

## I francesismi dell'italiano: un confronto fra i principali vocabolari dell'uso

Il presente contributo si propone di segnalare elementi utili alla riflessione su compiti, finalità ed eventuali limiti di un vocabolario dell'uso in edizione digitale. Tale strumento, con le nuove possibilità di ricerca addotte, è spesso scelto come base di indagini di carattere lessicologico; il lavoro proposto in questa sede, sperimentando alcune fra le tante possibilità di ricerca garantite dalla computerizzazione, mira a scoprire fino a quale punto un vocabolario può essere considerato autorità in materia di lessico, testando la sicurezza e l'univocità dei dati forniti dai vocabolari più celebrati. Le osservazioni di natura metalessicografica e gli spunti proposti sono emersi nel corso di una ricerca di interlinguistica che indaga l'influsso lessicale della lingua francese sull'italiana; la ricerca dei francesismi è condotta attraverso l'edizione digitale di quattro tra i principali vocabolari dell'uso: il *Gradit*, *Lo Zingarelli*, *Il Devoto-Oli* e *Il Sabatini-Coletti*<sup>1</sup>.

Il vocabolario elettronico consente una ricerca immediata attraverso la quale individuare direttamente il lemma cercato e una ricerca avanzata; questa sfrutta la piena potenza del motore di ricerca e offre la possibilità di indagini lessicologiche per campi e ambiti diversi. Le parole, intrattenendo relazioni sintagmatiche e paradigmatiche tra loro, possono essere raggruppate sulla base di un comune denominatore; è possibile ricercare, per esempio, tutti i tecnicismi settecenteschi della politica di provenienza francese. Tali tipologie d'indagine, prima dell'informatizzazione, avrebbero richiesto uno spoglio manuale dei vocabolari pressoché inattuabile. La computerizzazione inoltre mette in evidenza alcuni aspetti del lavoro dei lessicografi che un in vocabolario in volume, pensato per la consultazione, non possono rilevarsi. I confini tra i diversi settori del lessico non possono essere netti e decisi, ma il linguista e il

<sup>1</sup> Da qui i repertori lessicografici sono più spesso citati in sigla; di séguito se ne dà lo scioglimento: D = Giacomo Devoto & Giancarlo Oli, *Il Devoto-Oli 2012*, a cura di Luca Serianni & Pietro Trifone, Firenze, Le Monnier, 2011; DELI<sub>N</sub> = Manlio Cortelazzo & Paolo Zolli, *Dizionario Etimologico della lingua Italiana*, edizione in vol. unico a cura di Manlio Cortelazzo & Michele A. Cortelazzo, Bologna, Zanichelli, 1999; G = Tullio De Mauro (a cura di), *Grande Dizionario Italiano dell'Uso*, Torino, UTET, 2007; GDLI = Salvatore Battaglia, *Grande Dizionario della Lingua Italiana*, dal 1971 diretto da Giorgio Bàrberi Squarotti, Torino, UTET, 1961-2002, 21 voll.; *Supplemento 2004*, *Supplemento 2009*, a cura di Edoardo Sanguineti, ibid.; LEI = Max Pfister & Wolfgang Schweickard (dir.), *Lessico Etimologico Italiano*, Wiesbaden, Ludwig Reichert Verlag, 1979 e segg.; S = Francesco Sabatini & Vittorio Coletti, *Il Sabatini-Coletti*, Firenze, Sansoni, 2011 (ristampa dell'edizione 2007); Z = Nicola Z, *Lo Zingarelli 2012*, Bologna, Zanichelli, 2011 (ristampa della XII edizione);

lessicografo che approntano i dati per una ricerca digitale sono obbligati dal calcolatore ad impartire istruzioni puntuali ed esaurienti, dichiarando in tal modo (anche indirettamente) il proprio indirizzo teorico-concettuale. Il lessicografo assegna ad ogni singola voce una o più etichette che il software riconosce e interpreta nell'elaborazione di una risposta alla richiesta del fruitore. Per identificare i francesismi settecenteschi della politica, per esempio, l'utente compila tre campi di indagine differenti e il programma incrocia tre ricerche diverse: il software controlla infatti, per ciascuna voce, la data di prima attestazione, l'etimologia e l'etichetta che riconduce ad un particolare ambito d'uso; questi elementi fungono da *tag* che il programma, in prima analisi, riconosce per poi selezionare e restituire come *output* solo le voci che soddisfano tutte le richieste.

Non è sempre agevole riconoscere un forestierismo né è sempre facile stabilire con precisione la prima attestazione di un termine ma la possibilità di esprimere tali riserve non è prevista da un calcolatore elettronico o, quantomeno, dal programma (sicuramente non tra i più sofisticati) che supporta la versione digitale del vocabolario. Il corretto funzionamento di un vocabolario elettronico non dipende solo dalle capacità del programma quanto anche dal previo trattamento dei dati che il software deve gestire. Nel caso dei vocabolari digitali, è il lessicografo che predispone i materiali per l'informatizzazione, assegnando ad ogni voce uno o più *tag*; è infatti il lessicografo a stabilire previamente che *giacobino* è di origine francese, entra in italiano nel Settecento ed è un tecnicismo della politica. Spesso però la ragione di molte scelte lessicografiche da parte delle fonti in esame non sempre è facilmente desumibile dai criteri operativi e metodologici dichiarati dai vocabolari. Il presente lavoro, per mezzo di alcune indagini nel dominio della linguistica contattuale, cerca dunque di descrivere il funzionamento e le potenzialità di un vocabolario dell'uso in edizione informatizzata e di individuarne i possibili punti deboli. Attraverso i vocabolari dell'uso, l'interferenza esercitata in diacronia dal francese sulla lingua italiana si osserva nei suoi effetti in sincronia sull'italiano contemporaneo.

All'avvio del programma, l'utente del vocabolario, compilando i campi di interesse in una apposita maschera di interfaccia, inserisce l'*input* (la sua richiesta); il programma 'legge' e interpreta la domanda, elabora i dati e fornisce in tempi brevissimi una risposta (*output*). Con la ricerca avanzata, ogni vocabolario consente l'esplorazione diretta del proprio *corpus* di voci e articoli: con l'ausilio degli operatori della logica proposizionale, di particolari parametri di ricerca e caratteri jolly si può comporre un comando di ricerca complesso, lavorando a tutto testo o restringendo il campo d'indagine ad una sezione in particolare (definizione, etimologia, datazione, ecc.). Generalmente è prevista anche la possibilità di ricerche complesse predefinite: considerando la probabile evenienza di alcune richieste, il lessicografo, in fase di elaborazione dello strumento elettronico, segna le tappe di un percorso di ricerca guidato e dispone preventivamente un *output*. Tale previa selezione è interessante da un punto di vista metalessicografico in quanto, analizzando l'insieme di lemmi stabilito dal curatore, confrontando i dati con il risultato di ulteriori ricerche condotte sullo

stesso e su altri vocabolari, è possibile rintracciare l'indirizzo teorico che ha mosso la discriminazione. Ciò che si riscontra, in primissima analisi, è che strumenti aventi gli stessi compiti e obiettivi e spesso (soprattutto in fatto di etimologia e datazione) anche le stesse fonti, nella classificazione dei propri lemmi compiono scelte significativamente distanti. I vocabolari elettronici, nonostante i campi interrogabili per ognuno siano tra loro molto simili, differiscono in realtà anche formalmente, cioè nei diversi programmi informatici che li supportano. Il confronto fra questi strumenti si snoda anche su aspetti tecnici che sembrerebbero dominio dell'informatica ma che in realtà offrono spunti di riflessione di natura linguistica e lessicografica; tuttavia, in questa sede, non è possibile indagare oltre il funzionamento di tali strumenti<sup>2</sup>.

I campi interrogati per le ricerche oggetto di questo intervento sono quelli relativi a datazione ed etimologia.

I vocabolari, interrogati sui francesismi dell'italiano moderno<sup>3</sup>, presentano i seguenti risultati: 3928 lemmi in G, 4114 in Z, 2602 in D e 3084 in S. Allo scopo di ottenere un elenco unico, il totale di 13728 voci è stato privato di quei lemmi che, presentati come francesismi da più d'un vocabolario, necessariamente si ripetono nell'elenco totale. Ad ogni singola voce sono stati successivamente assegnati uno o più *tag* che, in fase di elaborazione statistica dei dati, hanno permesso di ricondurre la voce al vocabolario o ai vocabolari che hanno marchiato come francesismo l'unità lessicale in questione. Dall'integrazione dei dati provenienti dai quattro vocabolari risulta un totale di 6905 voci. Dal confronto puntuale e sistematico degli elementi delle quattro liste, emerge che solo nel 18% dei casi un lemma è presentato come francesismo da tutte le fonti; il 13% circa dei lemmi è un francesismo almeno in tre strumenti, il 17% circa almeno in due vocabolari; ma nel 52% dei casi il francesismo è attestato come tale solo in una delle quattro fonti. I numeri appena presentati individuano una situazione di estrema variabilità nel lavoro dei lessicografi; in poco più della metà dei casi, infatti, i vocabolari sembrerebbero in totale disaccordo.

Per una interpretazione più attendibile del dato numerico appena presentato, si deve tener conto del numero totale di voci che costituisce il *corpus* di ciascun vocabolario. Date le frequenze assolute (che corrispondono al numero di francesismi per vocabolario) è possibile calcolare le frequenze relative, ossia il rapporto (esprimibile anche in percentuale) fra il numero di francesismi per vocabolario e il numero totale di lemmi di ciascuno di essi<sup>4</sup>:

<sup>2</sup> Cfr. De Blasi 2013.

<sup>3</sup> La ricerca esclude l'italiano antico: i dati sono relativi all'arco temporale che va dal 1525 all'anno di pubblicazione dei vocabolari scelti come base.

<sup>4</sup> Più 250 mila entrate per G, circa 143 mila per Z, circa 98 mila per D e circa 87 mila per S. Solo i lemmi di G e Z sono dichiarati in copertina; il totale dei lemmi delle altre opere è stato recuperato mediante una ricerca (a mascherina vuota) nei diversi vocabolari. D dichiara 150 mila accezioni, Sabatini 110 mila lemmi e accezioni.

	Frequenze assolute	Frequenze relative	Frequenze percentuali
G	3928	0,015	1,5%
Z	4114	0,029	2,9%
D	2602	0,026	2,6%
S	3084	0,035	3,5%

In relazione al proprio *corpus* di voci, il S è il vocabolario che presenta il maggior numero di francesismi; sembrerebbe dunque il vocabolario con i criteri più inclusivi riguardo ai fenomeni di prestito: il 3,5% del lessico della lingua italiana, secondo il S, è di provenienza francese. Segue Z, sul cui risultato è necessario soffermarsi per un'analisi più approfondita. Sull'alto numero di voci influisce soprattutto la mancanza di una preliminare selezione di francesismi: in Z, infatti, la ricerca predefinita per l'individuazione dei forestierismi dell'italiano, in realtà corrisponde ad una ricerca a tutto testo, volta alla selezione di tutte le voci nella cui stringa dell'etimologia compare l'abbreviazione *fr.* Questo significa che il lessicografo non ha operato una scelta netta, marchiando alcuni dei suoi vocaboli come forestierismi; ed è comprensibile che il risultato sia viziato<sup>5</sup>. È difficile, dunque, dare un'interpretazione puntuale del risultato di Z; una ricerca avanzata secondo parametri personalizzati avrebbe consentito di ovviare in parte al problema ma in uno studio del comportamento dei lessicografi è fondamentale evitare ogni elemento di soggettività. Senza contare che, per interrogare direttamente il *corpus*, è necessario conoscere il sistema di classificazione dei suoi elementi e il complesso di abbreviazioni e diciture utilizzate nella compilazione delle voci. Ma, scorrendo l'elenco dei francesismi di questo vocabolario, si contano quasi un centinaio di formule differenti, alcune semanticamente molto vicine tra loro, per descrivere l'etimologia di una parola e si riscontra la oggettiva mancanza di univocità nel sistema di abbreviazioni adottato<sup>6</sup>.

Si è detto che nel 52% dei casi un lemma è presentato come francesismo solo in una delle quattro fonti; in particolar modo 1602 lemmi (sul totale delle 6905 voci dell'elenco unico) sono francesismi solo in G, 1239 lemmi solo in Z, 255 solo in D e 456 solo in S. Anche in questo caso è opportuno individuare le frequenze percentuali

<sup>5</sup> L'abbreviazione *fr.* può essere infatti utilizzata in diciture che non raccontano necessariamente la storia di un prestito come «sul modello del fr.», «cfr. il fr.», «prob. atrav. il fr.», ecc.

<sup>6</sup> Di seguito si danno alcuni esempi: «prob. atrav. il fr.», «prob. attraverso il fr.», «forse atrav. il fr.» e «forse attraverso il fr.»; «sul modello del fr.», «sul modello del corrispondente fr.», «sul modello della vc. fr.», «sul modello dell'equivalente fr.», «sul tipo dell'equivalente fr.», «sul tipo del corrispondente fr.», «sul tipo del fr.»; «calcato sul fr.» e «calcato dal fr.» oppure «preso dal fr.» e «ripreso dal fr.». A meno che non si debba attribuire una particolare sfumatura di significato ad ogni espressione per ciascun elemento diverso utilizzato in essa, forse il numero di tali diciture potrebbe essere ridotto uniformando tra loro le espressioni simili dal punto di vista del significato o almeno utilizzando un unico e univoco sistema di abbreviazioni.

calcolate prima sul numero totale di francesismi di ciascun vocabolario e successivamente sul numero totale dei lemmi di ognuno:

	Frequenze assolute	Frequenze percentuali (sul totale dei francesismi)	Frequenze percentuali (sul totale dei lemmi)
G	1602	41%	0,62%
Z	1239	30%	0,85%
D	255	10%	0,26%
S	456	15%	0,52%

In questo caso, le frequenze assolute esprimono l'apporto speciale di ciascun vocabolario all'elenco frutto dell'integrazione dei dati dagli stessi ricavati; il contributo maggiore è quello di G, segue ancora Z. I numeri di G e Z si motivano in base al fatto che, come si è visto, si tratta dei vocabolari con il maggior numero di lemmi in totale; nell'interpretazione del dato di Z è necessario ricordare ancora che questo vocabolario elettronico non prevede all'effettivo una ricerca predefinita.

Le frequenze relative (in percentuale), calcolate prima in relazione al *corpus* totale di voci di ciascun vocabolario e poi in base al numero di francesismi di ognuno, esprimono in questo caso il contributo di ogni opera all'elenco finale di voci ma in relazione alla propria potenzialità (ovviamente la mole del *corpus*). Z mostra che nella valutazione del 30 % dei suoi francesismi (che corrispondono allo 0,85 % del totale dei propri lemmi) non c'è accordo con gli altri strumenti; ancora una volta il risultato si giustifica sulla base del fatto che il vocabolario offre un elenco di francesismi senza operare un discernimento tra tutti i lemmi che in qualche modo sono in relazione con un loro equivalente francese. Il dato notevole invece è quello di S che, non avendo un *corpus* ampio né una scelta indiscriminata di voci, risulta un vocabolario con criteri elastici in fatto di catalogazione dei forestierismi. Il 41% dei francesismi presentati da G non trova conferma negli altri strumenti ma, ancora una volta, è necessario specificare che, essendo il vocabolario dalla mole maggiore, ha molte voci che le altre opere invece non registrano; si tratta soprattutto di lemmi marchiati dallo stesso lessicografo come BU (di basso uso), OB (obsoleti) oppure LE (di uso letterario). Lo dimostra il fatto che il 41 % dei francesismi, sulla valutazione dei quali G non si accorda con gli altri strumenti, corrisponde allo 0,62% del totale dei suoi lemmi mentre, come già visto, il 30% di Z corrisponde ben allo 0,85% di tutto il suo *corpus*. In ogni caso, D si rivela il più cauto nella valutazione dei fenomeni di prestito e quindi il vocabolario le cui notizie più spesso trovano conferma negli altri strumenti.

Si è detto che il 17% circa delle voci in esame mette d'accordo almeno due fonti mentre il 13% delle voci trova compatti nel giudizio ben tre vocabolari. Allo

scopo di meglio evidenziare il comportamento delle opere in esame è conveniente cercare di capire quali strumenti siano maggiormente in linea tra loro nelle scelte. Ma prima si devono individuare tutte le possibili coppie creabili a partire da un insieme di quattro vocabolari; a tal fine è utile avvantaggiarsi di una formula matematica del calcolo combinatorio<sup>7</sup>:  $C_{n,k} = D_{n,k} / P_k = n! / [k! (n-k)!]$ . Per comparare i quattro vocabolari due o tre per volta, si cerca il numero di combinazioni diverse possibili; in termini matematici una combinazione semplice di quattro elementi di classe due (o tre). Sono dunque possibili sei coppie e quattro combinazioni di tre vocabolari per volta<sup>8</sup>.

I vocabolari più vicini nelle scelte sono Z e S: si è già detto più volte che l'elenco di prestiti di Z è onnicomprensivo di tutti i termini nella cui stringa dell'etimologia compare l'abbreviazione *fr.*; S, proprio perché d'accordo con il vocabolario – in questo caso – meno selettivo, si conferma il più disposto a riconoscere l'influenza francese su alcune voci della lingua italiana. Per la stessa ragione, anche D (che insieme a S registra il minor grado di accordo) riconferma la sua posizione di vocabolario meno generoso nell'identificazione di fenomeni di prestito. Raggruppando i vocabolari tre per volta, non si ottengono risultati interpretabili diversamente rispetto a quanto si è detto finora: G, D e S sono i meno concordi tra loro; anche se considerati due per volta, infatti, questi tre strumenti sono quelli meno in sintonia<sup>9</sup>. Il dato relativo alle voci che mettono d'accordo Z, D e S si legge alla luce del fatto che si tratta – rispetto a G – di vocabolari in generale più vicini tra loro, perché prodotti commerciali che, tenendo necessariamente conto l'uno dell'altro, tendono ad allinearsi nelle scelte.

La statistica descrittiva si limita a raccogliere, organizzare ed elaborare i dati emersi nel corso di una determinata indagine; i risultati della ricerca devono dunque essere successivamente interpretati alla luce delle conoscenze sul fenomeno indagato. Confrontando sinotticamente tutte le fonti, nella particolarità di ogni singola voce, è possibile recuperare la ragione di tale inopinata situazione di variabilità; in fatto di linguistica del contatto la questione ruota attorno alla distinzione fra prestito lessicale «au sens strict du term» (Buchi 2010: 11) e fenomeno di interferenza. Un prestito, che spesso sembra immediatamente riconoscibile se frutto dell'interferenza tra lingue geneticamente distanti, non è sempre e facilmente individuabile nel caso di influenza fra lingue imparentate. Nella storia dei rapporti tra francese e italiano si registrano dunque molti casi definibili come *borderline* (Bombi 2005: 14-40, 21), cioè al limite fra «prestiti autentici e creazioni autonome» (Gusmani 1993: 10); questo perché nell'imitazione di un modello è possibile procedere secondo diversi gradi di fedeltà.

<sup>7</sup> In questa sede ci si limita a ricordare che il calcolo combinatorio è quella branca della matematica che studia i modi per raggruppare e/o ordinare, secondo regole precise, gli elementi di un insieme finito di oggetti.

<sup>8</sup> I dati complessivi: G, Z, D, S 1243; G, Z, D 222; G, Z, S 241; G, D, S 91; Z, D, S 329; G, Z 170; G, D 164; G, S 146; Z, D 162; Z, S 450; D, S 82; G 1602; Z 1239; D 155; S 456.

<sup>9</sup> Formano le tre coppie più distanti nelle scelte: G, S 162; G, D 164; D, S 82.

I francesismi recenti<sup>10</sup> sviluppano nuove modalità di prestito che attestano il carattere oramai moderno dell'interferenza linguistica: si pensi ai calchi o ai forestierismi crudi. Il processo imitativo, quando coinvolge anche il piano formale, consiste sostanzialmente nel ricondurre alle strutture della propria lingua (ai diversi livelli interessati) elementi estranei al proprio sistema linguistico. Tra lingue affini, come il francese e l'italiano, la forza assimilatrice è maggiore e segue norme precise (cfr. Weinreich 1974: 5). I calchi, classificati – secondo la terminologia di Gusmani (1993) – in strutturali e semantici, presuppongono nel mutuante una maggiore consapevolezza delle strutture lessicali, morfologiche e sintattiche delle due lingue in contatto. Il passaggio tra le lingue sfrutta nella maggioranza dei casi i canali della lingua scritta<sup>11</sup>. Un calco strutturale presuppone che, in sincronia, il modello sia percepito come composto da «unità minori interpretabili nel loro significato e funzione» (Dardi 1992: 77) e che la lingua d'arrivo, dopo il prestito, ottenga un nuovo elemento lessicale. Non sempre è possibile una corrispondenza perfetta tra le strutture delle due lingue; spesso è necessaria una riproduzione infedele del modello di partenza (fr. *battant-l'oeil* > it. *battilocchio*; fr. *compte-rendue* > it. *resoconto*) o una trasgressione (in sincronia) di alcune norme della lingua ricevente (fr. *équipage* > it. *equipaggio* non suffragato da \**equipare*). Al contrario, i calchi semantici non individuano un incremento lessicale nella lingua ricevente ma un'estensione del significato di una voce già esistente. La motivazione dell'interferenza è individuata dal parlante che riconosce (o istituisce, nel caso di «calchi per falsa motivazione», cfr. Gusmani 1993: 223) un rapporto tra il modello e un elemento lessicale autoctono, sulla base di tratti semantici comuni. Come per la categoria di calco strutturale, anche per i calchi semantici si pongono non pochi dubbi interpretativi. Nulla vieterebbe, nel caso di usi metaforici o figurati, che lo spontaneo ampliamento di significato registrato prima in una lingua possa poi verificarsi (altrettanto spontaneamente) in un'altra che alla prima è strettamente imparentata; in altre parole è difficile escludere uno sviluppo semantico poligenetico. Tali tipologie di interferenza (in misura maggiore nel caso di calchi semantici) non sono immediatamente riconoscibili, tanto che per la loro individuazione si fa necessario il ricorso a criteri extralinguistici (commenti e dichiarazioni esplicite, traduzioni, compresenza delle varianti non adattate). Quando una parola è utilizzata in un nuovo contesto può avere uno sviluppo ulteriore e autonomo rispetto al modello di partenza. La novità introdotta dal calco semantico, prima della stabilizzazione lessicale, si affianca dunque ai significati precedenti, contribuendo alla polisemia della voce (es. *abbordaggio* nell'accezione 'avvicinare qualcuno, accostarsi a qualcuno' è calco semantico sul fr. *abordage*; il significato si aggiunge a quello marinairesco di 'manovra di affiancamento di due imbarcazioni, spesso come atto ostile'). Si registra anche il caso in cui il nuovo significato si sostituisce alle altre accezioni: *eleggibile* nell'accezione 'che, chi può essere eletto a una carica' (calco semantico dal fr. *éligi-*

<sup>10</sup> Secondo la terminologia usata (Zolli 1977: 62).

<sup>11</sup> Si rilevano ovviamente anche alcune eccezioni; si pensi a un calco strutturale imperfetto (cfr. Gusmani 1986: 244 e ss.) come *tavoletta* che riproduce impropriamente il fr. *toilette*. Circolava anche *teletta* che ne è invece il calco perfetto, *toilette* infatti è letteralmente 'piccola tela'.



*ble*) soppianta l'accezione 'che, chi può essere preferito; utile, conveniente'; in ultima analisi, il calco potrebbe avere un successo limitato al quadro socio-culturale che ha motivato l'interferenza, decadendo di pari passo al venir meno del proprio contesto; ne è un esempio la *vc. espettorare* che contemplava (cfr. Dardi 1992: 530) oltre al significato d'ambito medico anche l'accezione 'parlare a cuore aperto', oggi non più registrata nei vocabolari dell'uso.

Alla luce di quanto detto, per uno studio del comportamento dei vocabolari in fatto di forestierismi, è necessario analizzare approfonditamente le liste di lemmi ottenute, confrontando gli strumenti in esame sulle singole voci. I dati presentati finora fanno riferimento all'intero periodo che va dal 1525 al contemporaneo ma «la visione su tempi lunghi [...] può diventare approssimativa e perfino fuorviante se non è fondata su tagli sincronici» (Dardi 1992: 49); per tanto, si è fatto riferimento *grossomodo* al secolo XVIII, che per l'età moderna, è il secolo di maggiore influenza della lingua francese sull'italiana. Dall'integrazione dei risultati ricavati dai diversi vocabolari si è ottenuta una lista di 1085 francesismi settecenteschi dell'italiano contemporaneo; ad ogni voce corrisponde un articolo in cui le notizie provenienti dai vocabolari dell'uso sono confrontate con la documentazione recuperata da vocabolari specialistici per mezzo di ricerche più approfondite. Il corpus di voci analizzate è costituito quindi da tutti i lemmi contrassegnati come francesismi da almeno uno dei quattro vocabolari dell'uso<sup>12</sup>.

Per ogni voce è prevista un'area dell'entrata, un'area della semantica e un'area delle informazioni complementari che si compone di più sezioni dedicate, in generale, a considerazioni di natura etimologica e storica (indicazioni di data d'ingresso e prima attestazione). L'area delle informazioni complementari offre un ventaglio di notizie utili per la ricostruzione delle vicende che hanno caratterizzato la storia della voce. I dati recuperati dai vocabolari dell'uso, in questo modo, sono puntualmente messe a diretto confronto tra loro e con tutta la documentazione recuperata da vocabolari specialistici e altri saggi e lavori tecnici (DELI<sub>N</sub>; GDLI; LEI; Dardi 1992 e 1995; Hope 1971, ecc.).

Si è già detto che, attraverso le ricerche predefinite, i vocabolari forniscono indirettamente una descrizione dei propri criteri lessicografici e mostrano la propria personale idea di prestito linguistico. E si è visto come, sulla base dell'indagine statistica precedentemente descritta, sia stato possibile fare alcune osservazioni di natura metalessicografica sulla struttura e l'impostazione di ciascun vocabolario. Il confronto puntuale, reso più agevole dal lemmario appena presentato, ha permesso

<sup>12</sup> Sono state scartate solo alcune voci (presenti unicamente nell'elenco di Z): *budget, cottage, gentleman, humour, miss, pudding, sloop, standard, tartan* e *test* (voci inglesi) e *rancio* e *toreador* (voci spagnole). Si tratta di voci dalla veste grafo-fonetica straniera ma non francese. È stata eliminata anche la *vc. chicco* (il lemma è registrato da D): nella stringa dell'etimologia della voce non è pervenuto alcun riferimento alla lingua fr.; è lecito pensare ad un *bug* del software.



invece di osservare il comportamento di ogni lessicografo e di individuarne l'orientamento teorico in fatto di interferenza fra lingue.

Si rileva, per esempio, che G, D e, in misura minore, S fra i rispettivi risultati non includono tutti i calchi.

In questi vocabolari, per esempio, la voce *chiaroveggente* è data come composto di *chiaro* con valore di avverbio e *veggente*, participio presente obsoleto di *vedere*; solo in Z è presentato come calco dal francese *clairvoyant*; l'opinione di Z trova conferma in DELI<sub>N</sub> e nel catalogo di francesismi di Andrea Dardi (cfr. Dardi 1992: 272-273). La provenienza francese della voce, d'altronde, è testimoniata dalla compresenza nel sec. XVIII del forestierismo crudo. Lo stesso può dirsi per le voci *capoluogo*, *controsenso*, *manodopera*, *paracadute*, *parafulmine*, *portabandiera*, *portafoglio*, *prestanome*, ecc.

D e soprattutto S, pur non annoverando tutti i calchi fra i prestiti, danno notizia del fenomeno nella sezione di ogni voce dedicata all'etimologia. Questi elementi sono spesso giudicati derivati o composti, frutto di fenomeni interni alla lingua italiana. Benché l'esclusione dei calchi dal novero dei prestiti non sia sistematica (soprattutto in Sabatini), è possibile pensare che questi strumenti si limitino, come già anticipato, «aux emprunts lexicaux au sens strict du term» (Buchi 2010: 11).

Si è detto che il 52% dei lemmi della lista di voci è un francesismo in uno solo dei vocabolari: i francesismi pervenuti solo in G, trattandosi del vocabolario maggiore, sono perlopiù voci non registrate negli altri strumenti (nella maggioranza dei casi, parole obsolete); in Z, senza un filtro predefinito per l'individuazione dei forestierismi, tra i veri e propri francesismi si contano anche voci per cui si segnala solo una minima corrispondenza con un equivalente francese della parola. I francesismi di D e S che non hanno riscontro negli altri strumenti sono soprattutto franco-latinismi e franco-grecismi: tecnicismi della scienza che G e Z catalogano, invece, come composti della lingua italiana di formazione neo-classica.

G esclude anche i derivati o composti che hanno come base un forestierismo: *abbordare*, per esempio, è considerato un francesismo ma *abbordaggio* è qualificato come derivato del ben acclimatato *abbordare* con il suffisso *-aggio* (prestito morfologico dal fr. *-age*), oramai standardizzato e produttivo; il sostantivo è perciò considerato un deverbale frutto di un fenomeno tutto interno alla lingua italiana. La voce è presentata invece come prestito dal francese negli altri vocabolari dell'uso, in DELI<sub>N</sub> e nel lavoro di Andrea Dardi (Dardi 1992: 242). In questo vocabolario sono numerosi e diversi i casi di voci escluse dal conto dei francesismi e ricondotte a fenomeni di derivazione; lo stesso accade anche nel caso di vere e proprie famiglie di voci per cui, specialmente con le date di prima attestazione pressoché coincidenti, si potrebbe pensare ad un unico e forte fenomeno di interferenza che assieme al prestito del capostipite ha promosso la formazione dei suoi derivati: *federale* (G, Z, D, S), *federalismo* (Z, D, S), *federalista* (Z, D, S), *federare* (Z, S), *federativo* (Z, S), *federato* (Z), *federazione* (Z, S). Si sono individuati anche casi particolari come quello di *voyeurisme* (prestito integrale dal fr.) e *voyeurismo* che in G è considerato derivato da *voyeur* con

il suff. *-ismo* e non invece adattamento (anche parziale) del primo; secondo il vocabolario le due voci sono per giunta attestate per la prima volta contemporaneamente nel 1964; la prima attestazione di *voyeur* è del 1911.

In tutti i casi appena citati, G non specifica il tipo di interferenza linguistica ma si limita a consigliare un confronto con l'equivalente francese della voce sotto esame, attraverso la dicitura "cfr. il fr.". Questa espressione dal carattere generico non chiarisce però il grado di influenza esercitato dalla lingua francese sull'italiana né il risultato che tale interferenza produce nella lingua che subisce l'influsso.

Le differenze tra i vocabolari investono anche il campo della datazione oltre a quello dell'etimologia. Salvo che per rarissime eccezioni, G e Z propongono sempre una data precisa d'ingresso per ogni voce, ovviamente la data di prima attestazione della voce in un qualsiasi scritto dell'epoca. D e S preferiscono informazioni più generiche, indicando più spesso il secolo in cui il termine ha presumibilmente circolato, anche prima della consacrazione precisa in uno scritto<sup>13</sup>. Raggruppando i vocabolari due per volta è possibile osservare quali vocabolari si sono dimostrati maggiormente d'accordo<sup>14</sup>: G e Z segnalano una data specifica, D e S indicano invece il secolo in cui la voce è attestata per la prima volta. Si segnalano tuttavia alcuni casi in cui i vocabolari sembrerebbero in totale disaccordo anche sulla datazione dei lemmi. Un esempio può essere la voce *manioca* per cui i vocabolari, in fatto di prima attestazione, forniscono i seguenti dati: sec. XVIII G; 1549 Z; sec. XVIII D; sec. XVI nella var. *mandioca*, sec. XVIII, nella var. *manioca* S. Come si evince facilmente, i vocabolari fanno riferimento a varianti diverse della stessa voce, in alcuni casi senza specificare. Nel caso di voci con più varianti grafo-fonetiche (spesso distribuite in diacronia), infatti, le più antiche, che non hanno avuto una significativa diffusione o che sono scomparse dall'uso, non sono messe a lemma e non sempre se ne dà notizia nella stringa dell'etimologia. Al momento di indicare la prima attestazione, come si vede nell'esempio precedente, è più spesso riportata, ma senza ulteriori indicazioni, la data d'ingresso della variante più antica (questo accade soprattutto in Z). In Devoto si preferisce, anche qui senza specificare ulteriormente, la data d'attestazione della variante più moderna. G segnala solo una data, specificando di quale variante si tratti. S, accanto alla data d'ingresso della variante in uso, spesso offre un'informazione ulteriore sulla data d'ingresso della variante più antica.

L'ausilio dell'elettronica ha permesso di quantificare l'apporto lessicale francese sulla lingua italiana dell'uso e di individuarne le stratificazioni in diacronia; è stato possibile inoltre osservare che le variabili possibilità operative previste dai software

<sup>13</sup> Alcuni esempi: *aneddoto* (av. 1729 G e Z; sec. XVIII, D e S), *azionista* (1960, G e Z; sec. XVII, D; sec. XVIII, S), *badiana* (1797, G e Z; sec. XVIII, D e S); *galletta* (1771, G e Z; sec. XVIII, D e S), *gallico* (1713, G; av. 1798, Z; sec. XVIII, D e S); *marmitta* (1598, G e Z; sec. XVI, D; sec. XVIII, S); *ratina* (1727, G e Z; sec. XVIII, D e S); *vampiro* (1749, G e Z; sec. XVIII, D e S); ecc.

<sup>14</sup> G, Z 273; G, D 173; G, S 59; Z, D 31; Z, S 23; D, S 252. I dati si riferiscono ai soli francesismi settecenteschi.

di ricerca, i differenti criteri di classificazione e indicizzazione degli elementi del corpus, le diverse sfumature teoriche sull'idea di prestito linguistico, non garantiscono l'omogeneità dei risultati.

Un dizionario generale è uno strumento d'informazione di base che ha l'obiettivo di orientare il fruitore. Si tratta di strumenti in evoluzione che si rinnovano continuamente, e non solo accogliendo numerosi neologismi di edizione in edizione. Ma se in questi strumenti sono spesso fornite notizie dettagliate per grammatica, semantica e ambiti d'uso, altrettanto spesso, per quel che riguarda l'area delle informazioni complementari, la notizia è solo un'indicazione generica e talvolta poco soddisfacente, in sostanza un punto da cui partire per ricerche più approfondite.

Il catalogo dei fenomeni fin qui proposti e analizzati, può dare un'idea di come, nella compilazione di un vocabolario, sia necessario stabilire criteri frutto di previe osservazioni e riflessioni di natura concettuale. Affidarsi ad uno solo tra questi strumenti significa dunque accettare anche i presupposti teorici che ne motivano le scelte; ma attraverso il confronto tra fonti diverse è possibile individuare gli elementi di soggettività presenti in ogni opera e, a seconda dei casi, o riconsiderarli – integrando i risultati delle diverse fonti – o accettarli – preferendo motivatamente, se si posseggono criteri di valutazione, una sola redazione rispetto a tutte le altre. Disse Kaspar Schoppe: «le persone colte i dizionari li adoperano non come padroni, ma come servitori che aiutano la memoria»<sup>15</sup>. In questo senso, si potrà forse convenire che l'autorità in materia di lessico non vada attribuita ad uno solo tra i principali vocabolari in circolazione ma alla categoria stessa di vocabolario dell'uso.

Università del Salento

Francesca DE BLASI

## Bibliografia

- Ajello, Roberto (a cura di), 1977. *Interferenza linguistica*. Atti del Convegno della Società Italiana di Glottologia, Perugia, 24 e 25 aprile 1977, Pisa, Giardini editori e stampatori.
- Aprile, Marcello, 2004. *Le strutture del Lessico Etimologico Italiano*, Lecce, Congedo.
- Aprile, Marcello, 2005. *Dalle parole ai dizionari*, Bologna, Il Mulino.
- Aprile, Marcello, 2011. *I vocabolari etimologici italiani*, in: *SLI XXXVII*, Roma, Salerno Editrice, 5-38.
- Bergantini, Gian Pietro, 1740. *Della volgare elocuzione illustrata, ampliata, facilitata. Volume primo contenente A-B*, Venezia, Gian-Maria Lazzaroni.
- Bertini Malgarini, Patrizia, 1993-1994. *L'Italiano fuori d'Italia*, in: Serrianni/Trifone, vol. 3: *Le altre lingue*, 883-922.

<sup>15</sup> «Viri docti lexicis non velut magistris, sed tamquam ministris in memoriae subsidium utuntur». Si cita la traduzione di Tullio De Mauro (De Mauro 2005: 38).

- Bezzola, Reto Raduolf, 1925. *Abbozzo di una storia dei gallicismi italiani nei primi secoli (750-1300)*. Saggio storico-linguistico, Heidelberg, Winter.
- Bielfeld, Antje, 1996. *Methoden der Belegsammlung für das «Vocabolario della Crusca»*, Tübingen, Max Niemeyer Verlag.
- Bombi, Raffaella, 2005. *La linguistica del contatto. Tipologie di anglicismi nell'italiano contemporaneo e riflessi metalinguistici*, Roma, Il Calamo.
- Buchi, Éva, 1996. *Les structures du «Französisches Etymologisches Wörterbuch»*, Tübingen, Max Niemeyer Verlag.
- Buchi, Éva, 2010. *Bolchevik, mazout, toundra et les autres. Dictionnaire des emprunts au russe dans les langues romanes (Inventaire – Histoire – Intégration)*, Paris, CNRS Éditions.
- Cella, Roberta, 2003. *I gallicismi nei testi dell'italiano antico (dalle origini alla fine del sec. XIV)*, Firenze, Accademia della Crusca.
- Coluccia Rosario (a cura di), 1991. *Riflessioni sulla lessicografia*. Atti del Convegno organizzato in occasione del conferimento della laurea *honoris causa* a Max Pfister (Lecce, 7 ottobre 1991), Galatina, Congedo Editore.
- Coluccia, Rosario/Cucurachi Adele / Antonella Urso, 1995. *Iberismi quattrocenteschi e storia della lingua italiana*, in: *Contributi di Filologia dell'Italia Mediana IX*, 177-232.
- Currel Aguillà, Clara, 2009. *Diccionario de Galicismos del español peninsular contemporáneo*, Éditions de linguistique et de philologie, Strasbourg.
- D'Angelis, Antonella, 2007. Lucia Toppino (a cura di), *Tendenze attuali della lingua e della linguistica italiana in Europa*, Roma, Aracne.
- Dardi, Andrea, 1984. *Uso e diffusione del francese*, in: Formigari, 347-372.
- Dardi, Andrea, 1992. *Dalla provincia all'Europa. L'influsso del francese sull'italiano tra il 1650 e il 1715*, Firenze, Casa Editrice Le Lettere.
- Dardi, Andrea, 1995. *La forza delle parole: in margine a un libro recente su lingua e rivoluzione*, Firenze, Stabilimento grafico commerciale.
- De Blasi, Francesca, (2011/2012). *I francesismi settecenteschi dell'italiano contemporaneo*. Tesi di laurea magistrale in "Linguistica italiana", relatore: Rosario Coluccia, Università del Salento, a. a..
- De Mauro, Tullio, 1994. *Capire le parole*, Bari, Laterza.
- De Mauro, Tullio, 2005. *La fabbrica delle Parole. Il lessico e problemi di lessicologia*, Torino, Utet.
- Formigari, Lia (a cura di), 1984., *Teorie e pratiche linguistiche nell'Italia del Settecento*, «Annali della Società italiana di studi sul secolo XVIII», Bologna, Il Mulino.
- Giovanardi, Claudio, 1987. *Linguaggi scientifici e lingua comune nel Settecento*, Roma, Bulzoni Editore.
- Gusmani, Roberto, 1987. *Interlinguistica*, in: Lazzeroni, 87-114.
- Gusmani, Roberto, 1993. *Saggi sull'interferenza linguistica*, Firenze, Le Lettere.
- Hope, Thomas E., 1971. *Lexical borrowing in the Romance languages. A critical study of Italianisms in French and Gallicisms in Italian from 1100 to 1900*, 2 vol., Oxford, Blackwell.
- Korzen, Iørn / D'Achille, Paolo (a cura di), 2005. *Tipologia linguistica e società. Considerazioni inter- e intralinguistiche*. Due giornate italo-danesi di studi linguistici (Roma, 27-28 novembre 2003), Firenze, Cesati.
- Lazzeroni, Romano (a cura di), 1987. *Linguistica storica*, Roma, La Nuova Italia Scientifica.

- Leso, Erasmo, 1991. *Lingua e Rivoluzione. Ricerche sul vocabolario politico italiano del triennio rivoluzionario 1796-1799*, Venezia, Istituto Veneto di Scienze Lettere ed Arti.
- Marazzini, Claudio, 2009. *L'ordine delle parole. Storia di vocabolari italiani*, Bologna, Il Mulino.
- Marello, Carla, 1996. *Le parole dell'italiano. Lessico e dizionari*, Bologna, Zanichelli.
- Massariello Merzagora, Giovanna, 1983. *La lessicografia*, Bologna, Zanichelli.
- Michel, Andreas, 1996. *Vocabolario critico degli ispanismi siciliani*, Palermo, Centro di studi filologici e linguistici siciliani.
- Mioni, Alberto M./Cortelazzo, Michele A. (a cura di), 1992. *La linguistica italiana degli anni 1976-1986*, Roma, Bulzoni.
- Morgana, Silvia, 1993-1994. *L'influsso francese*, in: Serianni-Trifone, vol. 3, *Le altre lingue*, 671-719.
- Mormile, Mario, 1986. *Storia polemica tra italiano e francese (1200-1800)*, Roma, Il Bagatto.
- Orioles, Vincenzo (a cura di), 1988. *Modelli esplicativi della diacronia linguistica*. Atti del Convegno della Società Italiana di Glottologia, Pavia, 15-17 settembre, Pisa, Giardini editori e stampatori.
- Orioles, Vincenzo, 2002. *Percorsi di parole*, Roma, Il Calamo.
- Orioles, Vincenzo, 2006. *I russismi nella lingua italiana con particolare riguardo ai sovietismi*, Roma, Il Calamo.
- Pfister, Max, 1992. *Lessicologia e lessicografia*, in: Mioni/Cortelazzo, 293-308.
- Rigual, Cesáreo Calvo, 2007. *Panorama della lessicografia monolingue italiana attuale*, in: D'Angelis-Toppino 2007, 9-78.
- Russo, Michela, 2004. *La sezione degli etimi galloromanzi nel LEI. Descrizione e prospettive*, in «Rivista Italiana di Linguistica e Dialettologia», IV, Pisa-Roma, Istituti editoriali e poligrafici internazionali, 191-208.
- Serianni, Luca, 1981. *Norma dei puristi e lingua d'uso nell'ottocento nella testimonianza del lessicografo romano Tommaso Azzocchi*, Firenze, Accademia della Crusca.
- Serianni, Luca, 1984. *La lessicografia*, in: Formigari, 111-126.
- Serianni, Luca, 2003. *Il lessico scientifico nei dizionari italiani dell'uso*, in: Adamo-Della Valle 2003, 19-44.
- Serianni, Luca/Trifone, Pietro, 1993-1994. *Storia della lingua italiana*, 3 voll. Torino, Einaudi.
- Sessa, Mirella, 1984. *La terminologia delle arti e dei mestieri. Appunti su Alberti di Villanuova*, in: Formigari, 205-224.
- Toso, Fiorenzo, 2009. *Nuove retrodatazioni da testi del periodo rivoluzionario (1797/1806)*, *Lingua e Stile* 44/2, 209-230.
- Varvaro, Alberto, 1973. *Notizie sul lessico della Sicilia Medievale. I francesismi*, «Bollettino del centro studi filologici e linguistici siciliani», 72-104.
- Weinreich, Uriel, 1974. *Lingue in contatto*, Torino, Boringhieri.
- Zolli, Paolo, 1971. *L'influsso francese sul veneziano del XVIII secolo*, Venezia, Istituto Veneto di Scienze, Lettere ed Arti.
- Zolli, Paolo, 1976. *Le parole straniere*, Bologna, Zanichelli.
- Zolli, Paolo, 1977. *I francesismi dei secoli XVIII e XIX nei dialetti italiani* in: Ajello.

