

Attraverso la consultazione di un lessico di frequenza (per esempio di uno come il LIPSI, il *Lessico di frequenza dell'italiano parlato nella Svizzera italiana*, il meritorio lavoro di cui è autrice Elena Maria Pandolfi e al quale sono dedicate queste pagine), è possibile non solo, banalmente, individuare le parole – ed eventualmente le forme del loro paradigma – che occorrono con maggiore o minore frequenza nei testi ed eventualmente nei generi e sottogeneri di discorso documentati dal *corpus* da cui esso è stato tratto: se ne possono anche trarre informazioni in merito a tendenze generali nella strutturazione del discorso (saggiando, per esempio, come ha fatto la Pandolfi, la relazione numerica che intercorre tra le cosiddette *forme vuote* e quelle *piene*; o il rapporto tra elementi nominali ed elementi verbali; o la consistenza di elementi funzionali quali i segnali discorsivi o le congiunzioni testuali; o, di nuovo, almeno per i *corpora* che rendono disponibile un'interfaccia digitale per l'interrogazione, la frequenza di particolari elementi che hanno un significato referenziale).

Inoltre, attraverso la consultazione di un lessico di frequenza (soprattutto nel caso in cui, come in quello del LIPSI, siano state generate liste di supporto a quella principale dei lemmi e delle forme) si possono rilevare grandi tendenze strutturali, quali, ad esempio, quella allo sfruttamento delle entità polirematiche, o all'impiego di determinati tipi derivativi o compositivi, o alla lessicalizzazione di elementi verbali o avverbiali cliticizzati; possibile è anche l'esame degli elementi a bassa frequenza (e in particolare degli *hapax*), utile nell'esplorazione di sottocodici e registri, oppure – tra l'altro – quando si voglia indagare la variazione diatopica.

Tra gli strumenti messi a disposizione dalla linguistica computazionale, del resto, i lessici di frequenza sono forse i più potenti e i più euristicamente stimolanti; e il LIPSI – che da questo punto di vista non fa eccezione – si presenta come particolarmente interessante per almeno due ragioni: *in primis* perché rende (programmaticamente) possibile – sia pure non senza qualche aggiustamento – un raffronto sistematico con altri lessici affini, il LIP e il C-Oral-Rom (sui quali v. *infra* e le indicazioni in bibliografia); *in secundis* perché si inserisce di diritto nel quadro degli studi sull'italiano parlato – studi notevolmente ricchi di risultati a partire dall'ultimo trentennio del secolo scorso – documentando tra l'altro anche quella varietà mediale così significativa che si può etichettare come *trasmesso*¹.

Oggi, anche se non si tratta sempre di risorse commensurabili né – anche quando lo siano – facilmente comparabili, i *corpora* (e, in qualche caso, i lessici) a disposizione dello studioso sono molto numerosi e si tratta, in molti casi, non solo di risorse utili allo studio dell'italiano scritto, parlato e – almeno in parte – trasmesso, ma anche di raccolte statisticamente rappresentative. Molte di esse sono *generali* o comunque tali da documentare ampi domini linguistici (tra queste si contano il LIP o il CoLFIS, v. *infra*), dedicate allo scritto o al parlato, sia sincroniche (come il LIP) sia diacroniche (come il LIPS o il *corpus Repubblica: infra*); altre sono invece *specialistiche* e orientate alla descrizione di una sola varietà di lingua o di una piccola costellazione di varietà tra loro correlate (una di queste è LABLITA, che raccoglie solo testi di parlato spontaneo; un'altra il ROM, che riunisce campioni di parlato della varietà romana; esistono anche raccolte di testi di apprendenti l'italiano, come il già citato LIPS: su tutti v. *infra*).

¹ Sono da vedere, sull'argomento, almeno, AA.VV. 1976; AA.VV. 1985; AA.VV. 1997.

Per quanto riguarda il parlato, in particolare, ha costituito un modello per le raccolte successive il LIP (*Lessico [di frequenza] dell'italiano parlato*). Il *corpus* impiegato nella sua costruzione è stato costituito tra il 1990 e il 1992 da Tullio de Mauro e dai suoi collaboratori e le liste di frequenza che ne sono derivate sono state generate in collaborazione con la Fondazione IBM Italia, dando vita al primo lessico di frequenza dell'italiano parlato (cfr la descrizione che se ne fa in De Mauro, Mancini, Vedovelli, Voghera, 1993). Il *corpus* include 469 testi parlati (corrispondenti a circa 490.000 parole) raccolti a Milano, Firenze, Roma e Napoli e distribuiti in cinque classi fondamentali (ancora: De Mauro, Mancini, Vedovelli e Voghera, 1993): quella dello scambio comunicativo bidirezionale con presa di parola libera faccia a faccia (che comprende fattispecie quali le conversazioni in casa e sul luogo di lavoro); quella dello scambio comunicativo bidirezionale con presa di parola libera non faccia a faccia (tra le quali vengono annoverate conversazioni telefoniche anche registrate); quella degli scambi comunicativi bidirezionali con presa di parola non libera faccia a faccia (assemblee, dibattiti, interrogazioni, interviste radiotelevisive); quella dello scambio comunicativo unidirezionale in presenza del/i destinatario/i (lezioni, relazioni, comizi, omelie, arringhe); e quella dello scambio comunicativo unidirezionale o bidirezionale a distanza o differito (trasmissioni radiotelevisive). I materiali (testi e liste) si possono consultare, oltre che nell'edizione ETAS che accludeva due *floppy* sui quali erano riversati i testi originali, anche tramite il sito del BADIP (*Banca Dati dell'Italiano Parlato*), ospitato dal *Language Server* della *Karl-Franzens-Universität Graz* (Austria) all'indirizzo web <http://languageserver.uni-graz.at/badip/badip/home.php>; dei materiali digitali è possibile anche lo scaricamento.

Altri *corpora* relativi all'italiano parlato (si tratta spesso di risorse lemmatizzate da cui si sono generate liste di frequenza, in genere disponibili a titolo gratuito attraverso il Web) sono i seguenti, in ordine di pubblicazione/diffusione cronologicamente inverso:

- LABLITA: si tratta di un *corpus* – costituito da Emanuela Cresti e depositato presso il Dipartimento di italiano dell'Università di Firenze – che riunisce testi di parlato spontaneo raccolti a partire dal 1965 e aggiornati regolarmente. La collezione, che dispone di registrazioni digitali dei testi che la compongono, è interamente trascritta in formato CHAT, quello impiegato nel progetto Childes², ed è stato parzialmente pubblicato in Cresti, 2000 e in Cresti e Moneglia, 2005.
- ROM (*Trascrizioni di campioni di parlato della varietà romana*), a cura di Antonella Stefinlongo, Riccardo Cimaglia, Silvia Di Blasi: si tratta di un *corpus* che raccoglie, secondo quanto precisato dagli autori, «le trascrizioni ortografiche di alcuni campioni di registrazioni di parlato effettuate nel corso del progetto *Parlare italiano* dell'Università Roma Tre» nel 2005; i testi, pubblicati nel 2007, sono disponibili a partire da www.parlaritaliano.it/parlare/.

² Il nome *Childes* è acronimo di *Child Language Data Exchange System* e fa riferimento a un insieme di risorse per la trascrizione e l'analisi informatica del parlato infantile. *Chat*, invece, sta per *Codes for the Human Analysis of Transcripts* e indica il formato di codifica dei dati testuali del progetto Childes. V. MacWhinney, 1997.

- LIPS (*Lessico italiano parlato di stranieri*), a cura di Massimo Vedovelli, Alessandro Pallassini, Sabrina Machetti, Monica Barni, Carla Bagna, Simone Pieroni, Francesca Gallina: i testi che lo compongono, resi pubblici nel 2007 e scaricabili come quelli della raccolta precedente da www.parlaritaliano.it/parlare/, rappresentano «le trascrizioni dei testi tratti dall'archivio delle prove d'esame CILS - Certificazione di Italiano come Lingua Straniera dell'Università per Stranieri di Siena». Si tratta di circa 2000 oggetti prodotti da apprendenti dell'italiano L2 tra il 1993 e il 2006 che ammontano a circa 100 ore di parlato e che comprendono circa 700.000 forme lessicali. Il LIPS raccoglie sia scambi bidirezionali faccia a faccia con presa di parola libera, sia scambi unidirezionali *in praesentia*; nella loro trascrizione i curatori si sono attenuti alle norme LIP. Il *corpus* è interamente lemmatizzato, operazione che ha reso possibile anche la generazione di liste di frequenza e di uso. Il *corpus* consente di verificare attraverso quali fasi passino i discenti di italiano come L2 a partire dai livelli elementari.
- CPT (*Corpus di parlato telegiornalistico. Anni Sessanta vs. 2005*), curato da Antonella Giannini, Massimo Pettorino, Ilaria Vitagliano: il *corpus*, pubblicato nel 2007, è costituito dalla trascrizione ortografica di testi telegiornalistici registrati tra il 1966-1969 e da quella di un notiziario fittizio del 2005, tutti etichettati; suo fine è quello di rendere possibile “un confronto tra il parlato degli anni ‘60 e quello odierno”. La documentazione e la base di dati è scaricabile a partire dal sito www.parlaritaliano.it. Si tratta di un interessante complemento del *corpus Repubblica* (sul quale v. *infra*).
- CLIPS (*Corpora e lessici di italiano parlato e scritto*), a cura di Federico Albano Leoni, Francesco Cutugno, Renata Savy *et alii*: raccolto tra il 1999 e il 2004 e pubblicato nel 2006, il *corpus* comprende la trascrizione di testi parlati di varia natura (radiotelevisivi, da interazioni telefoniche, in contesti di *map task*, da lettura di liste di parole ecc.) per complessive 100 ore di parlato (corrispondenti a un milione di forme). La base di dati è stata parzialmente trascritta, segmentata e marcata dal punto di vista fonetico. Gli informatori provengono da Bari, Bergamo, Bologna, Cagliari, Catanzaro, Firenze, Genova, Lecce, Milano, Napoli, Palermo, Parma, Perugia, Roma e Venezia e i materiali sono accessibili a titolo gratuito, previa registrazione, all'indirizzo <http://www.clips.unina.it/it/>.
- C-ORAL-ROM (*Integrated reference corpora for spoken romance languages*), a cura di Emanuela Cresti: è il risultato di un progetto europeo quadriennale (2000-2004) che ha portato alla costituzione di una collezione di *corpora* di parlato spontaneo in lingua italiana, francese, portoghese e spagnola di notevole ampiezza (si tratta nel complesso di 123 ore di parlato per circa 1.200.000 forme e circa 300.000 lemmi). Il *corpus* è disponibile a pagamento su DVD insieme a una monografia descrittiva (Cresti e Moneglia, 2005).
- API (*Archivio del parlato italiano*), a cura di Federico Albano Leoni, Claudia Crocco, Renata Savy, Francesco Cutugno: pubblicato nel 2003, il *corpus* include testi registrati tra il 1997-1999 a Napoli, Bari e Pisa per complessive 35.000 parole. L'archivio è disponibile su DVD (Crocco *et alii* 2003) e può essere richiesto anche inviando un messaggio di posta elettronica a cirass@unina.it.

Per ciò che attiene all'italiano scritto, invece, si devono ricordare almeno il *corpus* LIF, il *corpus* VELI, il *Corpus Italiano scritto L2*, il CoLFIS, il CORIS-CODIS e il *corpus Repubblica*. In dettaglio:

- il *corpus* LIF (*Lessico di frequenza della lingua italiana contemporanea*), raccolto ed elaborato dal *Centro Nazionale Universitario di Calcolo elettronico* per la costituzione del lessico omonimo e reso pubblico nel 1971 (Bortolini, Tagliavini e Zampolli 1972), è stato il primo tra quelli orientati a documentare le caratteristiche dell'italiano scritto. Il *corpus* raccoglie testi teatrali, narrativi, cinematografici, di istruzione e di informazione datati tra il 1947 e il 1968 per circa mezzo milione di occorrenze. Sulla base del LIF è stato compilato il *Vocabolario di Base* (VDB) di Tullio De Mauro.
- il *corpus* VELI (*Vocabolario Elettronico della Lingua Italiana*) – il secondo in ordine di costituzione tra quelli dedicati allo scritto – è un'ampia collezione di testi giornalistici (lanci di agenzia ANSA e pezzi pubblicati da *Il Mondo*, *Europeo* e dalla *Domenica del Corriere* tra il 1985 e il 1987) corrispondenti a poco più di ventisei milioni di forme, da cui è stato tratto nel 1989, sotto la guida di Tullio De Mauro, il VELI, un vocabolario elettronico/repertorio frequenziale di 10.000 parole, quelle che hanno maggiore rappresentanza.
- Il *Corpus Italiano scritto L2*, raccolto e analizzato da Miriam Voghera e Giusy Turco, è costituito, secondo quanto precisato dalle autrici, da testi scritti prodotti da apprendenti dell'italiano come lingua seconda presso l'Università di Greenwich, a Londra. Il *corpus* comprende un totale di oltre 20.000 parole ed è classificato in base al tipo di testo e al livello di competenza linguistica. In particolare, i testi che lo compongono sono fatti rientrare nelle tre categorie descrittivo, narrativo e argomentativo e distribuiti nei livelli di competenza *beginner*, *intermediate* e *advanced*. Tutti sono stati sottoposti a *tagging* sintattico e sono scaricabili gratuitamente, insieme a materiale esemplificativo e documentazione, da www.parlaritaliano.it/parlare/.
- Il CoLFIS (*Corpus e Lessico di Frequenza dell'Italiano Scritto*) è stato curato da Pier Marco Bertinetto, Cristina Burani, Alessandro Laudanna, Lucia Marconi, Daniela Ratti, Claudia Rolando e Anna Maria Thornton. Si tratta del risultato di un progetto cui hanno preso parte la Scuola Normale Superiore di Pisa, l'Istituto di Scienze e Tecnologie della Cognizione del CNR, l'Università di Salerno, l'Istituto di Linguistica Computazionale presso l'Unità Staccata di Genova del CNR e l'Università de L'Aquila. Il *corpus* di riferimento è costituito, secondo quanto dichiarano gli autori, da testi tratti da quotidiani (*La Repubblica*, *La Stampa* e *Il Corriere della Sera*), periodici e libri pubblicati tra il 1992 e il 1994 e consta di 3.798.275 ricorrenze lessicali. I testi sono stati selezionati ponderando «i dati dell'ISTAT relativi alle letture preferite dagli italiani, per ottenere un *corpus* che fosse il più possibile rappresentativo di ciò che gli italiani effettivamente leggono». Il materiale indicizzato è disponibile per lo scaricamento gratuito sia nel formato testo semplice, sia in quello proprietario di Microsoft, .mdb, all'indirizzo Web www.istc.cnr.it/material/database/colfis/.

- Il CORIS-CODIS (*Corpus di Italiano Scritto contemporaneo*), costituito nell'ambito di un progetto di ricerca avviato dall'Università di Bologna e coordinato da R. Rossini Favretti, si propone come un repertorio generale dell'italiano scritto, anche se è costituito prevalentemente da testi narrativi contemporanei (risalenti, vale a dire, agli anni '80-'90 del secolo scorso). Il *corpus* contiene circa 100 milioni di forme e dovrebbe essere aggiornato ogni due anni attraverso l'immissione di *sottocorpora* di monitoraggio linguistico. Al CORIS, che costituisce il *corpus*-base, si affianca il CODIS (*Corpus Dinamico dell'Italiano Scritto*) che contiene i medesimi testi del CORIS, ma che consente la selezione di *sottocorpora* tipizzati. L'accesso al *corpus* è possibile, gratuitamente all'altezza della stesura di queste note, attraverso la pagina Web http://corpora.dslo.unibo.it/coris_ita.html, previa richiesta scritta di consultazione.
- Il *corpus Repubblica* è stato creato dalla SSLMIT (Scuola superiore di lingue moderne per interpreti e traduttori) e raccoglie testi scritti giornalistici lemmatizzati per circa 380 milioni di parole; vi confluiscono i testi di tutti i numeri del quotidiano *La Repubblica* dati alle stampe tra il 1985 e il 2000. È interrogabile – anche per la produzione di concordanze – attraverso un'interfaccia Web al sito <http://sslmit.unibo.it/repubblica> previa registrazione gratuita.

Per ciò che attiene, infine, all'italiano come lingua di apprendimento, oltre al LIPS, in merito al quale si è scritto *supra*, si dovranno ricordare anche i seguenti:

- la *Banca dati di italiano L2*, consultabile su supporto digitale, che raccoglie testi parlati di immigrati e si fonda sui dati raccolti fra il 1985 e il 2000 da vari progetti condotti dall'università di Pavia e da quelle di Bergamo, Roma, Siena, Torino, Trento, Udine, Vercelli e Verona. Gli autori chiariscono che «il Cd-Rom contiene le trascrizioni delle interviste a venti soggetti di otto diverse madrelingue, per un totale di circa 120 ore di parlato; contiene inoltre una bibliografia degli studi compiuti sulla base di questi dati dai partecipanti al gruppo di ricerca e una bibliografia ragionata generale dei materiali prodotti in Italia per l'insegnamento dell'italiano come seconda lingua». Il disco può essere richiesto, dietro semplice rimborso delle spese di spedizione, al Dipartimento di Linguistica dell'Università di Pavia; informazioni si possono trovare all'indirizzo Web <http://lettere.unipv.it/diplinguistica/didattica.php>.
- *Valico*, un *corpus* di testi scritti che corrispondono a più di 500.000 forme che è stato raccolto dall'Università di Torino. La raccolta è interrogabile attraverso un'interfaccia Web a partire dall'indirizzo <http://www.corpora.unito.it/valico/valico.php>.

2. IL LIPSI, LE SUE CARATTERISTICHE, LA SUA STRUTTURA

Il LIPSI – *Lessico di frequenza dell'italiano parlato nella Svizzera italiana*, come si è visto – è il frutto del lavoro di Elena Maria Pandolfi all'interno di uno dei numerosi progetti di ricerca condotti dall'Osservatorio linguistico della Svizzera italiana (OLSI). L'autrice, che ha raccolto i materiali, ha trascritto i testi e ha prodotto le liste (poi faticosamente riviste manualmente), ha voluto produrre il lessico in un formato che rendesse possibili

raffronti con altri *corpora* di parlato, e segnatamente con il LIP e il C-Oral-Rom (v. il paragrafo precedente e cfr anche le indicazioni in bibliografia), a partire da un ampio *corpus* – aperto – di parlato spontaneo e di parlato trasmesso radiotelevisivo; gli informatori appartengono tutti a due aree ritagliate entro il *continuum* di quella italoфона della Svizzera: l'area del Ticino e del moesano da una parte e quella del bregagliotto e del poschiavino dall'altra; quelle, vale a dire, in cui stando ai dati messi a disposizione dall'Ufficio federale di statistica nel 2000, si ha una prevalenza di parlanti italiano.

Figura 2



Entro le due subaree geografiche, gli informatori sono stati selezionati sulla base del criterio «dell'essere nati e cresciuti nella Svizzera italiana e della reperibilità e disponibilità» (Pandolfi :31). Tuttavia, l'autrice sottolinea (*ibid.*) come, nel corso della selezione, «si *sia* cercato di avere parlanti» che fossero anche «rappresentativi di diverse classi sociali, di diversa cultura ed età».

I testi da cui è stato generato il LIPSI sono stati raccolti tra il 2004 e la fine del 2006, corrispondono a circa 60 ore di registrazione e ammontano a circa 400.000 parole, corrispondenti a un numero leggermente maggiore di forme e a circa 11.500 lemmi. Nel complesso, all'incirca un quinto delle parole del *corpus* (pari a un quarto delle ore di registrazione) proviene da testi di parlanti grigionesi; il resto da parlanti del Canton Ticino. Inoltre, *grosso modo* un sesto dei testi della raccolta ha origine radiotelevisiva; il resto proviene da conversazioni, registrate «quando è stato possibile a microfono nascosto e a volte invece a microfono a vista» (ivi, p.36), cercando comunque di elicitare una risposta non condizionata dal contesto attraverso la scelta di argomenti di qualche

impatto emotivo (facendo, cioè, discorrere gli informatori in merito a «da loro professione, la famiglia, la vita quotidiana, fatti sociali e culturali del paese di provenienza, eventi politici locali e di portata nazionale, viaggi e vacanze, eccetera»: *ibid.*).

Per la trascrizione dei testi – rinunciando a modelli come il CHAT (MacWhinney 1997), implementato invece nel *corpus* LABLITA (sul quale *infra* e v. anche Cresti 2000) – l'autrice ha scelto di impiegare un apparato minimale di segni: «non abbiamo adottato alcuna particolare tipologia di trascrizione. Non abbiamo segnalato con particolari segni grafici, se non sotto forma di interiezioni e solo nei casi più marcati, gli aspetti soprasegmentali del parlato, sovrapposizioni di turni, interruzioni, pause, esitazioni, fatti paralinguistici e legati all'emissione del suono (innalzamento del volume della voce o una risata, per esempio)» (*ivi*, p. 38). Si sono funzionalizzati all'individuazione degli enunciati i segni interpuntivi e a quella dei turni l'a-capo. Anche alcuni dati fonetici sono stati normalizzati, mentre «sono state mantenute nelle liste, con l'etichetta F (*fonosimbolo*) forme dotate di valore semantico preciso, come *eh* (assertivo, asseverativo), *beh*, *mah*, *ah* (assertivo, esclamativo)» (*ibid.*). Le parole in dialetto, trascritte seguendo le convenzioni di Moretti 1999, non sono state però inserite all'interno delle liste.

Il *corpus* è stato segmentato automaticamente e sottoposto a etichettatura e lemmatizzazione semiautomatica; le liste prodotte dagli applicativi impiegati nel trattamento dei testi, tuttavia, come si è scritto, sono state riviste a mano per eliminare gli inevitabili errori di classificazione categoriale e per arricchire la marcatura con indicatori relativi a caratteristiche sintattiche, morfosintattiche, sociolinguistiche e funzionali. Dalla lista principale dei lemmi/occorrenze sono state infine estrapolati nomi propri, sigle e simboli, fonosimboli e onomatopee, che sono confluiti in liste separate. Le polirematiche (o *entità polilessicali*, come le chiama l'autrice) sono state raccolte in un elenco a parte, ma sono presenti anche nella lista principale³.

Forse sarebbe stato utile non estrarre i nomi propri dalla lista principale (LIP e C-Oral-Rom, in effetti, come segnala l'autrice a p. 45, non li isolano dagli altri elementi lessicali); e ciò non solo perché – secondo le parole dell'autrice – «il lessico di frequenza» dovrebbe

“rappresentare ciò che è legato alla lingua e ciò che è espressione di una società e di una cultura», ma anche perché non si vede come si possa distinguere convincentemente ciò che fa parte della lingua da ciò che appartiene all'enciclopedia. E se è probabilmente vero che un nome come *Ada* (cito i primi nomi della lista) è poco utile a disegnare un quadro della cultura e della società elvetica, così non è per *Apollinaire*, *Alberto Sordi*, *Alessandro Baricco* e – perché no? – *Alex Britti*. Resta in ogni caso il fatto che, come osserva la curatrice, nomi e sigle «possono essere agevolmente re-introdotti nella lista di frequenza generale» (*ivi*, p. 55)”.

³ Il termine *polilessicale* o *espressione polilessicale* è giustamente preferito dalla Pandolfi a quello più corrente *polirematica*, che si è usato talora in questo articolo per semplice comodità d'uso; esso – del tutto corrente – appare in effetti poco giustificabile (“le unità polirematiche non veicolano più predicazioni distinte”, sottolinea Pandolfi); resta però vero che neppure gli elementi polilessicali rappresentano una predicazione unitaria, perché in molti di essi non si opera alcuna predicazione (così, per esempio, in *libro bianco o ponte di comando*).

Lemmi e forme appaiono uniti nella lista alfabetica principale (1.1 *Lemmi e forme*), le seconde elencate sotto ciascun *item* lessicale nella sua forma di citazione; un elenco dei lemmi in ordine di frequenza assoluta decrescente è poi contenuto nella seconda lista (1.2 *Lista dei lemmi*). I due elenchi principali sono seguiti da altri sei che elencano alfabeticamente, nell'ordine, nomi propri, fonosimboli e onomatopée, sigle, elementi polilessicali, avverbi in *-mente*, statalismi e forestierismi.

All'interno della lista principale, una serie di etichette indica, per i lemmi e – ove utile per le singole forme – la classe lessicale di appartenenza⁴, la struttura formantica nel caso di tipi verbali o avverbiali cliticizzati, la natura di elemento polilessicale, le caratteristiche di diffusione (Regionale Svizzera italiana e Regionale Grigioni italiano) e di provenienza nel caso di stranierismi.

Gli alterati sono ricondotti, come le forme flesse, a un lemma quando “poco lessicalizzati” (*ivi*, p. 51; è il caso di *chiesina*, ricondotto a *chiesa*), mentre appaiono come entrate autonome nei casi di importante divergenza semantica tra la forma-base e quella derivata (*ibid.*; è il caso di *lampadina* rispetto a *lampada*).

I nomi mobili corradicali sono stati lemmatizzati con entrate distinte per genere (*amico* segue, dunque, nella lista dei lemmi *amica* come lemma a se stante). La scelta lascia impregiudicato, in alcuni casi, il problema della *mozione di genere* (*ivi*, p. 52) che, se non si pone, com'è naturale, per i nomi “indipendenti” (*uomo*, *donna*), lo fa nel caso dei nomi che la grammatica tradizionale chiama *di genere comune*, per i quali la distinzione tra le classi non è possibile, non essendo presenti nelle liste del lessico etichette morfologiche diverse da quelle che indicano la parte del discorso (in *nipote*, ad esempio, il genere maschile è distinguibile dal femminile solo grazie alla marca esplicita di eventuali modificatori [*il nipote*, *nipote alto*]).

Le forme dialettali presenti nei testi del *corpus* non sono – lo si è scritto – registrate nelle liste; sono invece presenti, come è ovvio, i regionalismi, alcuni dei quali – che appartengono, come si vedrà, al sottocodice burocratico-amministrativo e che definiscono quello che l'autrice chiama, sulla scorta di Lurati (1976), *italiano federale* – sono *statalismi* e appaiono caratteristicamente e pienamente accettabili nella varietà svizzera dell'italiano.

La presenza di statalismi e di altre forme locali, caratterizzanti ma sostanzialmente prive di connotazione, insieme con la valorizzazione di una serie di fatti linguistici ed extralinguistici dei quali si scriverà *infra* nel paragrafo *L'italiano in Svizzera: una lingua “altra”?* fa sì che l'autrice proponga di considerare l'italiano svizzero come una varietà *statale* (nazionale) e non regionale dell'italiano, nello stesso modo in cui si considerano varietà nazionali dell'inglese quello britannico, quello nordamericano o quello scozzese.

I regionalismi sono poi marcati dalle etichette R (regionalismo ticinese e grigionese) o RG (regionalismo grigionese) e, in questo quadro, molto opportuna appare la scelta di mettere a lemma, anche entità polilessicali quali *in parte a* e molte altre delle quali si scriverà *infra*.

Il particolare interesse dei regionalismi raccolti nel LIPSI merita qualche considerazione aggiuntiva. Molte tra le forme considerate *ehvetismi* godono di un'ampia diffusione regionale e interregionale. L'interiezione *bon*, per esempio, molto frequente nei discorsi degli informatori (ricorre 587 volte e

⁴ A pagina 44 si conta però tra le etichette anche quella di *neologismo occasionale*.

ha rango 85) e marcata R, è ampiamente diffusa anche in Lombardia e in Piemonte; lo stesso si può dire per *ciola* 'accidenti' (e altro), diffusa anche in Lombardia (per quanto talora con fonetismo differente) almeno nei livelli più bassi del parlato popolare (oltre che, naturalmente, nel dialetto), anche se prevalentemente in situazioni di *code-mixing*.

In assenza di un lemma traduttore che affianchi le entrate – forse lo si sarebbe potuto inserire nella lista degli statalismi –, non è facile per il lettore italiano, soprattutto per il non specialista, identificare localismi che siano tali per deriva semantica (geomonimica)⁵ o in quanto neoformazioni (eventualmente geosinonimiche). Più semplice, evidentemente, riconoscere neoformazioni "assolute" [Petralli, 1990],⁶ anche se in mancanza di altri strumenti lessicografici può non essere facile risalire al referente. Così, può sfuggire l'esatta natura di un elvetismo semantico come *ghetta* 'calza di nylon per donna?' 'calzamaglia'? O come *generalista* 'generico' (forma che occorre per esempio in locuzioni come *medico generalista*, talora *generalista tout court*); e può sfuggire al lettore non altrimenti informato la ticinesità di *mantello* ('cappotto'), così come gli può essere difficile comprendere che una neoformazione *totale* quale la monorematica *autopostale* o la polirematica *corso di ripetizione* valgono, rispettivamente, 'autobus regionale della Posta' e 'chiamata annuale per il servizio di leva'⁷.

Molte sono le rilevazioni interessanti rese possibili dal LIPSI. Numerosi lemmi classificati tra gli statalismi sono geomonimi di parole italiane (e per questo anche *falsi amici*): *iniziativa*, ad esempio, che occorre 64 volte e ha rango 561, indica una proposta di legge di iniziativa popolare e corrisponde all'italiano *referendum propositivo*. *Apprendista* e *apprendistato* non indicano tanto il lavoratore che presti la propria opera, in genere a stipendio ridotto e comunque nel contesto di una disciplina contrattuale specifica, mentre apprende il suo mestiere, quanto piuttosto lo studente che frequenti la scuola secondaria e che abbia optato per un tirocinio attitudinale, impegnandosi così in una formazione in alternanza tra scuola e lavoro; in Italia lo si chiamerebbe forse *tirocinante*. *Conferenza* sta per 'commissione' ed *evidente* significa 'ovvio' (però nella locuzione *non è così evidente*: 'non è detto'; si tratta di un calco sull'espressione francese "*C'est pas évident*")⁸. *Domicilio* è la residenza legale illimitata concessa da un apposito permesso e può equivalere, in Italia a (*documento di*) *regolarizzazione*; *patrizio* è il membro di una *corporazione* locale di diritto pubblico cui compete l'amministrazione di beni comuni e che opera talora anche negli ambiti del sociale e della cultura.

Alcune tra le forme regionali documentate dal LIPSI sono elvetismi "assoluti": il più noto, forse, è il già citato *autopostale* 'autobus postale', voce dell'amministrazione che negli usi correnti è sostituito da quella scorciata *posta* (in effetti più frequente: 15 occorrenze contro 10, e un rango 1834 vs

5 Per fare esempi di parole non presenti nel *corpus*: *allocazione* 'sussidio', *brigadiere* 'generale di brigata'; nel *corpus* è invece documentato, insieme a molte altre forme, *dipartimento* 'ministero'.

6 Sono considerati *ticinesismi assoluti* forme che non hanno un riscontro né dal punto di vista del significante, né da quello del significato nell'italiano d'Italia. È tale, sempre per fare un esempio, *corso di ripetizione* 'richiamo annuale al servizio militare'.

7 A questi fini si può tuttavia consultare Pandolfi, 2006 o l'appena citato lavoro di Petralli e d'altra parte le caratteristiche fondamentali del lessico dell'italiano elvetico sono ben studiate e documentate.

8 Savoia e Vitale, 2002.

2426). Possono sembrare elvetismi assoluti anche i giovanilistico-studenteschi *espe* 'compito in classe' e *sore* 'prof' o il più trasversale *explo* 'scout' che hanno però anche diffusione insubrica e che derivano rispettivamente da *esperimento*, *professore* ed *esploratore*, il primo e il terzo troncati e il secondo aferesato (il primo, in particolare, è il risultato di una ridefinizione semantica, persuasa forse dal francese *expérience*). Elvetismo assoluto è certamente *buralista*, termine che indica l'impiegato sportellista in un ufficio postale di cui in qualche caso è anche responsabile; e lo sono numerose forme burocratico-amministrative, politiche e legali che rinviano a *denotata* caratteristici⁹. Forme interessanti sono *Natel* 'cellulare', acronimo di *Nationales Telefon* e, sempre per ciò che attiene agli oggetti della vita quotidiana, *cala* (anche *calla*, o *callaneve*), *spazzaneve*, che sarà certamente da collegare a *calada*, forma di vari dialetti lombardi che indica un attrezzo da traino per la rimozione della neve dalle strade, e a *calata*, forma (ma riflessa o popolare) dell'italiano regionale lombardo.

Anche molte tra le polirematiche presentano motivi di interesse: *caccia alta*, un elvetismo assoluto, ad esempio, non è l'inseguimento di prede sulle vette montane, ma piuttosto quello di prede pregiate (come i camosci)¹⁰. Molte poi sono le strutture polilessicali collegate a istituti amministrativi e a realtà politiche locali: il *Consiglio di Stato* – in Italia organo ausiliario del Governo dotato anche di funzioni giurisdizionali – è, in tutti i Cantoni eccetto che nei Grigioni, il Governo, mentre il *Gran Consiglio* è il parlamento cantonale; anche le etichette *Consigliere federale* e *Consigliere nazionale* rinviano a figure politico-amministrative elvetiche. Mentre, inoltre, la *cassa disoccupazione* corrisponde alla *cassa integrazione* dell'italiano d'Italia, la *cassa malati* indica l'assicurazione sanitaria obbligatoria gestita da uno degli enti privati che operano sul territorio svizzero. Hanno significato specifico anche *perequazione finanziaria* 'ridistribuzione perequativa dei proventi fiscali che opera a livello federale e cantonale'¹¹ e *scuola maggiore* 'scuola media'.

Tra le forme grammaticali sono da ricordare come locali gli usi di avverbi quali *già* e *ancora* o *volontieri* 'certo, grazie!¹²; di locuzioni avverbiali quali *da parte*, *dal lato*, *per solito*, *a corsa*, *a inaffiatoio* 'a pioggia' (detto per esempio di sussidi); o di locuzioni preposizionali come *per rispetto a*, *per rapporto a*, *da parte a* 'a parte'.

I forestierismi appaiono sia nella lista di frequenza fondamentale, sia in un elenco a parte; vi rientrano, precisa l'autrice, "anche quei lessemi che sono ampiamente acclimatati nel lessico italiano, come *sport*, *bar*, *hobby*", mentre "non sono stati considerati forestierismi i prestiti adattati al sistema morfologico italiano per mezzo di prefissi o

⁹ Se ne sono elencate alcune *infra*, nel capoverso dedicato alle polirematiche.

¹⁰ L'antonimo è *caccia bassa*, che si riferisce alla venazione di specie poco pregiate.

¹¹ *Perequazione* è ovviamente termine diffuso anche nella lingua della politica e della finanza italiana, ma la polirematica non mi pare lessicalizzata.

¹² *Volontieri* è noto anche all'italiano, ma si tratta di forma del tutto disusata se non in contesti marcati; il suo significato è, inoltre, differente da quello dell'omologo elvetico.

suffissi. Per esempio: *computerino, autostop, antidoping*¹³. Tra i forestierismi sono elencati anche i latinismi.

La lista include 479 elementi; il contingente delle forme a più alta frequenza è prevedibilmente quello degli anglicismi; seguono francesismi, germanismi, forme latine e, a grande distanza, ispanismi, nipponismi, turchismi e voci del persiano (turco e persiano entrano nel lessico con un solo vocabolo). Nel caso delle lingue meglio rappresentate, le parole più frequenti hanno lungo corso anche nell'italiano d'Italia (tra le prime cinquanta entrate si contano infatti anglicismi quali *film* 1905, *bobby* 1954, *sport* 1829, *okay* 1951, *computer* 1966, *bar* 1897, *bus* 1935, *weekend* 1905, *Sonar* 1950,¹⁴ *tunnel* 1839, *Hockey* 1927, *tennis* 1828, *partner* 1862¹⁵, *club* 1763¹⁶, *Radar* 1943¹⁷, *sketch* 1905, *stress* 1955, *marketing* 1957, *pullman* 1958, *record* 1884, *rock* 1957, *standard* 1764; vi appaiono anche anglolatinismi come *forum* 1956; francesismi come *camion* 1908, *garage* 1905, *dossier* 1895, *hotel* 1813, *cliché* 1837¹⁸; francolatiniismi come *referendum* 1892)¹⁹. Non mancano naturalmente nella lista voci entrate nell'italiano d'Italia in date relativamente recenti (per l'inglese, ad esempio: *trekking* 1979, *Internet* anni '90, *E-mail* anni '90, *down* 1982²⁰).

Tutti poco rappresentati nell'italiano d'Italia sono i germanismi, peraltro meno numerosi e meno frequenti rispetto alle altre forme alloglotte anche nel LIPSI. Poco numerosi sono i latinismi, tutti d'impiego normale anche in Italia (*tot, extra muros, super*); isolato appare un nipponismo, d'uso però comune nella Penisola (*tsunami* 1961).

Alcuni prestiti, d'altronde, non sono affatto diffusi nell'italiano d'Italia o hanno una connotazione spiccatamente colta; si tratta nella gran parte di forme che provengono dal francese e dal tedesco, come *mémoire* (che occorre 104 volte nel *corpus*), *à côté* 'e insieme, e in parallelo' (5 occorrenze), *buvette* (anche se il giornalismo politico lo usa talora in riferimento al bar degli elitari palazzi ministeriali), *classeur* 'raccoltore ad anelli'; oppure *Bonität* 'solvibilità', *Devise* 'valuta' (e *Devisenhandel, Devisentransaktionen*), *Zeitzone* 'fuso orario', *Hinterlegung* 'deposito', *Sonnenstube* 'Sala al sole', nomignolo del Canton Ticino, *Spätzli, Staubsauger* 'aspirapolvere', *Werkkleid* 'abbigliamento da lavoro', tutti documentati comunque con frequenza molto bassa. Gli anglicismi di questo tipo sono rari; tra questi si contano *skiman, truck, bachelor, Beamer* e la polirematica *working poor* 'lavoratore a basso reddito che non riesce a uscire dalla fascia della povertà', che si legge talvolta anche in Italia.

¹³ Mentre *computerino* è un alterato che mostra la produttività del prestito-base e *antidoping* un derivato neoclassico che può essere stato creato a partire dall'anglismo *doping*, *autostop* ha un antecedente francese noto al quale viene in genere ricondotto.

¹⁴ In effetti lessicalizzazione di un acronimo: *SOund Navigation And Ranging*.

¹⁵ Nell'accezione 'secondo di una coppia nello sport o nel teatro'.

¹⁶ Le accezioni settecentesche hanno senso politico, che sembra calcato sull'uso francese del termine inglese.

¹⁷ Acronimo: *RAdio Detection And Ranging*.

¹⁸ Nell'accezione 'lastra metallica usata nella stampa fotografica'.

¹⁹ La forma è stata introdotta in italiano attraverso la Svizzera.

²⁰ Si fa riferimento alla forma eponima, che può avere valore aggettivale; le altre – nominali, avverbiali – sono meno recenti.

3. L'ITALIANO IN SVIZZERA: UNA LINGUA “ALTRA”?

E bon anche a me m'andavano nell'angolo bon che si siedan lì [...] Eh chiaro Beh sì li abbiem fatti sedere cioè proprio poi ci son quelli che hanno proprio la palla [...] Dulcis in fundo gli orari per la palestra [...] Io prendo la posta tu prendi la posta Sì bon adesso hanno cambiato gli orari [...]
Okay siamo a posto qualcuno ha ancora qualcosa
[...] Cioè da sola non ce la faccio sì bon se c'è qualcuno in più sono perché non tutte le volte posso venire cioè adesso vado a scuola [...] E no è vero cioè tante volte perché cioè arrivi a casa da scuola che non stai bene perché sei rimbambito e tutto poi io sai che di malattie [...] Eh non vengo perché sono all'ospedale Sì adesso mi hanno operato adesso sono a posto [...]
Buono grazie ragazzi.

Lo stralcio di testo che precede è tratto da uno dei documenti raccolti nell'allegato digitale che accompagna Pandolfi (2006); si tratta di un brano di parlato conversazionale prodotto durante la riunione di un'associazione sportiva di Bellinzona il 10 gennaio 2005. Vi si riconoscono numerosi tratti genericamente marcati in diamesia e vari elementi propri dell'italiano dell'uso medio, magari nei registri informali. Si contano, tra i primi, a) la ricorrenza di segnali discorsivi con funzione emotivo-metadiscorsiva o fatico-conativa, che servono a rendere esplicito il *modus* del parlante in merito all'enunciazione, oppure a organizzare il flusso delle informazioni, o a gestire l'interazione con gli altri parlanti; b) la destrutturazione sintattica, secondaria alla progettazione *impromptu*; c) una certa ridondanza della segnaletica grammaticale funzionale all'indicazione della rilevanza e dello stato di attivazione dei referenti; tra i secondi, a) la presenza di qualche costrutto marcato e b) di qualche elemento moderatamente disfemico, oltre che – c) – di pochi stranierismi correnti, uno in funzione espressiva. Ciò che di regionalmente marcato vi riconoscerebbe un parlante italiano è veramente poco: praticamente solo l'ubiquitario *bon*, che ha, come si è osservato, coloritura settentrionale.

Qualche stralcio di testo mostra sicuramente un numero più alto di elementi connotati, come il seguente, tratto da un dibattito televisivo tenutosi il 23 novembre 2004:

G: Arrivederci questo è un'osservazione che viene dopo Il consiglio federale ha dato l'addio ha detto non ci interessa rispettivamente non è necessaria per affrontare i problemi linguistico culturali della Svizzera e credo che il parlamento ne riparleremo penso qui abbia un compito fondamentale e abbia già dato segni in questa direzione

D: Ma lei ha detto non si può credo che sia una frase che ha un contenuto politico chiaro non si può lasciare lo sviluppo alla deregolamentazione e al mercato delle lingue ma dobbiamo regolamentare Ora questa regolamentazione se non vado errato ha un costo Ora io vorrei capire se è proponibile sul piano politico e sul piano delle priorità un costo di questo tipo ma anche in termini e qui chiamerò in causa Gendotti di sforzo richiesto ai ragazzi perché è bello per il ticinese sapere tre lingue nazionali più o meno bene e in più l'inglese però è una grossa fatica e quando io ricordo che ai tempi lontanissimi in cui ero a Zurigo mi si diceva i nostri studenti al Politecnico come diceva il suo compianto collega Buffi sono costretti a correre con la bicicletta militare mentre gli altri avevano la bicicletta da corsa perché è vero che passano dallo Schnyzerdütsch al tedesco ma il salto è molto peggio dall'italiano al tedesco e magari allo Schnyzerdütsch Quindi non rischiamo di fare una legge che enuncia dei principi e che poi che ci

ricade addosso come un sassone pesantissimo Le posso fare un esempio che credo calzì a pennello Lo *Schnyzertütsch* che è stato nominato poco fa si constata attualmente che il dilagare dello *Schnyzertütsch* negli ultimi trent'anni supergiù conseguenza di una certa deregolamentazione che cosa ha prodotto ha prodotto un calo delle capacità di lettura di comprensione dei giovani della Svizzera tedesca questo dimostrato dalle analisi internazionali Cosa è successo o sta succedendo La conferenza svizzera dei direttori dell'educazione il consigliere di stato Gendotti lo potrà confermare ha cominciato a suggerire di reintrodurre in modo rigoroso il buon tedesco a partire dalle scuole dall'asilo fin su verso le alte scuole proprio perché questa è una misura contro la deregolamentazione è una misura politica che deve portare nella direzione di stabilire delle regole rispettivamente degli indirizzi sul piano della lingua da insegnare

Vi si nota – oltre ai tratti propri del parlato non progettuale e dell'italiano medio già rilevati nel brano precedente – a) la presenza di un certo numero di stranierismi (un germanismo ricorrente: *Schnyzertütsch*), b) di un calco sul tedesco (non ignoto anche all'italiano d'Italia: *bun tedesco*), c) di una forma avverbiale in funzione di connettivo disgiuntivo (*rispettivamente*), d) di un noto elvetismo ed e) di alcune forme amministrative altrettanto connotate (*direttori dell'educazione, consigliere di stato, alte scuole*).

Moltissimi altri tra i testi riportati in Pandolfi (2006) presentano caratteristiche analoghe e dunque, nel complesso, una marcatezza linguistica relativamente limitata. A fronte di questi dati di ordine linguistico, quali ragioni sostengono l'istanza dell'autrice del LIPSI (e in realtà anche quella di altri studiosi prima di lei), che – come si è scritto – considera l'italiano di Svizzera una delle due varietà paritetiche di un sistema linguistico policentrico?²¹ Non si tratta, evidentemente, di un problema nominalistico, ma di una questione di ordine sociolinguistico, per dirimere la quale si devono tenere in considerazione non solo fatti di sistema (attinenti, vale a dire, al sistema-lingua), ma anche extralinguistici, culturali, sociali e persino psicologici (in ciò che pertiene, per esempio, al vissuto dei parlanti²²), importanti – se del caso – per misurare la distanza tra lo *status* dell'italiano in Svizzera e, per esempio, quello dell'italiano settentrionale (e delle altre “regioni” d'Italia) o dell'italiano parlato, poniamo, nella penisola istriana o in Slovenia.

Un sistema policentrico è tale, per definizione, se vi si può identificare più di un centro e, nel caso di un sistema siffatto, si potranno definire *centri* realtà sociolinguistiche dotate della capacità di irradiare modelli linguistici, ovvero di produrre autonomamente strutture dotate di caratteristiche modellizzanti. L'effetto standardizzante di un centro linguistico, in una situazione di competizione normativa, è rafforzato dalla presenza di frontiere che ostacolano l'ingresso degli standard concorrenti e che creino adeguate condizioni di “segregazione”, intrinsecamente protettive e amplificatorie dei modelli

²¹ Sull'italiano usato in Svizzera sono ora disponibili numerosi studi, molti dei quali promossi proprio dall'Osservatorio Linguistico; oltre a Lurati 1980 e Bianconi 1980, si ricorderanno, tra i più recenti, Pandolfi, 2006, Moretti, 2004, Petralli, 1990. Indaga aspetti particolari della struttura sociolinguistica dell'italiano in Svizzera Antonini e Moretti, 2000; studia le dinamiche storico-linguistiche Bianconi, 2001.

²² Scrive Pandolfi, 2006, a p. 13: «Nei confronti dell'italiano d'Italia, l'italiano di Svizzera ha vissuto e vive sentimenti alterni nel rapporto con la 'norma', la varietà a base fiorentina sentita come la varietà alta, standard di lingua (cfr. Antonini-Moretti 2000, pp. 11-17) alla quale di volta in volta uniformarsi o dalla quale prendere le distanze per affermare la propria autonomia linguistica, culturale e territoriale».

locali. In queste condizioni, in genere, si instaura un differenziale linguistico riconoscibile a vari livelli del sistema (fonetico, prosodico, talora morfologico e morfosintattico e lessicale) che può fungere da spia induttiva al riconoscimento delle condizioni extralinguistiche di cui si è appena scritto. Un sistema linguistico policentrico, dunque, dovrebbe mostrare almeno alcune tra le seguenti caratteristiche: a) avere almeno due centri che abbiano potere modellizzante; b) presentare un certo numero di condizioni di segregazione; c) mostrare un differenziale linguistico riconoscibile tra le sue varietà che sia riconducibile ad a) e/o b).

L'individuazione del centro modellizzante (o *di radiazione?*)²³ dell'italiano svizzero rimane impregiudicata (si tratterà di Bellinzona, Lugano, Locarno o forse Berna? Più di una di queste località in sinergia?); lo è, tuttavia, anche per l'italiano d'Italia, in cui, nella realtà, Firenze non gioca più alcun ruolo rilevante, mentre esercitano un'influenza concorrenziale Milano e Roma, il primo un centro che irradia un modello sempre più potentemente koineizzante e neutralizzante. In effetti quella italiana – e credo si possa dire che ciò è vero anche per quella svizzero-italiana – è una situazione in cui convivono più norme di riferimento, che configurano dinamicamente equilibri e tendenze autorizzando e promuovendo idioletti e socioletti, esattamente come forze a orientamento vettoriale diverso che operino su un punto-uso: nel campo d'azione sociolinguistico, spinte solidaristiche e identitarie, tensioni antagonistiche, tendenze al riconoscimento dello *status*, ossequio ai detentori del potere culturale (D'Agostino, 2007) plasmano l'universo delle varietà e orientano le scelte maggioritarie. Terreno di sfida sono soprattutto i *media*, scritti e parlati, che diffondono un certo numero di modelli (in genere quelli già dotati di qualche prestigio) e ne determinano, così, la fortuna.

In Italia, dunque, come in Svizzera, la lenta deriva verso una condizione unitaria di equilibrio linguistico è orientata da modi d'uso e varietà che sono riverberate con maggiore o minore ampiezza ed efficacia variabile da televisione, radio e dalle interfacce telematiche; dall'interferenza, vale a dire, della lingua comune con quella pubblica delle amministrazioni e delle leggi, con i sottocodici tecnici a maggior impatto sociale, con le varietà regionali e i dialetti, con i gerghi ad alta commercializzabilità e, naturalmente, con alcuni codici alloglotti. L'italiano d'oltralpe, tuttavia, si differenzia dalle altre varietà regionali dell'italiano a) perché è diffuso in un'*enclave* extraterritoriale, b) perché è la lingua ufficiale di tale *enclave nazionale*, c) perché si trova in condizioni certamente speciali di contatto con altre lingue di cultura altrettanto ufficiali e con vitalissimi dialetti, e d) perché "mappa" una cultura diversa; fatti questi che costituiscono senza dubbio condizioni di segregazione e che rendono conto di un differenziale linguistico oggettivamente rilevabile, tanto a livello lessicale, quanto a livello morfosintattico (e, come è naturale, anche fonetico e prosodico; di queste differenze, che non sono rilevabili attraverso il LIPSI, non ci si occuperà in questa sede).

In particolare, per ciò che riguarda il differenziale linguistico, numerose peculiarità, come si è già visto, sono di pertinenza lessicale: vi sono parole nel LIPSI – e dunque presumibilmente anche nell'italiano parlato in Svizzera – che hanno frequenza (molto) più alta che, ad esempio, nel LIP o nel C-Oral-Rom (e dunque, presumibilmente, anche nell'italiano d'Italia) perché rinviano a *denotata* caratteristici: un lessema come *montagna* occorre per esempio 111 volte nella raccolta svizzera, ma solo 19 e 24, rispettivamente,

²³ Pandolfi, 2006: 35 e n. 40.

nel LIP e nel C-Oral-Rom; frequenze comparativamente molto alte hanno anche *caccia*, *neve*, *contadino*, *trekking*, *ghiaccio*, *mucca*, *stalla*, *sciare*, *ghiacciaio*, *cacciatore*, *capriolo* e altre ancora (Pandolfi, 2008: 71); è questo, però, un fatto che pertiene, più che al sistema-lingua, all'universo dei discorsi che con la lingua sono prodotti. È invece più direttamente implicata con la struttura linguistica una forma come *caccia alta* e altre simili (numerose rientrano nel novero degli elvetismi "assoluti" di cui si è già scritto e come tali non occorrono nell'italiano d'Italia), molte delle quali (mono- e polirematiche)²⁴ irradiano dall'amministrazione, sono dell'uso burocratico, presentano frequenza apprezzabilmente alta e costituiscono quello che l'autrice e gli studiosi dell'italiano di Svizzera chiamano *italiano federale*; si tratta, tra l'altro, di *iniziativa*, *apprendista*, *Conferenza*, *Dipartimento*, *patrizio*, *auto postale* e *buralista*; di *svizzero tedesco*, *consiglio federale*, *polizia cantonale*, *servizio pubblico*, *consiglio di stato*, *camera federale*, *cassa malati*, *consiglio nazionale*; e di altre forme ancora.²⁵ Si possono aggregare alla lista appena stilata anche pochi oggetti linguistici che si potrebbero considerare tecnologicamente determinati: si è già citato il caso di *Natel* 'cellulare', che è però affiancato dai più frequenti (e pan italiani) *telefonino* e *cellulare*.

Vige in realtà in questo campo, come in italiano e in altre lingue, una nomenclatura globalizzata su base inglese che lascia poco spazio a forme concorrenti. *Computer*, per esempio, occorre 44 volte e ha rango 778, *PC* occorre 5 volte e ha rango 3638, *calcolatore* è documentato solo in 4 casi e ha rango 4105; a differenza della seconda e della terza, la prima forma è produttiva e ha generato l'alterato *computerino* e la forma aggettivale *computerizzato*.

Internet ha rango 10 e non sembra avere concorrenti. *E-mail* occorre 15 volte (insieme al semplice *mail*, con 7 occorrenze), mentre *posta elettronica* non è documentato; *video game* occorre 6 volte e non lascia spazio all'antagonista *videogioco*. E gli esempi potrebbero moltiplicarsi.

Se poco caratteristiche sono molte tra le forme regionali, condivise molto spesso con gli italiani regionali settentrionali, lombardi in particolare, ha invece carattere distintivo e determinanti sociolinguistiche, piuttosto che l'uso di alcune forme (quelle regionali, per esempio), la frequenza con cui esse ricorrono, certamente più alta – stando a Pandolfi, 2006:47 – che nell'italiano dell'uso medio e forse anche che negli italiani regionali settentrionali²⁶. Il fatto è certamente un portato della tradizionale vitalità locale dei dialetti quali lingue della comunicazione ordinaria, che costituisce ancora un tratto di segregazione. A questo proposito, l'autrice calcola che possa essere considerato come regionalismo il 2.07 dei lemmi del *corpus*. Significativi a questo riguardo – perché caratterizzanti i "piani bassi" della lingua – sono soprattutto alcuni elementi funzionali – talora *polifunzionali* – come *bon* (che occorre 587 volte nel *corpus* e ha rango 85)²⁷, come

²⁴ Lurati, 1976; Pandolfi, 2006.

²⁵ Alcune di queste forme sono già state citate nei paragrafi precedenti.

²⁶ I dati di Pandolfi (2006) sembrano suggerire che i giovani siano tendenzialmente più poliglotti degli anziani, nel senso che introducono più frequentemente nel discorso elementi alloglotti, ivi compresi quelli dialettali. Ciò, più che essere il segno di una reviviscenza dei vernacoli presso i parlanti delle nuove generazioni, ne rende piuttosto evidente una spiccata inclinazione al mistilinguismo a fini espressivi.

²⁷ *Bon* è un elemento significativo non solo per via della sua frequenza assoluta, ma perché caratterizzato da un uso estremamente disperso in tutti i testi. Non si tratta, dunque, di elemento che contraddistingue uno o pochi dialetti, ma di una forma ubiquitaria. Va detto che la sua marcatezza diatopica, come osserva

l'interiezione *ciola* (rango 3638, frequenza 5) e come i combinati verbali o aggettivali/pronominali in cui occorrono le forme deittiche *lì, là, su, giù, qua, qui* (*andare giù, arrivare giù, marciare giù, esserci su e questo qui, dentro lì*), decisamente più rappresentate nel corpus LIPSI che nelle raccolte di confronto. Ha connotazione regionale anche l'uso, molto più esteso nel LIPSI che negli altri corpora, dell'elemento-wh *cosa*, normale nell'italiano regionale settentrionale e in progressiva diffusione anche in quello medio.

Contribuisce a delineare il carattere peculiare dell'italiano di Svizzera anche la frequenza dei forestierismi, che ammontano, secondo i calcoli dell'autrice, al 3,69% sui tipi lessicali e allo 0,42% sulle occorrenze; il dato è più alto – lo abbiamo già ricordato – di quello calcolato per il LIP (3,19%, 0,30% rispettivamente) e molto più alto di quello computato per il C-Oral-Rom (0,06% e 0,67%), anche se la differenza nei valori può essere giustificata entro certi limiti facendo ricorso ai criteri di categorizzazione (il LIPSI, per esempio, cataloga come esotismo le voci *film* e *camion*, che occorrono 84 e 50 volte e hanno rango 434 e 689; il LIP e il C-Oral-Rom, invece, considerano i lessemi come perfettamente adattati). In linea di massima, i parlanti svizzeri dell'italiano sembrano – e si potrebbe dire prevedibilmente – più inclini al forestierismo dei parlanti dell'italiano d'Italia. Ovviamente, più che la messe dei forestierismi pienamente assimilati cui si è fatto cenno nel paragrafo *Il LIPSI, le sue caratteristiche, la sua struttura*, ai fini della nostra analisi differenziale è significativo il numero di quelli che non hanno corso in Italia o che hanno una circolazione molto ridotta (e che sono spesso confinati in qualche lingua speciale): per trovarli basta scorrere l'elenco dei lemmi meno frequenti nella lista dei forestierismi (vi si leggeranno, tra gli *hapax* pretti tedeschi del tutto estranei all'uso del parlante medio d'oltralpe, francesismi che non hanno alcun corso o che hanno cittadinanza limitata o limitatissima in Italia [*panaché* 'bevanda composta birra e gassosa', *panachage* 'sistema elettorale per cui l'elettore può inserire in una lista il nome di candidati appartenenti a liste diverse'], anglicismi estranei all'uso comune della Penisola [*Side ways, settlement member, pet talk, bachelor*]), per quanto stranierismi ignoti all'italiano si trovino anche più in alto nella lista (alcuni sono già stati citati nel paragrafo *Il LIPSI, le sue caratteristiche, la sua struttura: Drug wipe*²⁸, *mémoire, à côté, Bonität, Devise, Zeitzone*)

Significativi sembrano anche gli criptostranierismi che risultano da adattamenti o calchi sul tedesco e sul francese, lingue che – per la loro diffusione e per il loro statuto amministrativo – hanno ovviamente un influsso fondamentale sull'italiano svizzero: si pensi, per quanto riguarda il tedesco, a *stanga* 'bicchiere di birra', *profilato* 'che spicca', o *rispettivamente*, che – come la congiunzione *beziehungsweise* – pare cumulare il valore distributivo di *rispettivamente* a quello disgiuntivo di *o* (Pandolfi, 2008: 95) e a quello epesegetico di 'per la precisione?'; per quanto attiene al francese si consideri anche il solo già citato e commentato *evidente*. Nel particolare contesto plurilingue svizzero – si è già sottolineato che esso configura, in maniera apparentemente paradossale, una specifica condizione di segregazione – acquisiscono un interesse specifico anche alcune forme persuase da più modelli concorrenti come *alte scuole*, polirematica che ha un parallelo in

già Pandolfi (2006: 47), è molto debole, considerata, tra l'altro, la sua convergenza con alcune forme interiettive del francese; e tuttavia proprio la sua presenza trasversale e le dinamiche di calcolo contribuiscono a innalzare la percentuale dei regionalismi.

²⁸ Questa polirematica è il nome commerciale di un prodotto con il quale è possibile effettuare test di positività a vari tipi di stupefacenti; ha guadagnato qualche – credo effimera – notorietà in Italia grazie a una trasmissione televisiva dai grandi ascolti in cui veniva proposta per l'uso a vari politici.

espressioni analoghe del francese e del tedesco, ‘scuole di livello universitario’; o *plafone*, che è persuasa dal francese ma forse anche dal sostrato dialettale (*plafon* per ‘soffitto’ è anche dei dialetti lombardi).

Caratterizzanti (e sistematicamente più rilevanti dei singoli elementi lessicali) sono pure alcune forme grammaticali (polirematiche o monorematiche, avverbiali o preposizionali), che hanno frequenza variabile e che, in qualche caso, sono state già citate nei paragrafi precedenti; molte calcano strutture del dialetto o di alcune varietà settentrionali dell’italiano; altre sembrano il risultato di un’evoluzione autonoma: si tratta, tra le altre, di *da parte a*, *per rispetto a*, *per rapporto a*, *per solito*, *dal lato*, *a corsa*, *rispettivamente*, ma l’elenco potrebbe allungarsi. Locali – e altrettanto caratterizzanti – sembrano anche alcuni usi preposizionali in collocazioni verbali come *mettere sotto discussione*, ed è certamente interessante dal punto di vista morfologico l’uso non connotato del femminile *ministra*: lo si sente anche nel giornalismo italiano d’Italia, ma mi pare che la forma non abbia ancora perso un’aura ironica o scherzosa, nonostante lo Zingarelli 2009 accolga il lemma senza marche (non così il GDU [De Mauro, 2000 sgg.], però); femminili sono anche *meteo* e *fine settimana*. Distinguono l’italiano svizzero da quello d’Italia anche alcuni fatti strettamente sintattici, come alcune reggenze verbali (per esempio l’uso transitivo di *pensare*, non ignoto però all’italiano d’Italia, almeno nell’accezione ‘immaginare’ [cfr. *penso questo, che cosa stai pensando?*], o quello pure transitivo di *avere bisogno*, diffuso anche nell’italiano regionale lombardo e nei dialetti).

Indicativi di uno statuto linguistico particolare²⁹, come sottolinea la stessa autrice, sono taluni fatti statistici che, come quelli citati nei capoversi precedenti, sembrano potersi giustificare convincentemente in un quadro di segregazione. Non si tratta, tanto, della *ratio type-token*, ovvero del rapporto tra numero di lemmi e numero di forme nell’unità di testo, per quanto tale valore possa in effetti essere utilizzato quale indicatore della ricchezza *media* del discorso (e, insieme ad altri, della sua *densità*, nei termini di Halliday, 1992; cfr. sull’argomento anche Chiari, 2007 e Lenci, Montemagni e Pirrelli, 2005), dal momento che il dato, secondo le rilevazioni dell’autrice, indica per il *corpus* LIPSI un valore solo leggermente più alto rispetto a quello del *corpus* LIP (il valore nei testi del *corpus* C-Oral-Rom, invece, è decisamente più alto e la discrepanza tra le due grandezze deve essere ancora spiegata adeguatamente). È piuttosto il caso del rapporto tra “parole piene” (dotate, vale a dire, di significato referenziale) e “parole vuote” (particelle e altri elementi funzionali che appartengono a classi chiuse): in questo caso i dati dell’autrice sembrano indicare che i testi del *corpus* LIPSI presentano sempre – sia nel caso in cui si effettui il computo prendendo in considerazione le forme, sia quando si considerino le parole – valori più alti rispetto a quelli registrabili per i *corpora* LIP e C-Oral-Rom (il differenziale è maggiore nel caso del rapporto tra forme). Ciò significa, in termini più tangibili, che nel *corpus* LIPSI il numero delle parole “referenziali” è in proporzione più alto di quello delle parole funzionali che nei *corpora* LIP e C-Oral-Rom, e il fatto sembra indicare che i testi LIPSI sono mediamente più *densi* di quelli delle altre raccolte. Il risultato può sicuramente essere giustificato tenendo conto di più di un fattore, ma non pare peregrina l’ipotesi di una maggiore formalità e scritturalità

²⁹ Si tratta di fatti che pertengono alla struttura del testo, nella quale però si riflettono evidentemente modi d’uso della lingua e rapporti tra utenti e struttura linguistica, che ci paiono rilevanti nella definizione della natura di un sistema linguistico (in quanto visto in relazione con i suoi utenti).

complessiva del parlato elvetico, più prossimo – per *segregazione*, ovvero per via delle modalità di impiego e di diffusione prevalenti almeno sino a qualche decennio fa in un quadro di vivace dialettologia – ai modelli “alti” promossi dalla scuola che a quelli più correvi dell’uso medio, come invece accade nel parlato comune dell’italiano contemporaneo d’Italia.

Avvalorata l’ipotesi di una maggiore complessità (e, forse, di una certa marcatezza tipologico-dialettale) dell’italiano svizzero anche il fatto che la percentuale delle congiunzioni coordinative sul totale delle congiunzioni presenti nel *corpus* sia meno alta nel LIPSI (*decisamente* meno alta nel raffronto con il LIP) che nei due *corpora* di raffronto. Quanto al resto, i dati estrapolati dal LIPSI mostrano invece una sostanziale convergenza con l’italiano d’Italia nel privilegiare un numero ridotto di elementi giunzionali ad alta ricorrenza.

Come è lecito attendersi, la prima posizione tra le congiunzioni coordinative – nel LIPSI come nel LIP e nel C-Oral-Rom – è occupata da *e*, mentre al primo posto tra quelle subordinative si trova *che*, noto per la sua polivalenza funzionale. Seguono *ma*, *poi*, *però*, *o*, *cioè*, *quindi*, *comunque*, *così*, *allora*, *oppure*, *né*; tra le subordinative *perché*, *quando*, *se*, *dove*, *anche se*, *visto che*, *mentre*, *perciò*, *per quanto*. Occorrerà osservare che molti di questi elementi giunzionali sono usati frequentemente in funzione “testuale”, ciò che evidentemente non può che concorrere a incrementarne la frequenza.

Una nota merita, per concludere, la struttura del vocabolario così come emerge dallo spoglio del LIPSI e dal raffronto dei lemmi (e delle forme) che appaiono avervi la frequenza maggiore con quelli omologhi del Vocabolario di base dell’italiano (De Mauro, 1980; De Mauro e Moroni, 1986): mentre il LIP, come ha mostrato Vedovelli (1993), presenta un quadro stratigrafico-distribuzionale sovrapponibile nelle sue linee essenziali a quello del Vdb, il LIPSI ne mostra uno piuttosto differente. Sono infatti assenti nel LIPSI forme del vocabolario fondamentale quali *accorgersi*, *arrabbiarsi*, *bontà*, *cortile*, *fedeltà*, *ferie*, *fiutare*, *graffio*, *grandine*, *guasto*³⁰; mancano forme del vocabolario d’alto uso quali *accarezzare*, *acconto*, *acido*, *acquedotto*, *afferrare*, *affondare*, *alcool*, *alleanza*, *amministratore*, *amoroso*, *annegare*, *arido*, *arnese*, *asciugamano*, *benedizione*; non sono attestate forme del vocabolario ad alta disponibilità quali *abbagliante*, *abbaiare*, *abete*, *accenno*, *adulta*, *afa*, *affogare*, *aglio*, *albicocca*, *allacciarsi*, *alluminio*, *alunna*, *ammuffire*, *analcolico*, *annusare*. Nel complesso, circa un terzo delle parole del VdB non sono attestate nel LIPSI³¹. Il dato, comunque interessante, si potrà giustificare in vario modo. Occorrerà, in primo luogo, tenere in considerazione la differenza essenziale che passa tra il vocabolario fondamentale e un lessico di frequenza estrapolato da un *corpus* di parlato, spontaneo e trasmesso radiotelevisivo: si tratta semplicemente di realtà lessicograficamente diverse.

Come si è notato nel paragrafo precedente, il Vocabolario di Base (VdB) è stato elaborato da Tullio De Mauro e dalla sua scuola (De Mauro, 1980; una lista delle parole che costituiscono il VdB appare poi anche in De

³⁰ Pandolfi (2007: 98) elenca tra le forme del vocabolario fondamentale assenti nel LIPSI anche *giovedì*, che è invece documentato alla posizione 1285, con 24 occorrenze; *ammalarsi* occorre una volta e ha rango 12331.

³¹ Si veda la tabella di Pandolfi, 2008: 97-99.

Mauro e Moroni, 1996) partendo dall'elaborazione statistica dei dati del *Lessico Italiano di Frequenza* (LIF: Bortolini *et alii*, 1972): dall'opera sono stati prelevati i primi 5000 lemmi e se ne è verificata sul campo la comprensibilità; dopo aver eliminato (perché essi non erano stati compresi da almeno il 50% degli informatori impiegati allo scopo) circa 250 termini, gli autori hanno provveduto ad accorpate i rimanenti nelle due classi delle parole che costituiscono il Vocabolario fondamentale (i primi 1.991 lemmi del LIF) e di quelle che formano il Vocabolario di alto uso (i successivi 2.750 lemmi). Una terza categoria – quella delle parole che compongono il Vocabolario di alta disponibilità (2.337 lemmi) – è stata costituita a partire da fonti diverse, soprattutto attraverso lo spoglio di dizionari dell'italiano comune e la validazione dei dati con un gruppo di informatori: il LIF, infatti, che è stato elaborato all'inizio degli anni Sessanta del secolo scorso, è il risultato dello spoglio di testi unicamente scritti (testi teatrali, romanzi, copioni cinematografici, quotidiani e settimanali, libri per le scuole elementari) e per questa ragione non include un insieme di lemmi che, pur essendo quasi del tutto assenti nella lingua scritta, sono conosciuti da tutti i parlanti³².

Si dovrà considerare, in secondo luogo, l'insieme degli universi di discorso attivati nel corso della raccolta dei dati – si ricordi che per neutralizzare l'“effetto academia” (la pulsione semiconscia alla selezione di un registro alto) e l'“effetto taverna” (la spinta semiconscia e modellizzante all'impiego di forme stereotipicamente collegate alla varietà, in genere marcata, che si vuole rappresentare) indotto dalle interviste a microfono a vista, gli informatori sono stati sollecitati a parlare della professione, della famiglia, della vita di tutti i giorni e degli avvenimenti che li toccano più da vicino –: si tratta di un dato che può legittimare l'assenza di alcuni referenti (cosa potrà mai dire un parlante svizzero in sostituzione di *balcone*, o di *bontà* o di *coscia* o di *Cristo* o di *acido* o di *alluminio?*); e si dovrà anche tenere in considerazione il fatto che la presenza di alcune forme può essere condizionata dalla loro connotazione diatopica (*babbo*, assente nel LIPSI ma parte del Vdb, è ormai confinato anche in Italia alle varietà di Toscana e di Sardegna) o diafasica (*battona*, pure non documentato nel LIPSI a differenza che nel Vdb, è disfemico e ha una leggera coloritura diatopica centrale). Altri casi di esclusione potranno essere giustificati tenendo in conto l'evoluzione diacronica (sarà il caso di *domattina* o di *giacché* o di *grembo* o di *guanciaie* o di *balia*, tutte forme che – segnala l'autrice – il LIPSI non documenta?)³³ o ragioni di tipo culturale (i *brigadieri* sono molto meno frequenti in Svizzera che in Italia! *Cin-cin* è spesso sostituito da *ciao*; l'*Adriatico* – a maggior ragione nella sua forma aggettivale – non rientrerà facilmente nei discorsi quotidiani dei ticinesi) o alcuni accidenti puramente formali (*dopoché* è certamente reinterpretato come la congiunzione polirematica *dopo che* e non è neppure nel LIP).

³² Secondo il *Grande dizionario italiano dell'uso* (Gradit: De Mauro, 2000 sgg.), invece, il Lessico fondamentale includerebbe 2.049 unità; il lessico di alto uso 2.576; il lessico di alta disponibilità 1897, sicché il vocabolario di base conterebbe 6522 parole.

³³ La seconda forma è presente nel LIP una sola volta, e occorre nel parlato di un napoletano; la terza in realtà manca. *Grembo* è indicato in De Mauro (1980) come AU e non come fondamentale; *guanciaie* è etichettato, in De Mauro (1980), come AD e non come fondamentale. *Balia* potrebbe essere sostituito dall'iperonimo *bambinaia*, che occorre due volte, e dall'appena più frequente *baby-sitter* (tre occorrenze).

L'italiano svizzero, dunque, è una lingua “altra” rispetto all'italiano d'Italia? La risposta è – in modo solo apparentemente paradossale – contemporaneamente positiva e negativa. È evidente – lo abbiamo appena rilevato – che l'italiano di Svizzera si è sviluppato in presenza di condizioni di isolamento, che hanno reso inevitabile l'affermazione di caratteri specifici; e altrettanto chiaro appare che essi costituiscono un sistema di varianti che realizza un differenziale linguistico riconoscibile.

Per riepilogare: i dati rilevabili attraverso l'ispezione del LIPSI – per ciò che attiene, come abbiamo visto, alla varietà del lessico, alla frequenza degli elementi referenziali, alla frequenza complessiva dei giuntivi subordinanti – sembrano effettivamente suggerire che l'italiano parlato in Svizzera sia mediamente più formale di quello parlato in Italia, e che i parlanti abbiano per modello piuttosto la scrittura che non il parlato paradigmatico; altri dati (cfr. per questo anche Pandolfi, 2006 e Berruto, 1980) paiono inoltre indicare «una certa propensione all'aulicità e all'arcaismo con una conseguente tendenza all'ipercorrettismo»³⁴.

L'italiano svizzero, inoltre – ed è anche questo tratto riconducibile a una situazione di segregazione – si caratterizza per la presenza di numerosi stranierismi (anglicismi, francesismi e germanismi, alcuni per la verità dotati di frequenza molto bassa) non adattati o calcati che non hanno alcun corso al di là del confine; senza contare le molte divergenze nella composizione del lessico addebitabili alla distanza culturale (ove si assuma il termine *cultura* in un senso estensivo, a includere tutte le manifestazioni sociali: amministrative, politiche, burocratiche), che giustifica la presenza di un ampio contingente di parole che non esistono nell'Italiano d'Italia o che, più frequentemente, vi hanno un significato diverso. Infine, il sistema linguistico sembra avere assunto una conformazione originale anche in alcune strutture meno superficiali di quella lessicale; nel dominio morfosintattico, per esempio, si registrano vari casi di reggenze preposizionali difformi da quelle dell'italiano d'Italia; né sono rari i casi di polirematiche preposizionali o avverbiali che non hanno riscontro nella penisola.

L'alterità dell'italiano di Svizzera rispetto agli italiani regionali della Penisola è in qualche modo anche tendenziale (e tale tendenzialità dipende, ancora una volta, dal contesto di confinamento – in senso anche propriamente etimologico – in cui esso vive e prospera): mentre gli italiani regionali sono sottoposti alla forte pressione koineizzante e neutralizzante esercitata soprattutto dai media audiovisivi, l'italiano di Svizzera sembra orientato alla stabilizzazione dei tratti propriamente regionali che lo caratterizzano e, insomma, all'affermazione di un nuovo standard che è pienamente *locale*.

D'altra parte, per dirla insieme con il Machiavelli e con il Trissino³⁵, anche se in una prospettiva rovesciata, a giudicare dai dati estrapolabili dal LIPSI e ancor più da quelli attingibili al *corpus* di Pandolfi (2006), è evidente che il parlare d'Italia e quello di Svizzera

³⁴ Caratteristiche che – segnala Pandolfi (2006: 61) – “sembra stiano però un po' scomparendo con la sempre maggiore competenza e familiarità di tutti i parlanti nell'uso dell'italiano”.

³⁵ Con i quali ritorniamo in effetti alla risposta paradossale che abbiamo dato alla domanda che ci siamo posti qualche capoverso più in alto: si tratta delle celebri affermazioni che il Segretario fiorentino ha affidato al *Discorso* (Trovato, 1982) in difesa della toscaneità della lingua di Dante e che il letterato vicentino ha fatto pronunciare (al Rucellai) nel *Castellano* a sostegno della sua tesi italianista (Castelvecchi, 1986).

presentano *molto più del comune che del proprio*, e per l'appunto ai livelli che determinano l'identità di una lingua. Ed è altrettanto evidente che la frontiera alpina è stata (ed è ancora) molto permeabile, come dimostra il fatto che l'italiano svizzero presenta una notevole affinità (talora proprio in alcuni dei punti in cui si scosta dall'italiano medio) con le varietà regionali del settentrione d'Italia, con quella lombarda soprattutto.

Unità di fondo dunque, ma nel contesto di una situazione sociolinguistica particolare e in presenza di segni oggettivi di distanziamento; per questo quello che possiamo considerare il *particolare elvetico* non deve essere né ignorato né giudicato in riferimento alla norma astratta dell'italiano d'Italia: deve essere invece tenuto in attenta considerazione – come suggerisce giustamente l'autrice nelle pagine conclusive della premessa al LIPSI – soprattutto ai fini di una didattica consapevolmente *differenziale*, che “deve iniziare fin dai primi livelli di apprendimento per quanto riguarda il lessico e approfondirsi e estendersi anche ad aspetti di morfosintassi nei livelli intermedio e avanzato” (106).

RIFERIMENTI BIBLIOGRAFICI

- AA.VV. (1977), *Atti del Seminario sull'italiano parlato*, Accademia della Crusca, Firenze.
- AA.VV. (1982), *La lingua italiana in movimento*, Accademia della Crusca, Firenze.
- AA.VV. (1987), *Gli italiani parlati*, Accademia della Crusca, Firenze.
- AA.VV. (1997), *Gli italiani trasmessi: la radio*, Accademia della Crusca, Firenze.
- Andorno C., Bernini G. (2003), “Premesse teoriche e metodologiche”, in Giacalone Ramat (a cura di), pp. 27-36.
- Antonini F., Moretti B. (2000), *Le immagini dell'italiano regionale: la variazione linguistica nelle valutazioni dei giovani ticinesi*, Osservatorio linguistico della Svizzera italiana, Bellinzona.
- Aprile M. (2005), *Dalle parole ai dizionari*, Il Mulino, Bologna.
- Barbera M., Corino E., Onesti C. (a cura di) (2007a), *Corpora e Linguistica in Rete*, Guerra, Peurgia.
- Barbera M., Corino E., Onesti C. (a cura di), (2007b), “Cos'è un corpus? per una definizione più rigorosa di corpus, token e markup”, in Barbera, Corino, Onesti (a cura di) 2007°, pp. 25-88.
- Berruto G. (1980), *Alcune considerazioni sull'italiano regionale ticinese*, Dipartimento della pubblica educazione, Ufficio dell'insegnamento medio, Bellinzona.
- Bianconi S. (1980), *Lingua matrigna: italiano e dialetto nella Svizzera italiana*, Il Mulino, Bologna.
- Bianconi S. (1998), *Plurilinguismo in Val Bregaglia*, Dadò, Locarno.
- Bianconi S. (2001), *Lingue di frontiera: una storia linguistica della Svizzera italiana dal Medioevo al 2000*, Casagrande, Bellinzona.
- Bianconi S. (a cura di), (1994), *Lingue nel Ticino*, Dadò, Locarno.
- Bianconi S. (a cura di), (1995), *L'italiano in Svizzera*, Dadò, Locarno.
- Biber D., Conrad S., Reppen R. (1998), *Corpus Linguistics. Investigating Language Structure and Use*, Cambridge University Press, Cambridge.
- Bindi R., Monachini M., Orsolini P. (1991), *Italian Reference Corpus. General Information and Key for Consultation*, Istituto di Linguistica Computazionale – CNR, Pisa.

- Bortolini U., Tagliavini C., Zampolli A. (1972), *Lessico italiano di frequenza* (LIF), Garzanti, Milano. [ripubblica i dati dell'edizione IBM del 1971].
- Bortolini U., Zampolli A. (1971), *Lessico di frequenza della lingua italiana contemporanea: prospettive metodologiche*, in Medici M. e Simone R. (a cura di.) 1971, vol. II, pp. 639-648.
- Bruni F. (a cura di), (1992), *L'italiano nelle regioni. Lingua nazionale e identità regionali*, UTET, Torino.
- Casoni M. (2003), *La presenza delle lingue in un repertorio di siti web elvetici*, Osservatorio linguistico della Svizzera italiana, Bellinzona [in linea all'indirizzo Web: <http://www.ti.ch/decs/dcsu/AC/OLSI/pubblicazionionline.asp>].
- Casoni M. (2005), *L'italianità nei giornali della Svizzera d'oltralpe*, Osservatorio linguistico della Svizzera italiana, Bellinzona. [in linea all'indirizzo Web: <http://www.ti.ch/decs/dcsu/AC/OLSI/pubblicazionionline.asp>].
- Castelvecchi A. (a cura di), (1986), *Giangiorgio Trissino. Scritti linguistici*, Salerno, Roma.
- Chiari I. (2004), *Informatica e lingue naturali. Teorie e applicazioni computazionali per la ricerca sulle lingue*, Aracne, Roma.
- Chiari I. (2007), *Introduzione alla linguistica computazionale*, Laterza, Bari.
- Cresti E., (2000), *Corpus di italiano parlato*, 2 voll., vol. I *Introduzione*, vol. II *Campioni*, Accademia della Crusca, Firenze.
- Cresti E., Moneglia M. (a cura di), (2005), *C-Oral-Rom. Integrated Reference Corpora for Spoken Romance Languages*, Benjamins, Amsterdam.
- Crocco C., Savy R., Cutugno F. (2003), *A.P.I.- Archivio del parlato italiano*, DVD-ROM, Università degli studi di Napoli "Federico II", distribuito da CIRASS, Università degli studi di Napoli "Federico II".
- D'Agostino M. (2007), *Sociolinguistica dell'Italia contemporanea*, Il Mulino, Bologna.
- De Mauro T. (1980), *Guida all'uso delle parole*, Editori Riuniti, Roma.
- De Mauro T. (2000 sgg), *Grande Dizionario Italiano dell'uso*, 6 voll., UTET, Torino, [due nuovi volumi, 7° e 8°, con CD Rom e chiave USB sono stati stampati nel 2007].
- De Mauro T., Mancini F., Vedovelli M., Voghera M. (1993), *Lessico di frequenza dell'italiano parlato*, ETAS, Milano.
- De Mauro T., Chiari I. (a cura di), (2005), *Parole e numeri. Analisi quantitative dei fatti di lingua*, Aracne, Roma.
- De Mauro T., Moroni G. G. (a cura di), (1996), *DIB. Dizionario di base della lingua italiana*, Paravia, Torino.
- Giacalone Ramat A. (a cura di), (2003), *Verso l'italiano. Percorsi e strategie di acquisizione*, Carocci, Roma.
- Giuliano L. (2004), *L'analisi automatica dei dati testuali. Software e istruzioni per l'uso*, LED, Milano.
- Halliday M. A. K. (1992), *Lingua italiana e lingua scritta*, La Nuova Italia, Firenze [edizione originale *Spoken and written Language*, Victoria, Deakin University Press, 1985]. <http://www.ti.ch/decs/dcsu/AC/OLSI/pubblicazionionline.asp>].
- IBM Italia (1989), *VELI. Vocabolario Elettronico della Lingua Italiana*, IBM Italia, Milano.
- Lenci A., Montemagni S., Pirrelli V. (2005), *Testo e computer. Elementi di linguistica computazionale*, Carocci, Roma.
- Lo Zingarelli 2009, Zanichelli, Bologna.
- Lurati O. (1976), *Dialecto e italiano regionale nella Svizzera italiana*, Banca Solari e Blum, Lugano

- Lurati O. (1980), *La lingua italiana in Svizzera*, Istituto Lombardo di Scienze e Lettere, Milano.
- Lurati O. (1992), *Il Canton Ticino*, in Bruni 1992, pp. 143-177.
- MacWhinney B. (1997), *Il progetto CHILDES - Strumenti per l'analisi del linguaggio*, edizione italiana a cura di Elena Pizzuto e Umberto Bortolini, Edizioni del Cerro, Pisa.
- Marello C. (1996), *Le parole dell'italiano. Lessico e dizionari*, Zanichelli, Bologna.
- McEnery T., Wilson A. (2006), *Corpus Linguistics*, Edinburgh University Press, Edinburgh.
- Medici M., Simone R. (1971), *L'insegnamento dell'italiano in Italia e all'estero*, Atti del IV convegno internazionale di studi della Società di Linguistica italiana, Roma, 1-2 giugno 1970, 2 voll., Bulzoni, Roma.
- Moretti B. (a cura di), (1998-2005), *La terza lingua: aspetti dell'italiano in Svizzera agli inizi del terzo millennio*, 2 voll., vol. I Norma e varietà di lingua in Ticino; vol. II Dati statistici e "varietà dialesche", Dadò, Locarno.
- Moretti B. (a cura di), (1999), *Ai margini del dialetto*, Dadò, Locarno.
- Moretti B., Antonini F. (2000), *Famiglie bilingui*, Dadò, Locarno.
- Pandolfi E. M. (2003) *L'italiano, il dialetto e le altre lingue nella pubblicità in Ticino*, Istituto di lingua e letteratura italiana dell'Università di Berna [in linea all'indirizzo Web: <http://www.ti.ch/decs/dcsu/AC/OLSI/pubblicazionionline.asp>].
- Pandolfi E. M. (2006), *Misurare la regionalità. Uno studio quantitativo su regionalismi e forestierismi nell'italiano parlato nel Canton Ticino*, Dadò, Locarno.
- Pandolfi E. M. (2009), *LIPSI. Lessico di frequenza dell'italiano parlato nella Svizzera italiana*, Osservatorio Linguistico della Svizzera Italiana, Bellinzona.
- Petralli A. (1990), *L'italiano in un cantone: le parole dell'italiano regionale ticinese in prospettiva sociolinguistica*, F. Angeli, Milano.
- Rossini Favretti R. (2000a), *Linguistica e informatica. Multimedialità, corpora e percorsi di apprendimento*, Bulzoni, Roma.
- Rossini Favretti R. (2000b), *Progettazione e costruzione di un corpus di italiano scritto: CORIS/CODIS*, in Ead. (a cura di), 2000°, pp. 39-56.
- Rossini Favretti R., Tamburini F., De Santis C. (2002), *A corpus of written Italian: a defined and a dynamic model*, in Wilson, Rayson, McEnery (eds.) 2002, pp. 27-38.
- Savoia S., Vitale E. (2002), *Lo Svizzionario*, Linguanostra, Bellinzona.
- Sciarone G. A. (1995), *Vocabolario Fondamentale della Lingua Italiana*, Perugia, Guerra [già Minerva Italica, Bergamo, 1977].
- Spina S. (2001), *Fare i conti con le parole. Introduzione alla linguistica dei corpora*, Guerra, Perugia.
- Thornton A. M., Iacobini C., Burani C. (1997), *VDVDB - Una base di dati sul Vocabolario di base della lingua italiana*, Bulzoni, Roma.
- Trovato P. (a cura di), (1982), Niccolò Machiavelli. *Discorso intorno alla nostra lingua*, Antenore, Padova.
- Vedovelli M. (1993), *Confronti tra il LIP e le altre liste di frequenza dell'italiano*, in De Mauro, Mancini, Vedovelli e Voghera 1993, pp. 119-147.
- Wilson A., Rayson P., McEnery T. (a cura di), (2002), *A rainbow of corpora: Corpus linguistics and the languages of the world*, LINCOM Europa, Munich.