

## La fonotassi statistica dell'italiano e del tedesco: i nessi consonantici

La fonotassi, o sintassi delle sequenze fonologiche, è un campo dell'analisi linguistica particolarmente rilevante per la caratterizzazione delle strutture linguistiche che permettono l'identificazione di una lingua tra le altre, e soprattutto per il suo specifico ruolo metalinguistico e psicolinguistico nella produzione e ricezione dei messaggi linguistici. In questo contributo si cercherà di mostrare come gli strumenti della linguistica quantitativa possono aiutare a fornire un quadro adeguato e fedele delle diverse stratificazioni che vanno a determinare il profilo fonotattico di una lingua, più conveniente rispetto alla rappresentazione offerta dagli approcci basati su regole. In particolare saranno osservati i nessi consonantici (tradizionale luogo di forte differenziazione dei sistemi fonologici delle lingue) presenti in forma testuale, permessi e vietati in due lingue, italiano e tedesco, così mettendo in luce le regole di restrizione che governano i diversi sistemi<sup>1</sup>.

---

<sup>1</sup> Lo studio sull'italiano e quello sul tedesco sono stati condotti in tempi diversi, ma con la stessa metodologia, senza tuttavia nessun obiettivo comparativo. Una parte dei dati forniti sulla lingua italiana da Isabella Chiari sono esposti e commentati nel volume *Ridondanza e linguaggio* (2002), soprattutto in relazione alla misura della ridondanza (che qui invece compare solo marginalmente), ove sono proposte diverse metodologie per il calcolo delle numerose forme di ridondanza linguistica. Del presente saggio il § 1 (Sistema e usi dei nessi consonantici), § 2 (Metodo di rilevazione e di analisi), § 3 (I nessi consonantici dell'italiano) e § 5 (Conclusioni) sono stati elaborati da Isabella Chiari; il § 4 (I nessi consonantici del tedesco) è stato elaborato da Silvia Castagna, come sintesi di una parte dei dati raccolti per la sua più ampia tesi di laurea su *La ridondanza fonologica in inglese e in tedesco* (La Sapienza, 2003). Si ringrazia inoltre Sabine Koesters Gensini per la lettura attenta del paragrafo sul tedesco. Mentre il metodo e gli strumenti usati per le rilevazioni del corpus italiano e di quello tedesco sono gli stessi, le tipologie testuali analizzate sono radicalmente diverse: nel caso dell'italiano il corpus è di lingua parlata, nel caso del tedesco è di lingua scritta. Non vi è dunque alcun intento comparativo nell'accostamento dei diversi materiali. Lo scopo è invece quello di mostrare come il metodo quantitativo possa render

## 1. Sistema e usi dei nessi consonantici

La presenza di restrizioni alla combinazione dei fonemi in sequenza è uno dei caratteri che in modo più evidente caratterizza la forma delle lingue. Le restrizioni, di qualunque tipo esse siano, generano uno scarto tra le combinazioni teoricamente possibili e le combinazioni effettivamente attualizzate in una data lingua<sup>2</sup>. Tale scarto definisce un tipo di ridondanza che chiameremo distribuzionale (cfr. Chiari, 2002). Ogni forma di restrizione produce dunque ridondanza distribuzionale e ha l'effetto di rendere possibile una parziale prevedibilità nelle sequenze articolate. Vi sarà una ridondanza distribuzionale sillabica, derivante dalle limitazioni nelle combinazioni di sillabe e nella formazione delle sillabe stesse; una ridondanza distribuzionale morfologica, sintattica e anche fonologica. La fonotassi da questo punto di vista può essere vista come un aspetto della ridondanza distribuzionale a livello fonologico.

In questo capitolo si osserverà un particolare aspetto della fonotassi dell'italiano e del tedesco, ossia la strutturazione e l'uso dei nessi consonantici in testi parlati e scritti. Il ruolo delle sequenze consonantiche è infatti particolarmente rilevante nell'identificazione, segmentazione e decodifica delle parole e degli enunziati (cfr. Cartwright e Brent, 1994; Coleman, 1996; Coleman e Pierrehumbert, 1997). Tale centralità è confermata anche da studi longitudinali sulle fasi di apprendimento del linguaggio, in particolare sullo sviluppo delle abilità discriminative dei suoni in neonati e bambini di pochi mesi.

La struttura delle sequenze fonotattiche (e dunque anche dei nessi consonantici) tuttavia non è omogenea e non è descrivibile con un numero molto limitato di regole di restrizione. Tale impossibilità dipende da diversi fattori, primo tra i quali l'evoluzione diacronica delle lingue che sovrappone strati evolutivi produttivi e popolari a strati più

---

conto di situazioni fluttuanti e gradualità, in modo molto più perspicuo rispetto alle descrizioni formali basate su regole.

<sup>2</sup> È possibile ulteriormente analizzare tale scarto in diversi livelli: combinazioni teoricamente possibili senza restrizioni di sorta; combinazioni possibili per l'apparato fono-articolatorio umano; combinazioni registrate in una qualunque lingua esistente; combinazioni registrate in una lingua data (tra queste ultime vi è una gerarchia d'uso rilevabile statisticamente). Il primo e l'ultimo livello sono più agevolmente identificabili, i livelli intermedi richiedono una complessità e ampiezza di dati più difficilmente valutabile.

conservativi (come per l'italiano testimoniano i cultismi, latinismi e grecismi, ripresi e messi in uso in epoche diverse). Agli strati naturali si sovrappone un altro fattore significativo: la presenza di prestiti non – o parzialmente – adattati da lingue straniere con cui il sistema linguistico è venuto in contatto nel corso dei secoli. Da quest'ultimo punto di vista, molte sequenze fonotattiche altrimenti vietate sono inserite a pieno titolo (e ad alte frequenze) nei testi parlati e scritti. La fonologia tuttavia, nella percezione (meta)linguistica dei parlanti, rimane spesso poco intaccata da questi contatti (certamente meno del lessico e della morfo-sintassi). I parlanti, pur producendo in effetti sequenze atipiche dovute a prestiti, riconoscono centralità a strutture tradizionali e marginalità – se non divieto – alle strutture atipiche.

La sovrapposizione di strati di origine e stratificazione diversa crea in ogni sistema fonologico una gradualità di nessi ammessi e vietati che non può essere colta se non in una prospettiva quantitativa, attraverso l'individuazione di diverse fasce d'uso suddivisibili per tipi attestati e solo successivamente per regole produttive, marginali o improduttive<sup>3</sup>.

## **2. Metodo di rilevazione e di analisi**

L'analisi dei nessi consonantici, sia per l'italiano che per il tedesco, è stata condotta su base testuale. Non si sono quindi osservate le forme nell'intero repertorio di forme possibili, ma nel modo in cui occorrono effettivamente negli enunziati e negli scritti prodotti, nelle forme testuali che si realizzano con frequenze molto diverse da quelle che possono essere registrate in un repertorio. Le sequenze fonologiche in questo modo non mostrano un'immagine falsata degli usi lessicali, permettendo una migliore rappresentazione delle fasce d'uso<sup>4</sup>.

---

<sup>3</sup> La distribuzione delle sequenze fonotattiche riflette indirettamente ma non sempre del tutto fedelmente il grado di integrazione dei diversi strati lessicali introdotti in una lingua. Da un certo punto di vista certamente la "spiegazione" dei fenomeni fonotattici risiede nelle caratteristiche del lessico. Tuttavia il punto di vista fonotattico ci mostra molto meglio gli appigli di struttura che guidano spesso molti scambi comunicativi soprattutto sul versante ricettivo.

<sup>4</sup> Le analisi condotte sui repertori, ossia su insiemi astratti (per esempio un vocabolario, anche nel caso riporti le forme flesse), non possono rappresentare l'uso effettivo. In un repertorio sono infatti registrate, poniamo, anche una gran quantità di forme rare e rarissime, che, come ben noto, possono costituire invece una porzione limitatissima, se non virtualmente assente, dei testi realmente prodotti.

I due corpora scelti per le analisi sono per tipologia molto diversi tuttavia di ampiezza pressoché uguale (circa 120.000 parole). Del dettaglio della composizione dei corpora si darà conto nei §3-4. In entrambi i casi tutto il materiale è stato tradotto in sequenze fonologiche, convertite in modo automatico in una prima fase, impostando una serie di *routines* grafema-fonema specifiche per le lingue analizzate e successivamente controllate e corrette manualmente<sup>5</sup>.

Per quanto riguarda invece la fase di interrogazione si è usato per entrambi i corpora un pacchetto di programmi di analisi testuale TACT (*Text Analysis Computing Tool*), versione 2.1 (1995), elaborato da John Bradley, Ian Lancashire, Lidio Presutti, Michael Stairs, e dal *Centre for Computing in the Humanities* dell'Università di Toronto. Le interrogazioni effettive sono state impostate e svolte con la formalizzazione in espressioni regolari, distinguendo la posizione del nesso nella parola: iniziale, mediana e finale.

Estratti i dati sugli incontri consonantici, si è proceduto all'ordinamento e alla classificazione del materiale in fasce di frequenza che mettersero in luce le regolarità e i differenti tassi d'uso dei nessi attestati.

### 3. I nessi consonantici dell'italiano

Sulla lingua italiana esistono diversi lavori che tentano una rappresentazione della fonotassi, incentrati su diversi aspetti: sulla sillaba (cfr.

---

<sup>5</sup> Come si è detto entrambi i lavori sono di impianto fonologico e non fonetico, quindi metodologicamente il problema più rilevante è stato il trattamento delle neutralizzazioni sia nel caso dell'italiano che del tedesco. In entrambi i casi si è proceduto a una trascrizione fonologica tradizionale, assegnando il fonema neutralizzato alla classe più vicina dal punto di vista articolatorio, evitando l'indicazione esplicita dell'arcifonema. Dunque la sequenza italiana [sp] viene assegnata al nesso /sp/, invece di segnalare con l'arcifonema la mancata distintività /Sp/. Questo sia per le sibilanti e nasali preconsonantiche dell'italiano, sia per le nasali preconsonantiche, gli assordimenti posizionali e le spirantizzazioni del tedesco. In realtà nel presente lavoro si tratta di un fatto puramente convenzionale, poiché sia nel corpus italiano che nel tedesco sono rilevate le posizioni dei soli nessi (non quelle assolute sulla presenza dei singoli fonemi), ossia di posizioni in cui vi è sempre neutralizzazione, i conteggi non cambierebbero, né risulterebbe falsata la relazione relativa tra i diversi elementi, attribuendo il singolo caso all'uno o all'altro fonema. Tutt'altra questione sarebbe se invece si volessero "contare" le frequenze assolute, nel qual caso sarebbe indispensabile una qualche attribuzione non convenzionale.

Bortolini, 1976 o Mancini e Voghera, 1994), sulla struttura e le restrizioni fonotattiche (anche Muljačić, 1972 e per certi aspetti Nespor, 1993), sui nessi consonantici (cfr. Klajn, 1967 e Bortolini e Zampolli, 1979)<sup>6</sup>. Di impianto specificatamente statistico è soprattutto il lavoro di Batinti (1993) che propone numerosi dati statistici su moltissimi aspetti della fonologia, pur tuttavia il lavoro si focalizza sulla statistica in un corpus (repertorio) derivato da lemmi di dizionario, rappresentativi del lessico, ma non della testualità con le sue diverse escursioni nelle frequenze d'uso. Il tentativo del presente lavoro è invece essenzialmente testuale, focalizzato sull'uso dei nessi consonantici calati in contesti enunziativi.

L'analisi per la lingua italiana è stata condotta su testi di lingua parlata, selezionati dal corpus che ha dato vita al *Lessico di frequenza dell'italiano parlato* (De Mauro, et alii, 1993), come rappresentativi di diverse tipologie parlate. I testi selezionati per l'analisi sono state tutte le registrazioni raccolte nella città di Roma (1/4 del LIP), per un totale di 129.056 parole (*tokens*). Dopo la conversione grafema-fonema si è ottenuto un corpus di 566.774 fonemi (*tokens*). I nessi sono stati conteggiati sia in posizione iniziale, interna e finale<sup>7</sup>.

L'inventario fonologico preso in considerazione ha incluso 19 fonemi consonantici /p/ /b/ /f/ /v/ /t/ /d/ /k/ /g/ /ts/<sup>8</sup> /tʃ/ /dʒ/ /m/ /n/ /ɲ/ /s/ /ʃ/ /l/ /ʎ/ /r/. I nessi consonantici totali analizzati nel corpus sono risultati 60.337 come si può osservare in Tabella 1.

---

<sup>6</sup> Esistono invece numerosi studi sulla composizione dell'inventario fonologico, alcuni dei quali integrati con elementi statistici. Per una discussione sull'inventario e una presentazione dei dati statistici sulla frequenza assoluta dei fonemi italiani, in un sotto-corpus del LIP vedere Chiari, 2002: 253-270 e il contributo di Albano Leoni e Clemente in questo volume (27-44).

<sup>7</sup> Per i nessi registrati al confine di parola si rimanda a Chiari (2002: 230).

<sup>8</sup> Trattandosi di una considerazione fonologica e non fonetica entro la categoria /ts/ sono stati inclusi entrambi gli allofoni [ts] e [dz]. Allo stesso modo le sibilanti in posizione preconsonantica essendo neutralizzate sono fatte rientrare nella stessa classe, così come le nasali preconsonantiche. Le approssimanti palatale e labiovelare invece sono state escluse dai nessi consonantici (e incluse invece tra i dittonghi, cfr. Chiari, 2002: 220-222), dato lo status incerto (Hall, 1971; Muljačić, 1972; Mioni, 1973; Serianni, 1988). Sono stati invece inclusi tra i nessi anche i cinque foni sempre lunghi in posizione intervocalica [ʃ ɲ ʎ ts dz], per omogeneità con le altre consonanti lunghe, in relazione al loro ruolo nella sillaba (diversamente da Chiari, 2002). L'inclusione delle consonanti lunghe nel conteggio è ovviamente dubbia, poiché tecnicamente non si tratta di nessi, tuttavia il dato è facilmente scorponabile.

**Tabella 1 - I nessi consonantici della lingua italiana**

Tipologia	Nessi attestati (tokens)				Nessi attestati (types)			
	totali	iniziali	interni	finali	totali	iniziali	interni	finali
Nessi biconsonantici	56.706	6.580	50.016	110	129	24	90	15
Nessi triconsonantici	3.631	269	2.972	-	38	5	33	-
<i>Totale</i>	<i>60.337</i>	<i>6.849</i>	<i>52.988</i>	<i>110</i>	<i>167</i>	<i>29</i>	<i>123</i>	<i>15</i>

Si procederà a ritroso, partendo dai nessi composti dal più grande numero di fonemi che presentano dunque maggiori restrizioni. Osserviamo che nessi composti da numerose consonanti sono rarissimi. I nessi formati da quattro consonanti sono assenti dal corpus e pure pochi sono quelli citati in letteratura<sup>9</sup>. Pochi sono i nessi triconsonantici in posizione iniziale e interna di parola. In posizione iniziale in particolare si manifesta una forte restrizione strutturale sui 267 *tokens* attestati: in prima posizione si trova sempre una sibilante (qui vi è neutralizzazione di /s/ e /z/). In seconda posizione compare sempre una occlusiva (/p//b//t//d//k//g/) o la fricativa /f/. Il terzo membro è sempre una liquida (laterale o vibrante alveolare), con una preferenza significativa per la vibrante (104 *tokens*, contro solamente 8 *tokens* con la laterale finale del nesso). Il nesso più rappresentato è /str-/ con 155 casi (*strada*), molto rappresentato anche in posizione interna (554 *tokens*).

In posizione interna di parola invece sono attestati più nessi triconsonantici (2.972 *tokens* su 33 *types*). Nel vocabolario della lingua italiana i nessi registrati sono molti di più (7.346 *tokens* su 68 tipi)<sup>10</sup>.

<sup>9</sup> Muljačić (1972: 98) segnala dodici casi di nessi quadriconsonantici nel lessico italiano (*superstrato*, *substrato*, *perscrutare*, *extra*, *feldspato*, *exploit*, *instradare*, *exclave*, *sanscrito*, *tungsteno*, *insclerosire*, *ad strato*). Una rapida occhiata al GRADIT ci permette di aggiungerne qualcuno. Solo per i nessi quadriconsonantici sopraccitati si arriva a circa 126 parole. Si tratta tuttavia quasi sempre di prestiti da lingue straniere o tecnicismi dei linguaggi settoriali (soprattutto dalla chimica).

<sup>10</sup> I dati sulla presenza dei nessi consonantici nel vocabolario sono tratti da un lavoro non pubblicato condotto sul *Dizionario della lingua italiana* di De Mauro (2000).

Come primo membro della tripletta possiamo trovare solo le nasali neutralizzate (1.310 *tokens*, soprattutto in 339 /mpr/ - *sempre* - 121 /mpl/ - *semplice* - 545 /ntr/ - *contratto*)<sup>11</sup>, le sibilanti neutralizzate (663 *tokens*, soprattutto con 554 /str/ - *nostra*) e le consonanti lunghe (428 *tokens* complessivi, soprattutto con l'occlusiva bilabiale sorda 66 /ppr/ - *rappresentante* - e 63 /ppl/ - *applausi* - e con l'occlusiva alveodentale sorda 136 /ttr/ - *quattro*) le liquide /l/ /r/ (38 *tokens* con la vibrante, 533 con la laterale, di cui 530 nella sequenza /ltr/ di *altro* e composti) Il secondo elemento è sempre una occlusiva o /f/ e il terzo membro sempre una liquida, soprattutto /r/ (2.570 *tokens* con la vibrante finale, solo 402 *tokens* con laterale finale del nesso).

Per quanto riguarda i nessi biconsonantici le attestazioni sono molto maggiori (56.706 *tokens*), con 6.580 nessi iniziali di parola, 50.016 nessi interni e 110 nessi finali (tutti derivati da forestierismi). Saranno di seguito riassunte brevemente le principali tendenze riguardanti questi nessi:

1. Le nasali sono molto produttive come primo elemento (in posizione neutralizzata) con 15.504 *tokens* per soli 15 *types*, seguite da quasi tutte le altre consonanti (eccetto /ʃ/ /ɲ/ /ʎ/ e le liquide /l/ /r/)<sup>12</sup>. Tra questi i nessi più frequenti sono quelli con la nasale seguita da occlusive alveolari (5.691 /nt/ e 2.666 /nd/), e seguita da occlusiva velare sorda (1.191 /nk/), con nasale alveolare lunga (1.135 /nn/) e con sibilante sorda (1.080 /ns/). Esempi sono *impero*, *interno*, *incontro*.
2. Le consonanti lunghe sono altrettanto frequenti (sia quelle con valore fonologico accertato, sia le cinque sempre lunghe in posizione intervocalica) con 21.579 *tokens* complessivi. Le più frequenti sono le consonanti lunghe con luogo di articolazione dentale o alveolare (/tt/ /ll/ /ss/ /tts/ /ddz/ e /nn/), per un totale di 14.499 *tokens*. Tuttavia questo caso rimane dubbio, non trattandosi tecnicamente di un vero e proprio nesso consonantico (cfr. nota 8). Esempi sono *sasso*, *palla*, *bocca*, *ragno*, *pazzo*.

<sup>11</sup> Per brevità di ogni tipo è dato solo un esempio, spesso il più rappresentato nel corpus. Dove è presente solo un caso viene esplicitamente indicato.

<sup>12</sup> L'esclusione delle palatali è generale per tutti i nessi, mentre le liquide sono trattate nelle tendenze successive.

3. Anche i nessi con la vibrante come primo elemento esibiscono tutte le combinazioni (eccetto che con le palatali), per un totale di 5.997 *tokens*, soprattutto se seguita da occlusiva (3.720 *tokens*), nasale (1.098) e laterale (466). Esempi sono *arto*, *sguardo*, *mar-mo*, *parlo*.

**Tabella 2 - Nessi consonantici italiani con frequenza tra 5.691 e 80 in ogni posizione**

/nt/	5.691	/mp/	748	/vr/	241
/st/	5.025	/sk/	715	/ng/	229
/tt/	4.426	/str/	709	/rv/	221
/ll/	4.271	/rd/	667	/mb/	209
/nd/	2.666	/rr/	639	/lm/	202
/ss/	2.514	/ddʒ/	633	/rtʃ/	193
/pr/	2.427	/gr/	610	/bl/	188
/tts//ddz/	2.153	/rm/	553	/ndʒ/	185
/rt/	1.513	/rn/	545	/dr/	160
/rk/	1.275	/ltr/	530	/vv/	150
/nk/	1.191	/ttʃ/	491	/skr/	137
/tr/	1.147	/rl/	466	/ttr/	136
/nn//mn/	1.135	/ntr/	454	/rg/	135
/kk/	1.111	/ʃʃ/	447	/zm/	132
/ms//ns/	1.080	/kr/	389	/mpl/	121
/lt/	905	/ff/	374	/rts/	120
/jɲ/	860	/lk/	352	/mbr/	109
/sp/	837	/mpr/	339	/rb/	107
/ʎʎ/	814	/mf/	284	/rdʒ/	87
/pp/	797	/fr/	279	/bbl/	84
/nts//ndz/	779	/ntʃ/	279	/dd/	83
/bb/	760	/mv/	267	/kl/	80
/mm//nm/	758	/br/	258		

4. Ancora nella fascia alta della produttività si trovano i nessi con le sibilanti come primo membro (qui neutralizzate) sia in posizione iniziale che interna di parola (4.469 *tokens*, esclusa la sibilante sorda lunga). Vi è una certa preferenza per le sequenze di sorde



con 4.311 *tokens* (3.528 /st/, 464 /sp/ e 319 /sk/), anche in posizione iniziale di parola (2.266 *tokens* totali, con 1.497 /st/, 396 /sk/ e 373 /sp/). Esempi sono *sperare, sdegno, asta, asma*.

5. Nessi con la laterale come primo membro sono abbastanza frequenti (1.731 *tokens*), con l'eccezione del nesso con occlusive bilabiali. In particolare la parte maggiore è giocata dal nesso /lt/ con 905 *tokens*. Esempi sono *alto, palco, elmo*.
6. Abbastanza frequenti sono anche i nessi con occlusiva o /f/ seguita da vibrante /r/, come già descritto in Klajn (1967) con 1.231 *tokens*. Esempi sono *treno, prima, quadro, acri*.
7. Relativamente frequente è anche la tendenza a formare nessi con /l/ come secondo membro preceduto dalle stesse consonanti che troviamo nella tendenza precedente (occlusive e /f/) con 723 *tokens*, in particolare il già citato nesso /rl/ con 466 *tokens* (tendenza 3). Tradizionalmente i nessi con /l/ e /r/ come secondo elemento sono assimilati in una tendenza unica, tuttavia i dati testuali mostrano una significativa differenza d'uso che privilegia i nessi con la vibrante. Esempi sono *merlo, blu, tarlo*.

In Tabella 2 si possono vedere i nessi bi- e triconsonantici più frequenti per tipologia (senza distinzione di posizione nella parola).

#### 4. I nessi consonantici del tedesco

Gli studi statistici e quantitativi precedenti della lingua tedesca compiuti a livello fonologico presentano come principale campo d'indagine il testo letterario. Forse il primo contributo di questo tipo è quello di Twaddell (1938); mentre R.D. King (1966) individua l'occorrenza di ogni singola unità fonematica in un *corpus* formato da parti di testi letterari di Thomas Mann, Franz Kafka e Theodor Storm letti ad alta voce. Anche David H. Chisholm (1987) compie uno studio statistico sul *patterning* fonologico di testi della letteratura tedesca del XX sec. usando come unità di analisi la sillaba.

L'intento dell'indagine presente sul tedesco è di descrivere statisticamente e quantitativamente la fonotassi della lingua tedesca d'uso, sulle orme della proposta già svolta per l'italiano e presentata nei paragrafi precedenti.

Per condurre l'analisi sui nessi consonantici tedeschi è stato costruito un corpus apposito, composto da 114.839 parole testuali e costituito da testi tratti dal quotidiano tedesco *Die Welt*. Con l'intento di analizzare la lingua tedesca contemporanea, sono stati presi in considerazione gli articoli scritti durante i mesi di febbraio e marzo del 2003. Il corpus così ricavato è stato sottoposto a una trascrizione fonologica semi-automatica<sup>13</sup>: dai grafemi consonantici si è passati ai fonemi che risultano essere in totale 369.362 *tokens*. La trattazione dei nessi consonantici tedeschi tiene conto solamente dei nessi iniziali, mediani e finali di parola, sono conteggiati anche i nessi che si trovano a cavallo dei confini morfologici<sup>14</sup>. Non si sono tenuti in considerazione gli eventuali nessi che si formano all'interno dell'enunziato al confine di parola.

I fonemi consonantici considerati<sup>15</sup> risultano essere 22: /p, b, t, d, k, g, pf, ts, tʃ, dʒ, f, v, s, z, ʃ, x, m, n, ŋ, l, r, h/<sup>16</sup>. I nessi consonan-

---

<sup>13</sup> Come si sa la lingua tedesca è particolarmente regolare nella corrispondenza tra struttura ortografica e fonologica standard: è stata dunque predisposta sequenza di conversione grafema-fonema, implementata automaticamente sul corpus, e successivamente rivista e corretta manualmente per tenere conto di alcune regole fonologiche di alternanza, delle neutralizzazioni, degli assordimenti e delle spirantizzazioni e dei prestiti da lingue straniere.

<sup>14</sup> La prospettiva che prende a unità di analisi la parola è classica, e trova dunque possibilità di comparazione con altri studi non solo sul tedesco: è implicito che anche l'unità sillabica fornisce una rappresentazione utile per illuminare gli stessi fenomeni, così come l'unità morfologica.

<sup>15</sup> Per la trascrizione fonologica sono stati presi come riferimento soprattutto gli inventari proposti da Trubetzkoy (1939); Philipp (1970); Mioni (1973); Wiese (1996); De Dominicis (1997); Ramers (1998).

<sup>16</sup> Come si può osservare in tale inventario le affricate compaiono come realizzazioni monofonematiche in accordo con Trubetzkoy (1939), nonostante per alcuni /tʃ/ venga considerato bifonemico, poiché tuttavia /ts/ è considerato monofonemico, per simmetria si è estesa la considerazione anche a /tʃ/ (cfr. Martinet, 1965); l'approssimante [j] non è stata inclusa in quanto variante allofonica di /i/ (cfr. Mioni: 1973); trattandosi di una prospettiva fonologica e non fonetica, tutte le varianti del fonema /r/ sono incluse nella stessa classe, così come le varianti di /x/. Non sono stati presi in considerazione nell'inventario fonologico lo straniero [ʒ], né [ʔ] cui solitamente non è attribuito valore fonologico. Come per il trattamento dell'italiano, anche nel tedesco le neutralizzazioni sono state trattate senza indicare l'arcifonema, ma mediante l'attribuzione del simbolo nel nesso al fonema più vicino dal punto di vista articolatorio (sia per le nasali preconsonantiche che per le desonorizzazioni e spirantizzazioni di <g>). Un caso particolare è il trattamento del nesso ortografico <NG>, in-

tici (*tokens*) presenti nel corpus sono in totale 73.655; di questi 62.651 sono biconsonantici; 9.357 triconsonantici; 1.465 quadriconsonantici; 173 formati pentaconsonantici; 9 esaconsonantici. Il numero di *types* realizzati dei nessi biconsonantici, indipendentemente dalla loro posizione nella parola, è 260; dei triconsonantici 621; dei quadriconsonantici 258; delle sequenze dei pentaconsonantici 62; degli esaconsonantici 4, come si può quindi osservare nella riassuntiva Tabella 3.

Per individuare le combinazioni fonotattiche “legali” e “illegali” e scoprire tra le prime le più produttive si inizierà con rilevare le sequenze fonologiche costituite da sei consonanti per arrivare infine all'analisi dei nessi biconsonantici. Per ragioni di spazio saranno illustrate solamente le tendenze generali della fonotassi consonantica del tedesco e si prenderanno in considerazione i nessi più significativi dal punto di vista statistico. La Tabella 4 illustra i 30 nessi consonantici più frequenti, mentre nella Tabella 5 sono riportati tutti i nessi con frequenza maggiore di 60.

Nel corpus compaiono 4 nessi formati da sei consonanti: in tutti compare come primo elemento una nasale: 3 presentano la nasale velare /ŋ/ seguita dall'occlusiva velare sorda /k/ e dalla sibilante sorda /s/. La sequenza /ŋks/ può essere seguita dalla fricativa /ʃ/ + /p/ + /r/; da /ʃ/ + /t/ + /r/ oppure da /t/ + /p/ + /s/. Il nesso /ŋksʃpr/ è attestato in *Bildungssprecher* /bɪldʊŋksʃprexə/ e in *Regierungssprecher* /regirʊŋksʃprexə/; /ŋksʃtr/ in *Entscheidungsstrukturen* /entʃaidʊŋksʃtrʊkturən/, *Hoffnungstreifen* /hɔfnʊŋksʃtraifən/ e in *Hochspannungsstromstöße* /hoxʃpanʊŋksʃtromʃtøʂə/; /ŋkstps/ in *Angstpsychosen* /aŋkstpsyxozən/. È ammessa anche la sequenza nasale /n/ + /d/ + /s/ + /ʃ/ + /p/ + /r/ nella parola *Vorstandssprecher* /foəʃtandsʃprexə/<sup>17</sup>. Analizzando i nessi formati da sei consonanti è quindi emersa la produttività della nasale velare /ŋ/ in posizione iniziale e del fonema /r/ in

---

terpretato in maniera controversa nella letteratura sull'argomento (cfr. la *g-deletion*, Wiese, 1996: 224-229). Le forme standard in cui a seguire il nesso ortografico vi sia una vocale centrale o indistinta è interpretato come /ŋ/; negli altri casi è interpretato come un nesso, dunque come sequenza di due fonemi (sottoposto a cancellazione di <g> dal punto di vista della realizzazione) (cfr. Mioni, 1973).

<sup>17</sup> È implicito che nel parlato spontaneo, dal punto di vista fonetico e non fonologico, questi nessi particolarmente impegnativi sono soggetti a fenomeni di riduzione molteplici.

posizione finale del nesso. Tendenzialmente i nessi così lunghi, abbastanza rari, contengono sempre confini morfologici<sup>18</sup>.

**Tabella 3 - I nessi consonantici della lingua tedesca**

Tipologia	Nessi attestati (tokens)				Nessi attestati (types)			
	totali	iniziali	interni	finali	totali	iniziali	interni	finali
Nessi biconsonantici	62.651	8.287	36.424	17.940	260	48	257	63
Nessi triconsonantici	9.357	325	7.927	1.105	621	7	618	44
Nessi quadriconsonantici	1.465	-	1.252	213	258	-	257	4
Nessi pentaconsonantici	173	-	173	-	62	-	62	-
Nessi esaconsonantici	9	-	9	-	4	-	4	-
<i>Totale</i>	<i>73.655</i>	<i>8.612</i>	<i>45.785</i>	<i>19.258</i>	<i>1.205</i>	<i>55</i>	<i>1.198</i>	<i>111</i>

Nelle sequenze formate da cinque consonanti, possono comparire in posizione iniziale del nesso le liquide /l, r/; le nasali /m, n, ŋ/, l'occlusiva velare sorda /k/ e le fricative /f, x/. Tuttavia la nasale velare /ŋ/ e le liquide /l, r/ sono quantitativamente privilegiate manifestandosi rispettivamente in 89, 32 e 28 *tokens*. La nasale /ŋ/ occorre in 19 *types* e tende ad essere seguita solamente dall'occlusiva /k/ + /s/ (è attestata una sola volta la sequenza iniziante con /ŋkʃ/). Il nesso più produttivo è /ŋkspr/ presente in *Stimmungsproblem* /ʃtmʊŋksproblem/. La consonante /r/ è attestata in 18 *types* e presenta come secondo elemento /p/, /k/, /ts/, /s/, /m/ o /n/. /r/ occupa la posizione iniziale nel nesso /rktpr/ presente in *Marktpreisen* /marktpraɪzn/. Il fonema /l/ occorre in 9 *types* ed è seguito da /p/ o /k/: è il caso della sequenza /lpstf/ della parola *Selbstfahren* /zɛlpstfərən/. In posizione finale del nesso, le liquide /l, r/ presenti rispettivamente in 34 e 89 *tokens* e in 12 e 35 *types*, sono le più frequenti.

<sup>18</sup> La manifestazione statistica dei nessi rari (e lunghi) risente di fatto della caratterizzazione lessicale del campione e delle sue eventuali idiosincrasie.

Nelle sequenze quadriconsonantiche il nesso più produttivo presenta come primo elemento la laterale /l/ seguita dalla sequenza /pst/; questo nesso si manifesta in fine di parola in *selbst* /zɛlpst/ ripetuta nel corpus 121 volte. In generale il fonema /l/, con frequenza uguale a 208 (*tokens*) in nessi quadriconsonantici, è molto frequente. Molto produttivo è il fonema /r/ che occorre come primo elemento in 336 *tokens* e 82 *types*. Il maggior numero di nessi (*types*) attestati si ha con /r/ davanti ai nessi triconsonantici che hanno come primo elemento l'occlusiva velare sorda /k/ (62 *tokens*; 20 *types*); la nasale /n/ (51 *tokens*; 6 *types*); le fricative /s/ (49 *tokens*; 14 *types*), /ʃ/ (49 *tokens*; 2 *types*) e /x/ (42 *tokens*; 9 *types*); la sequenza /rxʃn/ è attestata in *Durchschnitt* /dʊrxʃnɪt/. Le nasali /ŋ/ e /n/ sono le più produttive come primi elementi del nesso (*token*) manifestandosi 433 e 340 volte; frequente è l'aggettivo *künstlich* /kʏnstlɪx/. I nessi tendono a presentare come ultimo elemento il fonema /r/ (473 *tokens*; 105 *types*); /t/ (343 *tokens*; 26 *types*) e /l/ (158 *tokens*; 44 *types*). In posizione finale di parola tutti i nessi terminano in /st/.

**Tabella 4 - I primi 30 nessi della lingua tedesca per lunghezza e frequenza**

<i>Nessi esacons.</i>	<i>Nessi pentacons.</i>	<i>Nessi quadrics.</i>	<i>Nessi tricons.</i>	<i>Nessi bicons.</i>
/ŋksʃpr/ 4	/ŋkspr/ 18	/lpst/ 126	/rʃt/ 363	/nt/ 6.806
/ŋksʃtr/ 3	/ŋkskr/ 12	/ŋkst/ 97	/ntl/ 307	/st/ 3.413
/ŋkstps/ 1	/lpstb/ 10	/nsth/ 62	/ŋkt/ 304	/nd/ 3.094
/ndsʃpr/ 1	/ŋkspl/ 9	/ŋksf/ 54	/nst/ 289	/rt/ 2.953
	/ŋkʃtr/ 8	/ŋksm/ 44	/rst/ 277	/xt/ 2.334
	/ŋksʃt/ 7	/nʃtr/ 38	/ŋks/ 216	/lt/ 1.910
	/lpstf/ 6	/nstr/ 34	/rkt/ 216	/ʃt/ 1.655
	/ndspr/ 6	/rʃpr/ 33	/ʃtr/ 212	/kt/ 1.611
	/ŋksgr/ 6	/rxʃn/ 28	/ŋʃt/ 185	/rd/ 1.452
	/ŋkskl/ 5	/rstm/ 28	/ntv/ 177	/rn/ 1.391
	/lpstd/ 4	/ŋksv/ 25	/rtr/ 171	/pr/ 1.040
	/ŋksfl/ 4	/kstr/ 24	/kst/ 162	/ls/ 967

**Tabella 4** (*segue*)

<i>Nessi esaconsonantici</i>	<i>Nessi pentaconsonantici</i>	<i>Nessi quadriconsonantici</i>	<i>Nessi triconsonantici</i>	<i>Nessi biconsonantici</i>
	/ŋksʃp/ 4	/nʃpr/ 24	/str/ 144	/nts/ 946
	/lpstm/ 4	/ŋksb/ 23	/ntr/ 139	/gr/ 833
	/lpstʃt/ 3	/ntʃpr/ 23	/rkl/ 139	/rl/ 808
	/rktpl/ 3	/rnst/ 23	/nds/ 135	/tr/ 782
	/rktpr/ 3	/ŋksk/ 22	/xst/ 128	/rm/ 781
	/rstkl/ 3	/ŋksh/ 22	/ʃpr/ 119	/fr/ 781
	/rnsth/ 3	/ŋksr/ 21	/xsk/ 99	/rb/ 774
	/rnʃtr/ 3	/rnbr/ 21	/nft/ 92	/rk/ 766
	/ksʃtr/ 2	/ŋksl/ 20	/sʃt/ 83	/rg/ 764
	/ftsʃtr/ 2	/ŋksg/ 19	/ŋgl/ 78	/ft/ 746
	/nsʃtr/ 2	/rpst/ 16	/ŋkr/ 75	/nz/ 684
	/ŋksth/ 2	/rʃtr/ 16	/ksp/ 72	/rz/ 676
	/ŋkstl/ 2	/ŋksts/ 16	/rft/ 72	/br/ 670
	/ŋkstr/ 2	/ŋksp/ 15	/ntsl/ 71	/rf/ 657
	/ŋksgl/ 2	/rktf/ 14	/ndr/ 69	/ŋk/ 647
	/ŋksbr/ 2	/ndsf/ 14	/rbr/ 66	/kl/ 607
	/ŋksʃl/ 2	/ŋgfr/ 13	/rpr/ 60	/rh/ 555
	/lpsth/ 2	/nsth/ 11	/ŋʃl/ 58	/nf/ 553

Le tendenze della fonotassi tedesca riguardanti i nessi triconsonantici possono essere così schematizzate:

1. La sequenza più produttiva è formata da /r/ + /ʃ/ + /t/ attestata in 363 occorrenze. In generale il fonema /r/ è molto produttivo come primo elemento con 2.569 *tokens* e 136 *types* attestati; le uniche sequenze non ammesse sono /r/ + affricata /pʃ/<sup>19</sup> + consonante e /r/ + /z/ + consonante. Molto produttivi sono i nessi con /rʃ/ + C (494 *tokens*); /rk/ + C (482 *tokens*); /rs/ + C (334 *tokens*); /rt/ + C (318 *tokens*). Produttivi sono i nessi /rkt/ e /rkl/ che si realizzano in *Markt* /markt/ e *wirklich* /vɪrklɪx/.

<sup>19</sup> L'affricata /pʃ/ è molto rara e le uniche sequenze triconsonantiche attestate sono /m/ + /pʃ/ + C e /pʃdr/, /pʃʃt/, /pʃʃm/.

2. Molto frequenti sono le sequenze nasale /n/ + /t/ + /l/ (307 tokens) presente in *namentlich* /naməntliχ/ e /ŋ/ + /k/ + /t/ (304 tokens).
3. Si ha la tendenza a produrre nessi inizianti per nasale /n/ e laterale /l/, è il caso della parola *weltweit* /veltvaɪt/: questi occorrono rispettivamente in 2.544 e 638 tokens e in 80 e 87 types attestati. Anche in questo caso si evita l'incontro tra la laterale o la nasale e l'affricata /pf/ + C.
4. In posizione iniziale di parola compaiono come primo elemento solamente l'occlusiva velare sorda /k/; la fricativa /ʃ/ (*Straße*) /ʃtrasə/ e la sibilante sorda /s/.
5. in posizione finale di parola il numero maggiore di nessi presenta come primo elemento il fonema /r/ (408 tokens; 17 types); le nasali /ŋ/ (288 tokens; 2 types) e /n/ (274 tokens; 2 types); /l/ (32 tokens; 6 types). In questa posizione l'incontro più produttivo è formato da nasale velare /ŋ/ + occlusiva velare sorda /k/ + occlusiva dentale sorda /t/ (*bringt*) /brɪŋkt/. A fine di parola non sono attestati nessi triconsonantici inizianti con occlusiva dentale /t, d/; con le sonore /b, g/; con le affricate /pf, ts, tʃ, dʒ/; con le fricative /v, ʃ, h, z/.
6. la lingua tedesca ammette in posizione iniziale del nesso tutti i fonemi fatta eccezione per l'affricata /dʒ/ e la fricativa /h/.

Per quanto riguarda i nessi biconsonantici sono state osservate le seguenti tendenze:

1. La sequenza in assoluto più produttiva è formata dalla nasale /n/ + occlusiva dentale sorda /t/ ripetuta 6.806 volte. Tale sequenza occorre in parole grammaticali molto frequenti, si trova in posizione finale nella congiunzione *und* /unt/ presente 2.686 volte.
2. Sono molto produttivi i nessi che presentano l'occlusiva dentale sorda /t/ come secondo elemento. L'incontro /st/ ha frequenza di 3.413 e si manifesta nella terza persona singolare del verbo essere *ist* /ɪst/ presente 1.086 volte; il nesso /rt/ occorre 2.953 volte. In generale sono molto produttive le sequenze /n, s, r, l, x, ʃ, k, p/ + /t/ attestate con frequenza superiore a 1.000.

**Tabella 5 - Nessi tedeschi con frequenza superiore a 60 in ogni posizione**

/nt/	6.806	/pt/	548	/nh/	251	/ʃr/	154	/rp/	97
/st/	3.413	/kr/	536	/ʃv/	249	/zl/	144	/ŋkst/	97
/nd/	3.094	/ns/	532	/lm/	241	/str/	144	/sz/	92
/rt/	2.953	/rts/	524	/fl/	241	/lz/	143	/nft/	92
/xt/	2.334	/ʃp/	516	/xn/	239	/ntʃ/	139	/th/	90
/lt/	1.910	/bl/	497	/ln/	237	/ntr/	139	/zr/	90
/ʃt/	1.655	/ks/	488	/kts/	225	/rkl/	139	/fn/	90
/kt/	1.611	/gl/	479	/lf/	224	/lp/	138	/mpf/	85
/rd/	1.452	/dr/	479	/ŋks/	216	/xk/	137	/ŋkst/	85
/rn/	1.391	/tsv/	471	/rkt/	216	/gn/	135	/sʃt/	83
/pr/	1.040	/rv/	459	/ms/	212	/nds/	135	/lʃ/	78
/ls/	967	/tst/	441	/ʃtr/	212	/ps/	128	/sl/	78
/nts/	946	/ʃl/	366	/zg/	209	/xs/	128	/ŋgl/	78
/gr/	833	/rʃt/	363	/tv/	207	/xst/	128	/lv/	77
/rl/	808	/rx/	352	/rtʃ/	199	/lpst/	126	/ŋkr/	75
/tr/	782	/rs/	346	/nm/	194	/ʃn/	125	/bv/	74
/rm/	781	/nl/	333	/sp/	185	/ʃpr/	119	/zb/	72
/fr/	781	/mt/	326	/nʃt/	185	/tsl/	113	/ksp/	72
/rb/	774	/ld/	325	/sk/	183	/fg/	111	/rft/	72
/rk/	766	/nv/	319	/mp/	182	/bg/	110	/mz/	71
/rg/	764	/ntl/	307	/lk/	178	/sf/	109	/ntsl/	71
/ft/	746	/ŋkt/	304	/xts/	178	/sts/	109	/lh/	69
/nz/	684	/rʃ/	300	/tʃl/	177	/zm/	109	/ndr/	69
/rz/	676	/nst/	289	/ntv/	177	/zv/	106	/ʃm/	68
/br/	670	/rst/	277	/ŋg/	173	/nr/	104	/rbr/	66
/rf/	657	/pl/	274	/rtr/	171	/fs/	101	/fts/	64
/ŋk/	647	/mb/	265	/kst/	162	/lts/	99	/tg/	63
/kl/	607	/nʃ/	264	/lg/	161	/kn/	99	/xl/	63
/rh/	555	/tl/	260	/lb/	157	/xsk/	99	/tm/	62
/nf/	553	/nb/	256	/ŋx/	157	/ml/	98	/nstl/	62



3. Molto frequente è la sequenza formata dalla nasale /n/ e dall'occlusiva dentale sonora /d/ all'interno di parola (3.092 nessi) e si manifesta in parole come *Freundin* /frøyndɪn/. Tale nesso non compare a inizio di parola ed occorre solamente 2 volte in posizione finale. In posizione mediana di parola mostrano frequenza superiore a 1.000 gli incontri /s, n, r/ + /t/.
4. La laterale /l/ e il fonema /r/ sono tra le consonanti che entrano nel maggior numero di nessi anche se in generale l'incidenza dei casi con la laterale è inferiore rispetto a quelli di /r/.
5. In posizione iniziale di parola la sequenza composta dalla fricativa /ʃ/ e dall'occlusiva dentale sorda /t/ è molto produttiva (1.300 *tokens*); in generale in questa posizione sono ammesse le sequenze /ʃ/ + /p, t, m, n, l, r/ attestate per esempio in *stehen* /ʃteən/, *spulen* /ʃpulən/. Il fonema /r/ e la laterale /l/ tendono a comparire più frequentemente come secondo elemento del nesso attestati rispettivamente in 10 e 9 *types* e in 3.385 e 1.590 *tokens*.
6. In posizione finale di parola sono molto produttive le sequenze /x, s, n, r/ + /t/, la loro frequenza è superiore a 1.000; il nesso /xt/ è attestato in *recht* /rɛxt/.
7. Per quanto riguarda le restrizioni, in generale sono rari gli incontri in cui l'affricata /pf/ si manifesta come primo membro; sono del tutto periferici i nessi in cui l'affricata di origine straniera /dʒ/ occupa la seconda posizione; tale affricata e la fricativa /h/ non compaiono mai come primo membro; la nasale velare /ŋ/ mai come secondo.

Per completare l'analisi dei nessi consonantici della lingua tedesca, restano da raccogliere dati riguardanti specificatamente i nessi che si formano al confine di parola, e da un punto di vista invece fonomorfologico al contrario lo scorporamento dei nessi che si formano al confine di morfo (qui trattati insieme agli altri), prima di passare a una ulteriore prospettiva sulle strutture delle teste e delle code sillabiche.

## 5. Conclusioni

Nelle pagine precedenti si sono mostrati due esempi di individuazione quantitativa e dunque empirica delle principali tendenze della fonotas-

si di una lingua. Le caratterizzazioni quantitative sono state usate in entrambi i casi in maniera immediata e semplice, adoperando esclusivamente i dati di frequenza per osservare qualitativamente le possibili tendenze. È ipotizzabile tuttavia un uso più articolato dei dati statistici per l'individuazione semi-automatica delle principali tendenze: questo presupponendo la costruzione di un database dei tratti caratterizzanti i fonemi in nesso, e valutando l'impatto statistico incrociando le variabili articolatorie.

L'individuazione delle fasce d'uso, qui presentata solo nelle tendenze più frequenti per le due lingue prese in esame, contribuisce a dare una fisionomia precisa alla struttura fonotattica di una lingua, non ne esaurisce però il quadro. Usando come punto di partenza le fasce d'uso è possibile fornire una base per la formulazione delle più astratte regole di restrizione (mentre è più difficile, e criticabile il processo opposto).

Attraverso l'individuazione delle fasce di produttività è inoltre possibile individuare una gerarchia nelle regole, e inoltre caratterizzare e studiare la cronologia dei nessi in relazione alle diverse porzioni di lessico che si stratificano nel tempo nelle lingue naturali. Questo aspetto potrebbe essere interessante anche da un punto di vista comparatistico nell'osservazione dei diversi esiti fonotattici in lingue imparentate.

Lo studio dei nessi consonantici inoltre presenta solo una faccia della fonotassi. L'altra faccia, ancora più rilevante per la caratterizzazione del profilo fonotattico di una lingua, è lo studio delle strutture e dei tipi sillabici e della loro attestazione nell'uso. Data la stretta connessione tra possibilità di occorrenza dei nessi consonantici a inizio e fine parola e struttura delle teste e delle code sillabiche, lo studio dei nessi consonantici *tout court* resta una prima importante tappa verso la descrizione degli usi e della produttività dei tipi sillabici.