che esse tendono a essere suoni inattivi o trasparenti in processi fonologici, come ad esempio l'assimilazione o l'armonia vocalica (cf. Paradis – Prunet 1991; per una discussione nell'ambito della Teoria degli Elementi cf. Harris – Lindsey 1995, 68-69; Backley 1993).

Backley (2011) rappresenta la versione definita “standard”, che ha accolto alcune delle proposte di revisione presentate dagli esordi della Teoria degli Elementi a oggi. In merito agli elementi di luogo, tale modello esclude |@| come elemento per la rappresentazione delle vocali e delle consonanti velari, e |R| dalla rappresentazione delle consonanti coronali. Labialità e velarità convergono e vengono rappresentate entrambe con l'elemento |U| (|U| = labiale, |U| = velare) (Backley 2011, 79). La coronalità viene a essere espressa dall'elemento di palatalità |I| oppure dall'elemento di faringalità |A| (Backley 2011, 77-79). Secondo l'autore, l'osservazione empirica evidenzia da una parte che le coronali sono caratterizzate in alcune lingue da |I| e in altre da |A|, e dall'altra parte che, in una data lin-

gua, alcune coronali si comportano come segmenti contenenti [I] e altre come segmenti contenenti [A]. L'eliminazione di [R] persegue anche un obiettivo più generale, cioè quello di una rappresentazione completamente unificata del contenuto melodico di consonanti e vocali nel rispetto del principio che le categorie melodiche non sono di per sé consonantiche o vocaliche, e che ciò che fa di un composto segmentale una consonante o una vocale è il suo status sillabico, cioè l'associazione, rispettivamente, a un Attacco (o una Coda) o a un Nucleo.

Quest'ultima considerazione vale anche per gli elementi di modo (*manner*), il secondo aspetto su cui si sono incentrate le proposte di revisione. La Teoria degli Elementi standard prevede la fusione di [h], l'elemento esclusivamente consonantico contenuto nelle occlusive e nelle fricative, con [H]. Nell'analisi di Backley – Nasukawa (2009) e Backley (2011), [H], che caratterizza il tono alto nelle vocali, è contenuto nelle consonanti occlusive e nelle fricative in cui corrisponde al “rumore”, cioè alla turbolenza del rilascio dell'aria. In alcune lingue, tra cui ad esempio l'inglese, [H] determina anche la presenza di aspirazione nelle occlusive e l'assenza totale di sonorità nelle ostruenti più in generale. L'elemento di occlusione è invece mantenuto in Backley (2011), in quanto giustificato dall'esistenza di vocali glottalizzate, rappresentate come composti contenenti [ʔ]. L'ipotesi dell'eliminazione di [ʔ] risale però già alle origini della Teoria degli Elementi (Jensen 1994) ed è stata sviluppata in proposte più recenti (cf. Pöchtrager 2006; Pöchtrager – Kaye 2013) che, nell'ambito di una più articolata teoria della struttura dei segmenti, considerano le proprietà segmentali corrispondenti a [ʔ] e [H] come caratteristiche determinate dalla struttura stessa piuttosto che dal contenuto melodico dei segmenti.

Il terzo cambiamento riguarda la rappresentazione della nasalità, che nella teoria standard converge con la sonorità nell'elemento [L]. Anche per questo passo le motivazioni sono sia di tipo acustico, sia relative al comportamento dei segmenti nei processi fonologici. Dal punto di vista acustico, il “mormorio” caratteristico dei suoni nasali e il segnale glottidale della sonorità consistono entrambi in energia periodica a bassa frequenza. In una prospettiva tipologica, si osserva che molte lingue sono caratterizzate da restrizioni fonotattiche o processi che evidenziano uno stretto legame tra le occlusive sonore e le occlusive nasali, chiaramente distinte, come insieme, dalle occlusive sorde (cf. Nasukawa 2005; Backley – Nasukawa 2010).

In sintesi, l'insieme della Teoria degli Elementi standard contiene i sei elementi [A], [I], [U], [H], [L], [ʔ]. Revisioni ulteriori, sviluppando la proposta di Jensen (1994) sull'elemento di occlusione, spostano proprietà dei segmenti dal piano della melodia a quello della

struttura e arrivano a un insieme ridottissimo costituito dai soli tre elementi [I], [U], [L] (ad esempio Pöchtrager – Kaye 2013).

4. La ricerca dell'economia

La tendenza alla riduzione del numero degli elementi ha come obiettivo fondamentale un'adeguata rappresentazione delle categorie fonologiche pertinenti nella composizione dei segmenti. In questa prospettiva, l'esatta forma fonetica dei suoni ha un'importanza limitata, dato che, come già osservato, il contenuto fonologico dei segmenti è rivelato dal loro comportamento all'interno degli enunciati e dalle caratteristiche acustiche essenziali. È per questo che, ad esempio, mentre sonorità e nasalità sono caratteristiche fonetiche diverse, ci sono ragioni convincenti per considerarle entrambe come determinate da una stessa proprietà di un livello più astratto.

Esiste tuttavia un argomento di altra natura, di fatto generalmente menzionato nelle proposte in favore della riduzione, quello che fa riferimento al pericolo della “*overgeneration*”. Il termine, riferito a modelli di sistemi fonologici, indica la capacità di generare insiemi di segmenti molto più grandi degli inventari effettivamente osservabili nelle lingue naturali.

Nella letteratura che fa capo alla Teoria degli Elementi tale questione ha ricevuto fin dall'inizio notevole attenzione, e non è un caso che quasi tutte le proposte di diminuzione degli elementi menzionino tra le motivazioni il rischio di ipergenerazione. Ad esempio, Backley (1993, 301) osserva: «The ten elements inventory [...] still has the capacity for generating a significant number of expressions which are either unobserved or non-contrastive in the world's languages»; Jensen (1994, 72-73): «Most of the segments generated by the fusion operation [...] seem to have no interpretation. Even the [...] Revised Theory with only {A, I, U, H, L, ?} and no dependency nodes generates a total of 256 representations. But languages typically seem to support between 50 and 100 linguistically significant contrasts in their phonologies»; Charette – Göksel (1994, 2): «The main idea [of Revised Theory of Elements] is to overcome the overgeneration of phonological expressions by reducing the number of elements»; Pöchtrager (2006, 14): «P[honological] E[xpression]s only encode what is *phonologically relevant*, and current estimates are that the number of expressions needed will be well below 100. Any theory generating more than that is certainly wrong»; Backley (2011, 142): «There is no doubt that [employing a 'noise' element |h| in addition to |H|] increases the richness and expressive power of rep-

resentations; however, a larger set of elements always brings with it the risk of overgeneration».

Si deve osservare tuttavia che da parte degli autori che esprimono preoccupazione al riguardo, la questione della ipergenerazione non è realmente approfondita, e non è spiegato in che senso questo sia un tema pertinente per una teoria della grammatica.

Naturalmente quella di evitare la ipergenerazione può essere considerata una esigenza generica di economia e di eleganza di un modello del sistema segmentale, così come di altri moduli della grammatica. Da questo punto di vista, la Teoria degli Elementi è sicuramente migliore delle teorie che utilizzano tratti binari, dato che anche l'insieme più ricco, costituito da dieci elementi, può produrre 6.144 combinazioni, cioè molte meno delle 1.048.576 combinazioni possibili con una matrice di venti tratti binari. Ma se l'obiettivo è quello più preciso di escludere sistemi linguistici non attestati dall'insieme delle grammatiche possibili, i dieci elementi della teoria classica sono ancora troppi, perché 6.144 è un numero che supera non solo la dimensione dell'inventario dei segmenti delle singole lingue, ma anche l'insieme di tutti i segmenti osservabili nelle lingue del mondo.

È necessario d'altra parte considerare che, per quanto il pericolo dell'ipergenerazione sia ritenuto generalmente una motivazione sufficiente e prioritaria, l'eliminazione di elementi è un cambiamento non privo di costi.

Una prima, ovvia conseguenza è il pericolo di "ipogenerazione", cioè che il totale delle combinazioni segmentali ottenibili con gli elementi a disposizione non sia sufficiente a esprimere i contrasti osservabili nelle lingue del mondo. Questo problema può essere almeno in parte superato adottando una concezione in cui gli elementi hanno una «risoluzione fonetica» piuttosto bassa (Breit 2013, 42-43), e sono quindi compatibili con interpretazioni fonetiche piuttosto diverse tra loro. In tal modo si prevede l'esistenza di lingue che, pur avendo lo stesso insieme di composti di elementi, presentano variazioni apprezzabili al livello della forma fonetica più dettagliata; questo può essere il caso, ad esempio, del composto |A I| che corrisponde a [ɛ] in certe lingue e a [æ] in altre (Backley 2011, 20). Il problema della ipogenerazione ha peraltro incoraggiato la ricerca di dispositivi che garantiscano la capacità espressiva del modello permettendo l'incremento delle possibilità combinatorie a parità di numero di elementi. Con questo obiettivo è stata sostenuta un'interpretazione molto meno restrittiva ed economica del concetto di testa di un composto. Il modello standard ammette infatti composti senza testa o con più di una testa; inoltre, un'espressione costituita da un solo elemento può corrispondere a due diversi segmenti, uno in cui l'elemento è testa e uno in cui l'elemento è dipendente. È significativo che gli stessi sostenitori sottolineino come questa concezione "permissiva" della di-

pendenza sia funzionale alla capacità espressiva del modello; ad esempio Backley – Nasukawa (2009, 54): «This headship relation serves two important purposes: first, it increases the number of possible melodic expressions – and thus, the number of contrasts – that the model can generate»; Backley (2012, 82): «The obvious advantage of asymmetric fusion is that it increases the number of contrasts that can be expressed using a given number of elements».

L'eliminazione di elementi può causare anche un secondo svantaggio, alquanto paradossale considerando gli argomenti che la motivano, cioè quello di impedire generalizzazioni significative. Argomentando contro i modelli colpevoli di ipergenerazione, Backley (2011, 82) osserva: «Once again, using two elements instead of one has the obvious advantage of greater expressive power. [...] In such a situation, theorists must consider the question of cost». Un buon esempio di questo stato di cose è quello della già menzionata unificazione di velarità e labialità, espresse entrambe da [U] nella teoria standard. Secondo Backley, l'unificazione è coerente con il fatto che, da una parte, le due classi di segmenti condividono una proprietà acustica (il tratto “grave” nei termini di Jakobson) e dall'altra si comportano come una classe naturale in alcune lingue. Inoltre, nella maggior parte delle lingue, nelle vocali labialità e velarità sono proprietà correlate, in quanto le vocali velari sono anche arrotondate (cf. Scheer 1999). Tuttavia si deve notare che, tipologicamente, non emerge un'evidenza univoca che labialità e velarità costituiscano una stessa categoria. Scheer (1999) menziona al riguardo l'esistenza in alcune lingue di vocali velari non arrotondate insieme al fatto che le consonanti velari non sono arrotondate. Ancor più importante è il comportamento fonologico delle vocali posteriori, [u] in particolare, che interagiscono sia con le consonanti labiali sia con le consonanti velari; ad esempio in ceco, la forma del suffisso nominale di vocativo è determinata dalla qualità della consonante precedente, nel modo seguente (cf. Scheer 1999, 209-210):

(1)		NOMINATIVO	VOCATIVO	
<i>dopo consonante palatale:</i>	/-i/	kuuj	kɔj-i	‘cavallo’
<i>dopo consonante velare:</i>	/-u/	ptaak	ptaak-u	‘uccello’
<i>dopo altre consonanti:</i>	/-e/	fírat	fírad-e	‘castello’

Nell'esempio in (1), la presenza della vocale posteriore [u], che è velare e arrotondata, si correla alla presenza di una consonante velare, che non è arrotondata. Scheer (1999, 210) afferma quindi che «consonantal representations should reflect the affinity of [u]

[...] with both labials and velars» e propone due diversi primitivi per la velarità e l'arrotondamento labiale, in modo tale che [u] contiene sia |U| sia |B|, le consonanti velari contengono |U| e le consonanti labiali contengono |B|.

Il caso problematico della distinzione velarità/labialità illustra una situazione, che è ragionevole pensare non sia isolata, in cui l'unificazione di proprietà segmentali differenti in un'unica categoria di livello più astratto può consentire generalizzazioni in qualche caso o per alcune lingue, ma impedirle in altri casi o per altre lingue.

Alla luce di queste considerazioni, possiamo legittimamente chiederci se l'ipergenerazione sia davvero una questione pertinente per la teoria segmentale.

Per argomentare in merito a questa domanda fondamentale, consideriamo brevemente la prospettiva tipologica. Gli studi di questo ambito indicano che la maggior parte delle lingue del mondo ha inventari segmentali che contengono da 20 a 40 unità. Nel *UCLA Phonological Segment Inventory Database* (UPSID), le lingue con 20-40 segmenti costituiscono il 78% delle 451 lingue del campione analizzato. La cardinalità minima nell'UPSID è 11 (in due lingue) e quella massima è 141 (in una lingua); inoltre, solo otto lingue hanno un numero di segmenti inferiore o uguale a 15, e solo nove lingue ne hanno un numero uguale o superiore a 60. Per quanto riguarda la diversità segmentale, il campione UPSID contiene 919 segmenti diversi, ma l'80% di essi è limitato a dieci lingue al massimo; ciò significa che quasi tutte le lingue formano i propri inventari attingendo da un insieme molto più piccolo di segmenti, nel caso del campione UPSID un insieme da 184 unità. Questi dati evidenziano che, per quanto riguarda gli inventari segmentali, le lingue hanno una forte preferenza per un sottoinsieme di tutte le cardinalità attestate (20-40) e per un sottoinsieme di tutti i segmenti attestati (184). È però altrettanto evidente che, al di là di limiti e preferenze molto chiari, nell'ambito dei sottoinsiemi preferiti la variazione interlinguistica è ampia e non riconducibile a plausibili fattori non arbitrari.

Questo tipo di dati produce un quesito fondamentale per la teoria fonologica e la teoria del linguaggio più in generale, cioè se questo quadro tipologico sia determinato da condizioni puramente grammaticali o dipenda invece da fattori più generali, non esclusivamente linguistici, anche riguardanti proprietà cognitive e motorio-percettive della specie umana.

Le proprietà della percezione e dell'articolazione sono ovviamente responsabili dei limiti riguardanti la diversità segmentale delle lingue naturali: il numero complessivo di segmenti diversi utilizzabili dalle lingue del mondo è limitato dalla capacità umana di produrre e udire suoni.

Per quanto riguarda la cardinalità minima e massima, è ragionevole che il limite inferiore alla dimensione degli inventari sia dovuto al fatto che un numero ancora più piccolo di segmenti sarebbe insufficiente alla esternalizzazione delle differenze lessicali. Il limite massimo può invece rispondere a esigenze di computazione ottimale, non specificamente linguistica, che interagisce con caratteristiche dell'articolazione e ancor più della percezione uditiva. Dato lo spazio acustico e articolatorio consentito dal sistema motorio-percettivo umano, un numero di segmenti superiore ai limiti osservati determinerebbe una minore differenziazione tra i segmenti e quindi una discriminazione troppo difficile, e comunque non ottimale per la computazione.

La differenza significativa tra cardinalità massima e numero di segmenti attestati nelle lingue del mondo (nell'UPSID 141 contro 919; ma è plausibile che anche campioni analizzati con criteri diversi produrrebbero valori molto distanti) è in accordo con l'osservazione di Jakobson (1941), poi confermata sperimentalmente (cf. Guasti 2002 e bibliografia ivi citata), che la notevole capacità di discriminazione, anche fra suoni linguistici non familiari, presente nel bambino già dal primo mese di vita, decresca rapidamente nel progredire dell'acquisizione della madrelingua. In altri termini, il numero di suoni contrastivi acquisibili è ben maggiore del numero di fonemi utilizzati da una singola grammatica. A tale riguardo bisogna osservare che i parlanti plurilingui gestiscono complessivamente un numero di contrasti fonemici maggiore rispetto ai monolingui; resta il fatto, tuttavia, che il numero di contrasti all'interno di uno stesso sistema resta nei limiti osservati.

Considerazioni non specificamente linguistiche possono spiegare anche l'esistenza di intervalli preferiti dalle lingue del mondo per quanto riguarda la cardinalità e la diversità degli inventari. È plausibile che un inventario di 30 segmenti costituisca un miglior compromesso tra le esigenze espressive e l'efficacia della computazione rispetto a inventari di 11 o di 141 segmenti, che sono tipologicamente rari. Rispetto alle diverse probabilità di attestazione dei vari segmenti, la ragione per cui la nasale bilabiale sonora è presente nel 94,2% delle lingue del campione UPSID mentre la fricativa laterale dentale/alveolare palatalizzata compare solo in una lingua deve essere cercata in principi della percezione uditiva e dell'articolazione dei suoni in generale.

Va sottolineato che queste considerazioni non si richiamano a una generica rappresentazione di tipo funzionalista basata sulla dinamica tra le contrapposte esigenze di "minimo sforzo articolatorio" e "massima differenziazione". Consideriamo ad esempio l'universale tipologico per cui le lingue che hanno tre o più vocali hanno le vocali *a*, *i*, *u*, e quindi la presenza di vocali medie implica la presenza di *a*, *i*, *u* (cf. Crothers 1978).

Un'osservazione intuitiva è che, dal punto di vista percettivo, le tre vocali al vertice del triangolo vocalico sono più distanti tra loro e quindi massimamente differenziate; ma a un esame più approfondito emerge una caratteristica meno ovvia di *a*, *i*, *u* ovvero la loro natura quantica (cf. Stevens 1989), che fa sì che la loro qualità acustico-uditiva sia molto stabile, nonostante la notevole variabilità che caratterizza la loro articolazione. In tal senso, una delle cause della priorità delle vocali di vertice rispetto alle altre vocali dipende da un'interazione specifica e niente affatto ovvia di fattori uditivi e articolatori.

Riassumendo, le lingue mostrano limiti assoluti e preferenze riguardo alla dimensione e alla qualità degli inventari segmentali, che vanno attribuite tanto a caratteristiche del sistema percettivo-motorio, quanto a proprietà cognitive generali. Tutti questi fattori non appartengono specificamente alla facoltà del linguaggio, e gli effetti che producono non rientrano nell'ambito di interesse di una teoria della grammatica. All'interno dei limiti stabiliti da tali restrizioni di tipo non linguistico, le lingue mostrano invece un'ampia variazione riguardo alle combinazioni di elementi, differenziandosi molto rispetto alla dimensione degli inventari segmentali e alla loro qualità.

Pertanto, se le combinazioni di elementi logicamente possibili ma in effetti non attestate sono escluse per effetto di fattori non linguistici, allora non è la grammatica a determinare quanti e quali segmenti possono trovarsi negli inventari segmentali. La conseguenza per una teoria della grammatica è che non esiste un problema di ipergenerazione.

Una critica alla preoccupazione diffusa nella linguistica generativa nei confronti di modelli troppo ricchi e troppo potenti è in Reiss (2012), il quale osserva che la «esplosione combinatoria» (180) dei sistemi di tratti fonologici non è un problema se i sistemi sono caratterizzati intensionalmente, cioè nella prospettiva della lingua interna, anziché estensionalmente. L'osservazione, riportata alla Teoria degli Elementi, significa che i bambini in fase di acquisizione devono costruire l'inventario segmentale della lingua a cui sono esposti, e per fare questo esplorano il segnale alla ricerca delle “acoustic signatures” (nel senso di Kaye 2005) corrispondenti agli elementi che, isolati o in combinazione, compongono le consonanti e le vocali della lingua. Non c'è ragione di ipotizzare che i bambini partano dall'assunzione che ciascun elemento debba combinarsi con tutti gli altri, e ancor meno che si pongano l'obiettivo di valutare le migliaia di inventari segmentali logicamente possibili per arrivare a scoprire come è fatto quello della propria lingua. Un'ipotesi più plausibile è che partendo dalla conoscenza – che possiamo ipotizzare faccia parte di una dotazione genetica – dell'esistenza di primitivi segmentali dotati di specifiche proprietà (privatività, interpretabilità fonetica autonoma, composizionalità, dipendenza) e utilizzando una capacità di percezione uditiva altamente sviluppata fin dai

primi giorni di vita, i bambini analizzano il parlato che costituisce il loro *input* alla ricerca delle categorie corrispondenti agli elementi, per arrivare a individuare i segmenti che compongono le parole della madrelingua.

Riguardo alla natura della dotazione genetica, l'ipotesi minima è che essa consista nella conoscenza dell'esistenza degli elementi e delle loro proprietà; in questa ipotesi, non è necessario attribuire alla Grammatica Universale l'identità degli elementi stessi. In altri termini, non è necessario pensare che l'acquisizione dei segmenti parta dagli elementi come categorie preesistenti, ciascuna dotata di una propria interpretazione fonetica; possiamo limitarci a ipotizzare che la conoscenza iniziale sia semplicemente che i suoni linguistici sono fatti di elementi, in accordo con i principi più generali di discretezza e combinatorietà che informano la facoltà del linguaggio. A partire da questa base, il segnale acustico viene esaminato alla ricerca dei primitivi segmentali, nella forma di schemi acustico-uditivi presenti individualmente o all'interno di composti, per arrivare all'individuazione degli elementi e delle loro proprietà.

Un'ipotesi come questa è estremamente parsimoniosa per quanto riguarda il ruolo assegnato alla Grammatica Universale ed è per questo in sintonia con l'approccio "bottom up" alla conoscenza della facoltà del linguaggio che Chomsky (2007) affianca all'approccio "top down", che è stato per decenni esclusivo nella linguistica generativa (cf. Reiss 2012 per una discussione di questo approccio applicata a modelli segmentali binari). Nella prospettiva «top down» la domanda è: «How much must be attributed to UG to account for language acquisition?», mentre l'approccio «bottom up» si chiede: «How little can be attributed to UG while still accounting for the variety of I-languages attained, relying on third factor principles?» (Chomsky 2007, 4).

Nell'ipotesi qui prospettata, la Grammatica Universale contiene solo l'informazione minima relativa all'esistenza dei primitivi segmentali. Le caratteristiche comuni ai sistemi segmentali delle lingue naturali sono in parte determinate da principi non linguistici del terzo fattore (Chomsky 2005). Per quanto riguarda la diversità interlinguistica, per cui ad esempio alcune lingue hanno inventari di 20 e altre di 40 segmenti, e alcuni suoni che sono presenti in molte lingue mancano in molte altre, è semplicemente una questione di variazione lessicale, una conoscenza che i parlanti costruiscono attraverso l'acquisizione del vocabolario della loro lingua.

Alla luce di queste considerazioni, gli sforzi di costruire modelli segmentali usando insiemi piccolissimi di primitivi, quando travalicano un'esigenza generica di economia ed eleganza della rappresentazione formale, appaiono vani e non mirati a un obiettivo di

economia cruciale nella teoria del linguaggio, quale è quello individuato da Chomsky (2007).

5. Conclusioni

La discussione e gli argomenti presentati in questo articolo ci portano a concludere che la riflessione teorica per la costruzione di modelli segmentali basati su elementi, così come l'analisi di fenomeni fonologici condotta all'interno di questo quadro teorico, non hanno come proprio obiettivo quello di ridurre al minimo il numero degli elementi da utilizzare nel modello. Questa conclusione è rinforzata dal fatto che le strategie adottate per preservare l'adeguatezza descrittiva dei modelli, indeboliti dalla diminuzione degli elementi, sollevano a loro volta problemi gravi di adeguatezza descrittiva e di arbitrarietà e non sembrano permettere generalizzazioni cognitivamente appropriate.

Sostenere questo punto di vista non equivale, ovviamente, ad autorizzare la proliferazione degli elementi con il fine della descrizione di caratteristiche fonetiche minute: l'economia della rappresentazione è comunque una delle caratteristiche di una buona teoria, ed è un obiettivo che la Teoria degli Elementi, anche nelle sue formulazioni meno restrittive, ha le caratteristiche per raggiungere. Rispetto a questo, il compito della ricerca fonologica è quello di stabilire quanto delle caratteristiche osservabili dei suoni linguistici appartenga alla sola dimensione fonetica e quanto invece faccia parte del contenuto linguisticamente pertinente e quindi corrisponda a categorie mentalizzate che costituiscono la competenza dei parlanti.

In questa prospettiva la *overgeneration* non è un tema pertinente per una teoria fonologica della lingua interna, e l'individuazione dei primitivi segmentali non deve fondarsi su considerazioni e adeguarsi a esigenze che sono proprie degli approcci esternalisti al linguaggio.

Laura Bafile
Università di Ferrara
Dipartimento di Studi Umanistici
laura.bafile@unife.it

Riferimenti bibliografici

ANDERSON – EWEN 1987

J.M. Anderson – C.J. Ewen, *Principles of dependency phonology*, Cambridge.

ANDERSON – JONES 1974

J.M. Anderson – C. Jones, *Three theses concerning phonological representation*, «Journal of Linguistics» X 1-26.

BACKLEY 1993

P. Backley, *Coronal the undesirable element*, «UCL Working Papers in Linguistics» V 301-323.

BACKLEY 2011

P. Backley, *An introduction to Element Theory*, Edinburgh.

BACKLEY 2012

P. Backley, *Variation in Element Theory*, «Linguistic variation» XII/1 57-102.

BACKLEY – NASUKAWA 2009

P. Backley – K. Nasukawa, *Headship as melodic strength*, in K. Nasukawa – P. Backley (eds.), *Strength relations in phonology*, 44-77.

BACKLEY – NASUKAWA 2010

P. Backley – K. Nasukawa, *Consonant-vowel unity in Element Theory*, «Phonological Studies» XIII 21-28.

BREIT 2013

F. Breit, *Formal aspects of Element Theory*, London, MRes dissertation.

CHARETTE – GÖKSEL 1994

M. Charette – A. Göksel, *Vowel harmony and switching in Turkic languages*, «SOAS Working Papers in Linguistics and Phonetics» IV 31-52.

CHOMSKY 2005

N. Chomsky, *Three factors in language design*, «Linguistic Inquiry» XXXVI 11-22.

CHOMSKY 2007

N. Chomsky, *Approaching UG from below*, in U. Sauerland – H.M. Gartner (eds.), *Interfaces + Recursion = Language? Chomsky's Minimalism and the view from syntax-semantics*, Berlin, 1-24.

CHOMSKY – HALLE 1968

N. Chomsky – M. Halle, *The sound pattern of English*, New York.

CROTHERS 1978

J. Crothers, *Typology and universals of vowel systems*, in J.H. Greenberg – C.A. Ferguson – E.A. Moravcsik, (eds.), *Universals of human language: phonology*, vol. II, Stanford, 93-152.

GUASTI 2002

M.T. Guasti, *Language acquisition. The growth of grammar*, Cambridge.

HARRIS 1994

J. Harris, *English sound structure*, Oxford.

HARRIS – LINDSEY 1995

J. Harris – G. Lindsey, *The elements of phonological representation*, in J. Durand – F. Katamba (eds.), *Frontiers of phonology: atoms, structures, derivations*, Harlow, Essex, 34-79.

HOOPER 1976

J.B. Hooper, *An introduction to Natural Generative Phonology*, New York.

JAKOBSON 1941

R. Jakobson, *Kindersprache, Aphasie und allgemeine Lautgesetze*, Uppsala.

JENSEN 1994

S. Jensen, *Is ? an element? Towards a non-segmental phonology*, «SOAS Working Papers in Linguistics & Phonetics» IV 71-8.

KAYE 1989

J. Kaye, *Phonology: a cognitive view*, Hillsdale.

KAYE 2005

J. Kaye, *GP, I'll have to put your flat feet on the ground*, in H. Broekhuis – N. Corver – R. Huybregts – U. Kleinhenz – J. Koster (eds.), *Organizing grammar. Studies in Honor of Henk van Riemsdijk*, Berlin, 283-288.

KAYE – LOWENSTAMM – VERGNAUD 1985

J. Kaye – J. Lowenstamm – J.-R. Vergnaud, *The internal structure of phonological elements: a theory of charm and government*, «Phonology» II 305-328.

KAYE – LOWENSTAMM – VERGNAUD 1990

J. Kaye – J. Lowenstamm – J.-R. Vergnaud, *Constituent structure and government in phonology*, «Phonology» VII 192-231.

KIPARSKY 1982

P. Kiparsky, *Explanation in phonology*, Dordrecht.

LINDSEY – HARRIS 1990

G. Lindsey – J. Harris, *Phonetic interpretation in generative grammar*, «UCL Working Papers in Linguistics» II 355-369.

NASUKAWA 2005

K. Nasukawa, *A unified approach to nasality and voicing*, Berlin-New York.

PARADIS – PRUNET 1991

C. Paradis – J.-F. Prunet (eds.), *The special status of coronals: internal and external evidence. Phonetics and phonology 2*, San Diego.

PÖCHTRAGER 2006

M.A. Pöchtrager, *The structure of length*, Wien, Dissertation zur Erlangung des Doktorgrades der Philosophie aus dem Fachgebiet Sprachwissenschaft.

PÖCHTRAGER – KAYE 2013

M.A. Pöchtrager – J. Kaye, *GP2.0*, «SOAS Working Paper in Linguistics» XVI 51-64.

REISS 2012

C. Reiss, *Towards a bottom-up approach to phonological typology*, in A.M. Di Sciullo (ed.), *Towards a biolinguistic understanding of grammar: essays on interfaces*, Amsterdam.

SCHANE 1984

S. Schane, *The fundamentals of Particle Phonology*, «Phonology» I 129-156.

SCHEER 1999

T. Scheer, *A theory of consonantal interaction*, «Folia Linguistica» XXXII 201-237.

STEVENS 1989

K.N. Stevens, *On the quantal nature of speech*, «Journal of Phonetics» XVII 3-45.

TRAUNMÜLLER 1994

H. Traunmüller, *Conventional, biological and environmental factors in speech communication: a modulation theory*, «Phonetica» LI 170-183.

UPSID

(*UCLA Phonological Segment Inventory Database*) compiled by I. Maddieson – K. Precoda, http://web.phonetik.uni-frankfurt.de/upsid_info.html#seg_freq (ultima consultazione 15 novembre 2017).