



Pier Marco Bertinetto

Sillaba

(contributo destinato alla *Enciclopedia dell'italiano*, curata da R. Simone per i tipi dell'Istituto dell'*Enciclopedia Italiana*)*

1 La struttura della sillaba

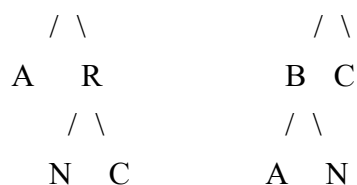
Ogni lingua si articola in sillabe, per effetto dell'alternarsi di foni consonantici e vocalici. I secondi sono caratterizzati da maggior sonorità intrinseca e maggior facilità articolatoria (potendo essere sostenuti fintantoché il fiato lo consente); d'altra parte, una mera successione di vocali mancherebbe di quegli episodi di discontinuità, che rendono meglio percepibile il contenuto di una sequenza fonica. Il linguaggio umano si è costruito sfruttando al meglio queste possibilità e su di esse si è affinata la nostra capacità percettiva (→ fonetica italiana).

Idealmente, il modo ottimale per sfruttare l'alternanza fra consonanti e vocali è rappresentato da sequenze del tipo CV.CV (con C = “consonante”, V = “vocale” e il punto come simbolo di confine di sillaba). Ma benché esistano lingue che ammettono soltanto questo genere di sequenze, il campionario dei tipi sillabici è assai più ampio. In polacco si possono avere fino a 10 consonanti (/mpstfspstr/) nello spazio fra due vocali, come nella sequenza (semanticamente incongrua) *następstw z pstrąga* ‘conseguenze (gen.plur.) di trota’.

La sillaba si costruisce attorno ad un “nucleo”, tipicamente occupato da una vocale. Alla sua sinistra si ha l' “attacco” (ing. onset), a destra la “coda”, entrambi occupati da elementi consonantici. Per es., nella sillaba *lin* si ha /l/ in attacco, /i/ nel nucleo, /n/ in coda. Una sillaba senza coda è tradizionalmente detta “aperta”; con coda, “chiusa”. Se invece manca l'attacco, si dice che la sillaba è “nuda”. La maggior parte degli studiosi concorda nel ritenere che la sillaba sia fondata su una struttura gerarchica, orientata come in figura (1a), con σ = sillaba, A = attacco, N = nucleo, C = coda, R = rima (N+C), B = “corpo” (A+N [ing. body]):

(a) σ (b) σ

* L'autore ringrazia Emanuele Saiu per i molti e preziosi suggerimenti.



La concezione gerarchica della sillaba rivela la propria utilità ogniqualvolta si possa mostrare che un determinato componente strutturale manifesta delle regolarità di comportamento. Per es., in una certa fase del francese si è avuta la nasalizzazione delle vocali, ma solo quand’esse erano seguite da consonante nasale in coda (che è poi caduta), non quando la nasale era nell’attacco della sillaba seguente (cf. bon /bõ/ vs. bonne /bõn/). Si può dunque dire che la rima delimitava il dominio della nasalizzazione. Benché la ramificazione a destra sia più frequente, non mancano casi di ramificazione a sinistra del tipo in (1b). Per es., mentre nei lapsus spontanei italiani gli attacchi tendono a scambiarsi con gli attacchi e le rime con le rime (cf. pento invece di tempo, con automatica assimilazione della nasale, cf. <http://linguistica.sns.it/BancheDati.htm>), in giapponese e coreano prevalgono gli scambi del tipo “corpo” con “corpo” e coda con coda.

Di solito, la suddivisione in sillabe è intuitivamente semplice. Per es., la parola Linda si scandisce in Lin.da. Per render conto del processo di scansione, sono state elaborate opportune scale di “sonorità” o “forza consonantica” (→ fonologia italiana), che regolano la gerarchia dei foni in base ad uno specifico criterio. Le scale di sonorità assegnano il primo posto ai foni maggiormente atti ad occupare la posizione di nucleo (le vocali) e l’ultimo ai foni meno attrezzati a comparirvi (le occlusive e le affricate sorde). Le scale di forza consonantica sono l’esatto inverso delle precedenti, ma ottengono lo stesso risultato. Ad accrescere la capacità predittiva di queste scale, esse vengono corroborate dall’ispezione dei nessi consonantici ammessi ad inizio o fine di parola, calibrando i principi di sillabazione sulla singola lingua. Per es., poiché in polacco i nessi iniziali e finali di parola possono comprendere fino a 4 e, rispettivamente, 5 consonanti, si può avere un’idea circa gli attacchi e le code potenzialmente ammissibili all’interno di parola (cf. /fskʃ/ all’inizio di wskrzesić ‘risorgere’ e /mpstf/ alla fine di przestępstw ‘crimine (gen. plur.)’; per paradosso, la grafia è qui ingannevole, con rispettivamente 5 grafemi consonantici iniziali e 4 finali). E’ importante comunque comprendere che non è la struttura sillabica a regolare la fonotassi, bensì il contrario.

La formulazione più articolata è quella di Vennemann (1988), che propone una serie di “leggi preferenziali”. In base ad esse, per es., un attacco è tanto meglio formato, quanto più bassa è la sonorità (o più alta la forza consonantica) dell’elemento che lo costituisce; se poi l’attacco è formato da due elementi, il criterio di buona formazione prevede che dal primo al secondo ci debba essere una forte crescita dell’indice di sonorità. Ad es., /pV/ è meglio di /pnV/, per non dire di /lpV/; difatti, nelle varie lingue il primo tipo di attacco si trova molto più spesso. Per converso, una coda è tanto meglio formata, quanto maggiore è il suo indice di sonorità, e quanto più bruscamente esso decade allontanandosi dal nucleo (ad es., /Vlp./ è meglio di /Vpl./). Ciò spiega perché le code del lessico tradizionale italiano siano tipicamente costituite da sonoranti (es. can.to, sal.to, lam.po), e perché si avesse in passato la tendenza a correggere la struttura sillabica in presenza di ostruenti in coda (cf. Giosafatte per Giosafat).

Si noti, tuttavia, che le leggi indicate da Vennemann sono “preferenziali”, non assolute; ossia, segnalano delle tendenze che, per quanto robuste, restano violabili. In effetti, benché i principi di scansione producano il risultato corretto nella maggior parte dei casi, essi non risolvono tutti i problemi. Quando il livello di complessità fonotattica supera una certa soglia, emergono situazioni indecidibili. Per es., ing. reptile ammette per la maggior parte degli analisti più di una scansione (/ˈrep.taɪl/ o /ˈrept.aɪl/), dato che il lessico attesta ad inizio e in fine di parola ciascuna delle suddette sequenze (cf. parole come tap o apt). Occorre dunque distinguere due diversi problemi: il computo delle sillabe e l’individuazione dei loro confini. Il primo è di soluzione generalmente agevole ed intuitiva, il secondo presenta maggiori insidie.

In realtà, perfino il primo problema non è privo di insidie, come ben sa chiunque si sia chiesto quante siano le sillabe in it. viòla. Per molti parlanti (come lo scrivente), questa parola è bisillabica, ma per un fiorentino essa è trisillabica, com’è confermato dal grido di incitamento dei tifosi della Fiorentina (vi-o-la, vi-o-la!). La difficoltà è qui costituita dalla scelta fra iato /iˈɔ/ e dittongo /iˈjɔ/; non tutti, del resto, sanno cogliere la sottile differenza tra la quale /la.ˈkwa.le/ e lacuale /la.ku.ˈa.le/. Non così in tedesco – lingua peraltro non priva di dittonghi – dove ad es. eventuel presenta di solito un netto iato, facoltativamente enfaticizzato da un’occlusiva glottidale (cf. ted. /e.ven.tu.[ʔ]ˈeɪ/ con it. /e.ven.ˈtwa.le/, almeno per molti parlanti).

Non c’è dubbio, comunque, che le difficoltà maggiori si presentino in merito alla fissazione dei confini di sillaba, che perfino in lingue con fonotassi abbastanza semplice come l’italiano può presentare insidie. Una di queste è costituita dalla

scansione dei nessi /S/+C, tradizionalmente denominati “esse impura” (la S maiuscola indica l’arcifonema derivante dalla neutralizzazione del tratto di sonorità; → fonetica italiana). Le regole ortografiche impongono per consolidata tradizione di dividere come in, per es., pa.sto, mentre la maggior parte dei fonologi suggerisce la scansione /pas.to/. Il problema non è però di facile soluzione. Il lessico italiano presenta infatti varie parole inizianti in /sC/ (es. stile) e molte meno terminanti in /s/ (cf. gas), il che va in favore della scansione ortografica. Quanto alle scale di sonorità o forza consonantica, esse sono qui di scarsa utilità perché entrano tipicamente in crisi proprio di fronte ai nessi di questo tipo; i quali, a rigore, non dovrebbero ricorrere ad inizio di parola, ma solo all’interno, dove le leggi preferenziali di sillabazione vengono rispettate. Eppure, siffatti nessi iniziali compaiono in molte lingue. Una possibile via d’uscita consiste nel considerare tali sequenze – in posizione interna – come casi di scansione indecidibile. Secondo tale concezione, il parlante italiano oscilla tra pas.to e pa.sto, allo stesso modo in cui il parlante inglese oscilla fra rep.tile e rept.ile. Questo significa avere una concezione probabilistica, anziché deterministica, della scansione sillabica; se poi in molti casi la scansione può apparire deterministica, ciò è solo dovuto al fatto che spesso si verificano condizioni di elevata probabilità positiva (Bertinetto 2004).

2 La sillaba in italiano

La struttura della sillaba italiana è relativamente semplice, almeno per quanto riguarda il lessico autoctono.

La posizione di coda è principalmente assegnata alle sonoranti /r l N/ (dove la maiuscola indica l’arcifonema nasale, con luogo diaframmatico determinato dalla consonante seguente). Altra consonante tradizionalmente ammessa in coda è l’arcifonema /S/, e ciò anche per chi ritenga indeterminata la scansione dei nessi /S/+C (vedi sopra; l’indeterminatezza non esclude infatti una tale eventualità). Infine, in italiano la coda può essere occupata dalla prima parte di una geminata distintiva o intrinseca (→ fonetica italiana): cf. fatto /^lfat.:o/, bagno /^lbaj.:o/. La scansione eterosillabica delle geminate italiane non è stata quasi mai revocata in dubbio. Per quanto riguarda le geminate intrinseche, un argomento diacronico in tal senso deriva dal fatto che, davanti a parole inizianti con tali foni, è rimasto l’allomorfo lo dell’articolo maschile, mentre si è generalizzato il davanti alle scempie o ai nessi del tipo ostruente + liquida/legamento (cf. lo gnomo, lo sciocco, lo zoccolo, ma il nano, il

treno, il t[wɔ]no). Ogni altra possibilità, per quanto riguarda le code italiane, va considerata come dovuta alla pressione dei prestiti, alcuni dei quali di lunga data: cf. rit.mo, naf.ta, mana[ɕ.m]ent. In ogni caso, la coda comprende normalmente un solo elemento; le code biconsonantiche sono rare e ristrette o al lessico non autoctono o a neoformazioni (es.: trans.frontaliero).

Il fatto che l'italiano sia divenuto oggi più tollerante circa i tipi di coda ammessi non deve stupire: nel corso del tempo, la struttura sillabica può evolversi. Un esempio è rappresentato proprio dai nessi /S/+C italiani. Benché si assuma qui una posizione teorica aperta (di tipo probabilistico), non v'è dubbio che, inizialmente, la scansione di tali nessi fosse eterosillabica. Così si spiega il fatto che davanti ad essi non si sia mai avuta la dittongazione di /ε ɔ/, che in italiano antico avveniva appunto in sillaba aperta: cf. tosto, vesti vs. cuore, viene. Del resto, ad inizio di parola questi nessi venivano un tempo regolarizzati con l'inserzione di una /i/ prostetica (es. in Ispagna), come è accaduto con regolarità in spagnolo (cf. escuela 'scuola'). Altro indizio è costituito dal fatto che l'allomorfo dell'articolo maschile sia rimasto lo anche davanti a tali nessi (lo spago vs. il sale). Tuttavia, questi dati diacronici non possono essere direttamente trasferiti all'italiano contemporaneo. Per es., la /i/ prostetica è fortemente regredita; oggi (tranne in locuzioni fossilizzate, come per iscritto) si sente spesso in Spagna; dove, per ragioni strutturali, si può escludere che vi sia una coda biconsonantica */ins.pa/ (semmai, per certi parlanti può esservi semplificazione: /is.pa/). In ogni caso, a parte le considerazioni teoriche (Bertinetto 1999), vari indizi sperimentali portano a dire che in simili casi il comportamento dei parlanti non è univoco (Bertinetto 2004, Turchi & Bertinetto 2000).

La posizione di attacco è assai meno rigidamente regolata in italiano, dato che ogni consonante può comparirvi, compresi i legamenti (cf. [wɔ]vo, [jɛ]ri). Esistono tuttavia delle restrizioni combinatorie per i nessi tautosillabici in attacco. Per es., mentre occlusive e fricative possono essere seguite dalle liquide (pra.to, flo.ra), nasali e liquide ammettono dopo di sé soltanto i legamenti (liu.to, nuo.ra; non ad es. */nlV./, /lnV./). Sequenze di questo tipo sono semmai ammesse, come sopra indicato, con scansione eterosillabica (Car.lo, San.Remo; non *Santo Remo, come invece Santo Stefano). Va notato, a tal riguardo, che i legamenti e le liquide preceduti da consonante non possono occupare da soli l'attacco; ciò violerebbe una delle "leggi preferenziali" di buona formazione sillabica (→ paragrafo precedente), poiché si creerebbe una coda meno sonora dell'attacco seguente; si ha pertanto e.quo e li.tro (non */lit.ro/). Non a caso, nei nessi con legamento o liquida si è spesso avuta la

geminazione della consonante precedente (come in OCULU > OCLU > occhio /^hɔk.:jo/); infatti, mentre i nessi latini “muta cum liquida” – la cui scansione eterosillabica è spesso attestata dalla metrica – violavano la suddetta tendenza, quest’ultima è stata ristabilita in volgare. Le sequenze geminata + liquida/legamento non sono rare in italiano, perché tra le geminate distintive troviamo in tale posizione le occlusive e /f/: cf. sof.fio, sof.fro, cap.pio, lab.bro. Le geminate intrinseche subiscono tuttavia maggiori costrizioni, dato che possono tutt’al più essere seguite da un legamento (quasi sempre labio-velare). Si veda la tab. 1, in cui compaiono anche, per affinità con le altre affricate, le affricate palatali, che peraltro presentano la correlazione di quantità consonantica distintiva (cf. cacio /^hka.tʃo/ vs. caccio /^hkat.:ʃo/):

	ts	ɟ	tʃ	ɟʃ
_j	a[t.:s]ione	a[d.:z]ienda	-	-
_w	ca[t.:s]uola	gar[dz]uolo	la[t.:ʃ](u)olo	[dʒ](u)oco
	ʎ	ɲ	ʃ	
_j	-	-	-	
_w	fi[ʎ.:](u)olo	gra[ɲ.:](u)ola	u[ʃ.:](u)olo	

Tabella 1

Se precedute da altra consonante in coda, le geminate intrinseche si degeminano, accomodandosi nell’attacco (cf. tab. 2 e garzuolo in tab. 1). In particolare, /ʎ ɲ ʃ/ hanno, in tali contesti, una distribuzione molto ristretta, fatte salve le ampie possibilità combinatorie del clitico plurale gli (la tab. 2 elenca, per completezza, anche le affricate palatali):

	ʎ	ɲ	ʃ	ts	ɟ	tʃ	ɟʃ
r _	dar[ʎ]i	Gar[ɲ]ano	Mar[ʃ]ano	mar[ts]o	ar[ɟ]illo	mar[tʃ]o	ar[ɟʃ]ento
l _	-	-	Vol[ʃ]i	mil[ts]a	el[ɟ]eviro	fal[tʃ]e	al[ɟʃ]ido
n _	pon[ʎ]i	-	con[ʃ]o	can[ts]one	man[ɟ]o	pan[tʃ]a	man[ɟʃ]are
s _	-	-	-	-	-	dis[tʃ]inesi	dis[ɟʃ]elo

Tabella 2

Il fatto che i nessi italiani ostruente + liquida/legamento occupino la posizione di attacco è confermato, come sopra notato, dalla scelta dell’articolo. Restano peraltro

esclusi in posizione iniziale gli ardui nessi */t/ dl/, anche interlinguisticamente inconsueti. All'interno di parola essi offrono un altro caso di scansione indeterminata, potendo risolversi in modo sia eterosillabico (*at.lante*), sia tautosillabico (cf. la possibile gorgia fiorentina in [a.θ^llan.te]). Ulteriori sequenze rare s'incontrano in altre parole di importazione, come *tme.si*; la cui scansione ad inizio di enunciato differisce però da quella eterosillabica di *rit.mo*, come avviene in altri casi analoghi (cf. *psi.cologia* ma *Pep.si*, *pte.rodattilo* ma *cop.to*). La relativa estraneità alla fonotassi italiana di questi nessi è dimostrata dal fatto che in certe pronunce substandard si ha la tendenza ad acclimatarli mediante epentesi di vocale (*pissicologia*) o, intervocalicamente, assimilazione (*raddomante* per *rabdomante*). Altra sequenza inconsueta è /stʃ/ di *dis-cinesi* (cf. tab. 2), che compare anche ad inizio di parola in certe varietà di pronuncia (cf. *scentrato* e *scervellato*, regolarizzati peraltro in /ʃ/ nello standard). Si mantiene invece dappertutto /zɲ/ (cf. *sgnacchera*), altra sequenza creata da un prefisso. Quanto ai nessi triconsonantici in attacco, gli unici ammessi in italiano sono quelli ad inizio di parola del tipo /S/+ ostruente+liquida/legamento, come in *scrofa* o (con tre diversi luoghi diaframmatici) [spj] *etato*.

Il nucleo sillabico consiste tipicamente di una vocale. I dittonghi “discendenti”, ad es. /aj/, si distribuiscono fra nucleo (la vocale) e coda (il legamento; circa la convenzione consistente nell'indicare i legamenti postvocalici come semivocali /j̣ ụ/ e quelli prevocalici come semiconsonanti /j w/ → fonetica italiana). Peraltro, tali dittonghi rappresentano in italiano un problema non semplice, perché non è agevole distinguere fra dittongo e iato. Ad es., in pronuncia accurata /nej̣/ (preposizione articolata) differisce da /^lne.i/ (plurale bisillabico di *neo*); ma in pronuncia veloce – a parte la diversa vocale e il ruolo dell'accento – la struttura sillabica delle due parole viene a coincidere. Lo stesso dicasi di *fi.ni.i*, che può diventare bisillabo in pronuncia veloce alla stregua di *fà.rei* o *da.rai*; e non è certo scontato che un confine di parola, come in *chiosa* vs. *chi osa*, possa fornire un appiglio. Conviene dunque distinguere tra pronuncia accurata (a bassa velocità di elocuzione o in parola isolata o alla fine di sintagma intonativo) dove *po.i*, *lu.i* e *ma.i* sono bisillabi, e pronuncia veloce, dove le medesime parole diventano monosillabiche. Il loro oscillante trattamento è del resto attestato nella tradizione metrica italiana. Perfino gli iati contenenti vocali atone diverse da /i u/ possono essere realizzati, in pronuncia veloce, come dittonghi anomali, ad es. in *ster[ɛo]tipato* o *c[qa]lizione* (cf. Gili Fivela & Bertinetto 1999). Ciò non deve sorprendere, perché il ruolo di semivocale può essere ricoperto da qualsiasi vocoide, non soltanto dai tipici legamenti che tradizionalmente creano dittongo in italiano. In

certe parole funzionali contestualmente deaccentate può anche accadere che si riduca a semivocale una vocale normalmente accentata, come in *il m[jo]/[wo] libro*. Di solito, comunque, l'accento costituisce un valido discriminante fra dittongo (*fai.da, Lau.ra*) e iato (*fà.i.na, pa.u.ra*).

I dittonghi ascendenti costituiscono un caso solo in apparenza più semplice, con i legamenti /j w/ in attacco (talvolta unico elemento, cf. *ie.ri, uo.vo*) e la vocale nel nucleo. Marotta (1988) suggerisce peraltro una differenza strutturale fra i tipi /jV/ e /wV/, con il secondo da interpretarsi come nucleo sillabico complesso. Benché la maggior parte degli studiosi non adotti tale soluzione, è indubbio che /w/ costituisca un elemento fonemico dal comportamento non lineare. Lo dimostra la scelta dell'articolo davanti a parole inizianti per /j w/. Mentre /j/ seleziona l'allomorfo preconsonantico (*lo iodio, lo iato, la iella*, a fronte tuttavia del tradizionalmente preferito *l'ieri*), /w/ mostra un comportamento ibrido: cf. *l'uomo, l'uovo* – come *l'oro, l'oste* – vs. *il whisky, il walkie-talkie* (ma per taluni: *lo whisky* o, per converso, *l'whisky*). Molti preferiscono del resto *lo swatch* a *il swatch*, mostrando di considerare il nesso iniziale come un caso di “esse impura” (in contrasto con *il suono, il siero*; circa le oscillazioni nella scelta dell'articolo davanti a parole non autoctone, cf. Marotta 1993).

Anche per i dittonghi ascendenti possono esservi situazioni di incertezza circa la distinzione rispetto allo iato. Certe parole sono indicate nello standard come contenenti iato, e tali esse sono per la maggior parte dei toscani, mentre per altri parlanti contengono un dittongo. Oltre agli esempi indicati nel paragrafo precedente (*viòla, lacuale*), si vedano *vi.ale, attu.are, bi.ennale, manu.ale*. In pronuncia veloce, tuttavia, queste parole possono perdere una sillaba per dittongazione, anche se ciò non appare sempre altrettanto naturale (cf. infatti *bi.ennio, spi.are, sci.are, vi.abile, fri.abile*).

A differenza di altre lingue (come le germaniche), la sillabazione italiana è un processo post-lessicale, ossia soggetto a ricalibrazione in contesto; cf. *in alto* /i.'nal.to/, *la psiche* /lap.'si.ke/. Peraltro, data la rarità delle parole uscenti in consonante (a parte gli elementi funzionali, come preposizioni o articoli), accade spesso che si produca una geminata per evitare che la consonante migri interamente sull'attacco della sillaba seguente (cf. *tra[m.:]azzurro, su[b.:]atletico, lap[s.:]appuntito*; cf. Camilli & Fiorelli 1965). Quanto ai casi di sillabazione alternativa di sequenze fonotatticamente identiche ma diversamente articolate sul piano morfemico, in italiano sono eccezionali e non da tutti avvertiti (*su.blime* vs. *sub.liminale, sub.lunare*).

Riferimenti bibliografici

- Bertinetto, Pier Marco (1999). La sillabazione dei nessi /sC/ in italiano: un'eccezione alla tendenza 'universale'? In Benincà, Paola & Mioni, Alberto M. & Vanelli, Laura (curr.), *Fonologia e morfologia, Atti del 31 Congresso Internazionale SLI*, 71-96, Roma, Bulzoni.
- Bertinetto, Pier Marco (2004). On the undecidable syllabification of /sC/ clusters in Italian: Converging experimental evidence, in *Italian Journal of Linguistics / Rivista di Linguistica*, 16: 349-372.
- Camilli, Amerindo & Fiorelli, Piero (1965). *Pronuncia e grafia dell'italiano*, Terza edizione riveduta a cura di Piero Fiorelli, Firenze, Sansoni.
- Gili Fivela, Barbara & Bertinetto, Pier Marco (1999). Incontri vocalici tra prefisso e radice (iato o dittongo?), in *Archivio Glottologico Italiano*, 84: 129-172.
- Marotta, Giovanna (1988). The Italian diphthongs and the autosegmental frame work, in Bertinetto, Pier Marco – Loporcaro, Michele (curr.), *Certamen Phonologicum*, Papers from the 1987 Cortona Phonology Meeting, 399-430, Torino, Rosenberg & Sellier.
- Marotta, Giovanna (1993). Selezione dell'articolo e sillabazione in italiano: un'interazione totale? In *Studi di Grammatica Italiana*, 15: 255-296.
- Turchi, Laura & Bertinetto, Pier Marco (2000). La durata vocalica di fronte ai nessi /sC/: un'indagine su soggetti pisani, *Studi Italiani di Linguistica Teorica e Applicata*, 29: 389-421.
- Vennemann, Theo (1988). *Preference Laws for Syllable Structure and the Explanation of Sound Change*, Berlin etc., Mouton de Gruyter.