





ACCADEMIA DELLA CRUSCA

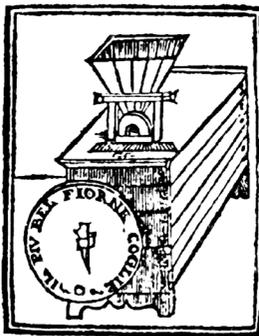
**OSSERVATORIO
DEGLI ITALIANISMI
NEL MONDO:
PUNTI DI PARTENZA
E NUOVI ORIZZONTI**

Atti dell'incontro OIM

Firenze, Villa Medicea di Castello

20 giugno 2014

a cura di
Matthias Heinz



FIRENZE
2017

Tutti i diritti riservati

Sono rigorosamente vietati la riproduzione, la traduzione, l'adattamento, anche parziale o per estratti, per qualsiasi uso o con qualsiasi mezzo effettuati, compresa la copia fotostatica, il microfilm, la memorizzazione elettronica, ecc., senza la preventiva autorizzazione scritta dell'Editore. Ogni abuso sarà perseguito a norma di legge.

Cura editoriale: Ufficio Pubblicazioni dell'Accademia della Crusca

© Accademia della Crusca
Via di Castello, 46
50141 Firenze
www.edizionidicrusca.it

ISBN 978-88-89369-80-7

PREMESSA

Questo volume raccoglie gli Atti di una giornata di studio dedicata all'*Osservatorio degli italianismi nel mondo* (OIM) tenutasi il 20 giugno 2014 presso l'Accademia della Crusca. L'iniziativa dell'osservatorio è strettamente legata alle attività dell'Accademia perché sorta da un'opera pubblicata nel 2008 nella collana "Grammatiche e lessici", il *Dizionario degli italianismi in francese, inglese, tedesco* (DIFIT). Nasce presto l'idea della sua digitalizzazione con apposita presentazione in rete e dall'originario nucleo di dati lessicografici prende spunto il progetto di un osservatorio degli italianismi di più ampio respiro, accolto di recente tra i progetti strategici dell'Accademia.

Ai dibattiti di questa Giornata, formalmente aperta dal Presidente della Crusca, Claudio Marazzini, assisteranno da parte della stessa Crusca Marco Biffi, Domenico De Martino, Massimo Fanfani, Nicoletta Maraschio, Matilde Paoli, Maurizio Rago, Giovanni Salucci, Luca Serianni e Raffaella Setti, e dall'estero Giuseppe Brincat (Malta), Stefan Engelberg e Peter Meyer (Mannheim), Matthias Heinz (Salisburgo), Elzbieta Jarmozik (Varsavia), Franco Pierno (Toronto), Gesine Seymer (Dresda) e Harro Stammerjohann (Francoforte sul Meno).

I miei sentiti ringraziamenti vanno a Anne-Kathrin Gärtig, Eleonora Doria e Katharina Kofler per aver partecipato a varie fasi redazionali dei presenti atti, a Harro Stammerjohann per la rilettura delle bozze e per aver contribuito una presentazione del volume, così come a Domenico De Martino per averne accompagnato la realizzazione e all'Accademia della Crusca per aver voluto accogliere gli atti tra le sue pubblicazioni.

MATTHIAS HEINZ
Universität Salzburg



SALUTO

L'Accademia della Crusca, come tutti sappiamo, pur essendo straordinariamente legata a Firenze, la città dove è nata e dove ha svolto, in modo ininterrotto per oltre quattrocentotrenta anni, la sua attività, ha sempre, fin dall'inizio, guardato fuori dalle mura cittadine, in direzione non solo dell'Italia, ma del mondo. L'Accademia l'ha fatto perché era convinta del valore della lingua italiana (per secoli raccolta e trasmessa dal suo grande *Vocabolario*), ma anche perché era consapevole del diffuso apprezzamento di cui essa godeva in tutta Europa. Per quanto riguarda l'età presente, emblematica è la parte iniziale del nuovo Statuto entrato in vigore nel 2011: «L'Accademia della Crusca nel perseguire i suoi scopi fondamentali, considera la lingua italiana in tutti i suoi aspetti cognitivi, culturali e sociali, nella realtà storica e attuale dell'Italia e nel quadro del multilinguismo europeo e della comunicazione mondiale» (art 2.1). Si tratta di un'apertura internazionale molto ben documentata nel tempo: dalla folta e qualificata presenza di accademici esteri, ai frequenti rapporti con studiosi d'oltralpe, dalle convenzioni/protocolli d'intesa con importanti istituzioni linguistiche di varie parti del mondo, alla collaborazione con enti/associazioni (come l'EFNIL e la REI) che operano in Europa in favore di una politica linguistica realmente rispettosa del multilinguismo. Tre studiosi, Giuseppe Brincat, Elisabetta Jamrozik e Franco Pierno offrono nel loro intervento un saggio molto istruttivo della fortuna delle parole italiane in tre realtà diverse: Malta, Polonia e Canada.

Eppure, come è testimoniato da questi Atti, l'*Osservatorio degli italianismi del mondo* (OIM) è un progetto del tutto

nuovo per l'Accademia e lo è da diversi punti di vista; innanzi tutto da quello organizzativo e metodologico. Per la prima volta un gruppo di studiosi di diverse nazionalità lavorerà a un'impresa lessicografica di Crusca, una grande banca dati degli italianismi nel mondo. Il primo nucleo di tale banca dati è rappresentato dal *Dizionario degli italianismi in francese, inglese e tedesco* (DIFIT, diretto da Harro Stammerjohann, pubblicato in versione cartacea dalla Crusca nel 2008 e nel 2013 nel nostro sito, all'interno del repository digitale VI-VII. Vivi italiano, in una versione elettronica predisposta da Gesine Seymer). Ma il nucleo originario sarà notevolmente incrementato dai risultati di ricerche estese ad altre lingue. Alcuni risultati sono già disponibili grazie al progetto *Censimento degli italianismi nel mondo*, avviato nel 2004 (per la casa editrice Utet di Torino) da Luca Serianni con una nutrita squadra di collaboratori (censimento che ha prodotto una cospicua serie di pubblicazioni particolari, senza comunque trovare la conclusione prevista). Ma altri italianismi saranno indicati da gruppi di ricerca specializzati, italiani e stranieri, sparsi per tutto il mondo. Inoltre il carattere di *Osservatorio* implicherà che il sistema possa essere via via aggiornato da segnalazioni di utenti comuni, che confluiranno in un serbatoio e saranno opportunamente vagliate da esperti prima di essere pubblicate. Il dialogo tra le diverse parti del sistema richiederà una ricerca informatica specifica e l'approntamento di strumenti adeguati di interrogazione.

Ma la novità dell'OIM non è solo organizzativa e metodologica. Come appare evidente da questa giornata di studio, la novità più significativa consiste in un vero e proprio rovesciamento di prospettiva rispetto a quella tradizionale della Crusca. È importante infatti ricordare due fatti: l'iniziativa di una ricerca sistematica sugli italianismi è partita qualche anno fa da fuori dei confini nazionali, dall'accademico Harro Stammerjohann (coordinatore per il DIFIT di un gruppo di studiosi di varie nazionalità); l'allargamento della ricerca stessa è stato invece proposto da un suo allievo, Matthias Heinz

dell'Università di Salisburgo. A loro si sono uniti successivamente altri accademici italiani e stranieri, altri collaboratori dell'Accademia e altri studiosi che, come Pierno, si stanno già muovendo in parte autonomamente. Il loro impegno è stato premiato dall'Accademia della Crusca che ha deciso di considerare l'OIM un proprio progetto strategico, destinato quindi a caratterizzarne significativamente l'attività dei prossimi anni.

NICOLETTA MARASCHIO

Firenze



PRESENTAZIONE

Sono grato a Matthias Heinz per aver portato alla pubblicazione gli atti di una Giornata di studio dedicata all'*Osservatorio degli italianismi nel mondo* (da ora in poi OIM) tenutasi il 20 giugno 2014 presso l'Accademia della Crusca. Mentre sei delle relazioni tenute da Marco Biffi, Giuseppe Brincat, Stefan Engelberg e Peter Meyer (quest'ultimo è l'autore responsabile della pubblicazione qui presente), Matthias Heinz, Elżbieta Jamrozik, Matilde Paoli, Maurizio Rago e Giovanni Salucci, Franco Pierno, Luca Serianni, Raffaella Setti e Gesine Seymer, sono state riunite in questo volume, Marco Biffi, Raffaella Setti e Gesine Seymer, invece di pubblicare i loro interventi per intero, mi hanno mandato dei riassunti che sono lieto di riprodurre in questa presentazione.

Infatti, sia l'ideazione dell'OIM che i problemi della sua realizzazione risultano molto chiaramente dal riassunto di Biffi: «L'OIM prende l'avvio dal progetto del portale VIVIT (Vivi italiano) e in esso trova la sua origine storica e metodologica. Nel VIVIT infatti si ritrova il primo nucleo di un dizionario di italianismi sul web, con la realizzazione informatica del DIFIT elettronico (www.italianismi.org); e d'altro canto in esso si rinviene l'embrione di un vero e proprio osservatorio, con la possibilità, offerta ai vari consultatori del portale sparsi nel mondo, di segnalare italianismi venendo incontro alla filosofia WEB 2.0 presupposta dal Ministero finanziatore.

L'OIM si configura infatti come una grande banca dati, di cui le voci già ereditate dal DIFIT costituiscono la base. Tali voci devono essere incrementate nel numero, aggiungendo italianismi provenienti da altre lingue, le cui schede siano cu-

rate da gruppi di ricerca specializzati; e devono essere adattate nella struttura, in modo che il nuovo dizionario elettronico sia completato da nuovi campi descrittivi che possano rendere conto delle specificità emergenti da altre realtà linguistiche e del confronto sistematico con i corpora di riferimento.

Ma va anche potenziato il sistema di raccolta di segnalazioni degli utenti comuni, che devono confluire in una banca dati che serva da serbatoio di lavoro e di studio, e che una redazione di esperti linguisti possa compulsare in vista della redazione di nuove voci. Fondamentale per il progetto è la ricerca informatico-linguistica, visti i complessi problemi posti dalle numerose interazioni fra piattaforme diverse; problemi che richiedono la progettazione e la creazione di strumenti informatici innovativi che potrebbero incrociarsi con le linee strategiche di alcuni progetti europei Horizon 2020».

Matthias Heinz (*Dal DIFIT all'OIM. Sfide lessicografiche e prospettive di implementazione*) illustra quali «opzioni di ricerca» (Biffi) si offrono dal momento che i dati raccolti nel DIFIT sono resi accessibili in una banca dati, base dell'OIM. Per esempio, ha potuto stabilire che i prestiti diretti, senza modificazione formale, sono di gran lunga i più numerosi, mentre i calchi formali, parziali e altri sono molto meno frequenti; poi, che dei prestiti polirematici ben un terzo sono del tipo N+agg (*salto mortale*), seguiti dal tipo Prep+N (*a battuta*). In conclusione del suo intervento Heinz presenta OIMap, il progetto della rappresentazione cartografica dell'irradiazione della lingua italiana, progetto in via di realizzazione presso la Università di Salisburgo.

Luca Serianni (*L'italiano nel mondo. Intenti e propositi di un progetto editoriale sugli italianismi*) riassume la storia di un progetto di pubblicazione che non andrebbe neanche chiamato un fallimento, considerando i materiali già consegnati quando l'editore lo interruppe, forse meno per impazienza con i ritardi di alcuni dei contributori, del tutto normali in un'opera collettiva di questo ordine di grandezza, che per aver mutato politica editoriale. Erano previsti due volumi, uno

dedicato alle aree storico-geografiche, dalla Svizzera al Corno d’Africa, e a quelle semantiche, dalla musica alla gastronomia ecc., il secondo al censimento degli italianismi in 78 lingue – il primo affidato a più di 60, il secondo a più di 80 autori. Gran parte dei saggi previsti per il primo volume e alcuni dei regesti degli italianismi in varie lingue del mondo sono stati pubblicati altrove. Dopo aver raccontato la (s)fortuna del progetto editoriale di sopra, Serianni aggiunge riflessioni di metodo, in particolare su quello che lui chiama il “peso di un prestito”, cioè la sua effettiva presenza nell’uso della lingua di arrivo, che spesso non risulta dai dizionari.

Sono problemi che hanno incontrato anche Stefan Engelberg e Peter Meyer, ideatori del *Lehnwortportal Deutsch* (Peter Meyer, *The limits of lexicographical abstraction. Some strengths and problems of the architecture in the Lehnwortportal Deutsch*). Il *Lehnwortportal Deutsch* è simile all’OIM, essendo anch’esso un «data-base driven ‘inverted’ loanword dictionary», cioè un dizionario di germanismi, sia diretti che indiretti, in altre lingue, a cominciare da quelli nelle lingue slave. In quanto si basa sui dizionari di germanismi già esistenti, che cerca di collegare tra di loro «under a single roof and with a uniform access structure», piuttosto che un dizionario, è una rete di dizionari, difficile da realizzare a causa della eterogeneità delle fonti incluse.

Seguono tre visioni d’insieme degli italianismi nel maltese, nel polacco e nelle lingue del Canada. Come ricorda Giuseppe Brincat (*Gli italianismi nella lingua maltese: non solo prestiti*), fuori d’Italia il maltese è sì la lingua più italianizzata di tutte, ma in modi molto particolari. Per esempio, se sono pochi i sicilianismi veri e propri, per un processo ancora attivo gli italianismi assumono una forma sicilianeggiante, come fanno anche gli anglolatinismi, tanto che i maltesi che non sanno l’inglese li considerano italianismi. Ogni lingua, sembra, ha subito l’influsso italiano in un modo diverso, a seconda della situazione storico-geografica. Per il polacco, Elżbieta Jamrozik (*Gli italianismi in polacco: storia e realtà attuale*) distingue tre

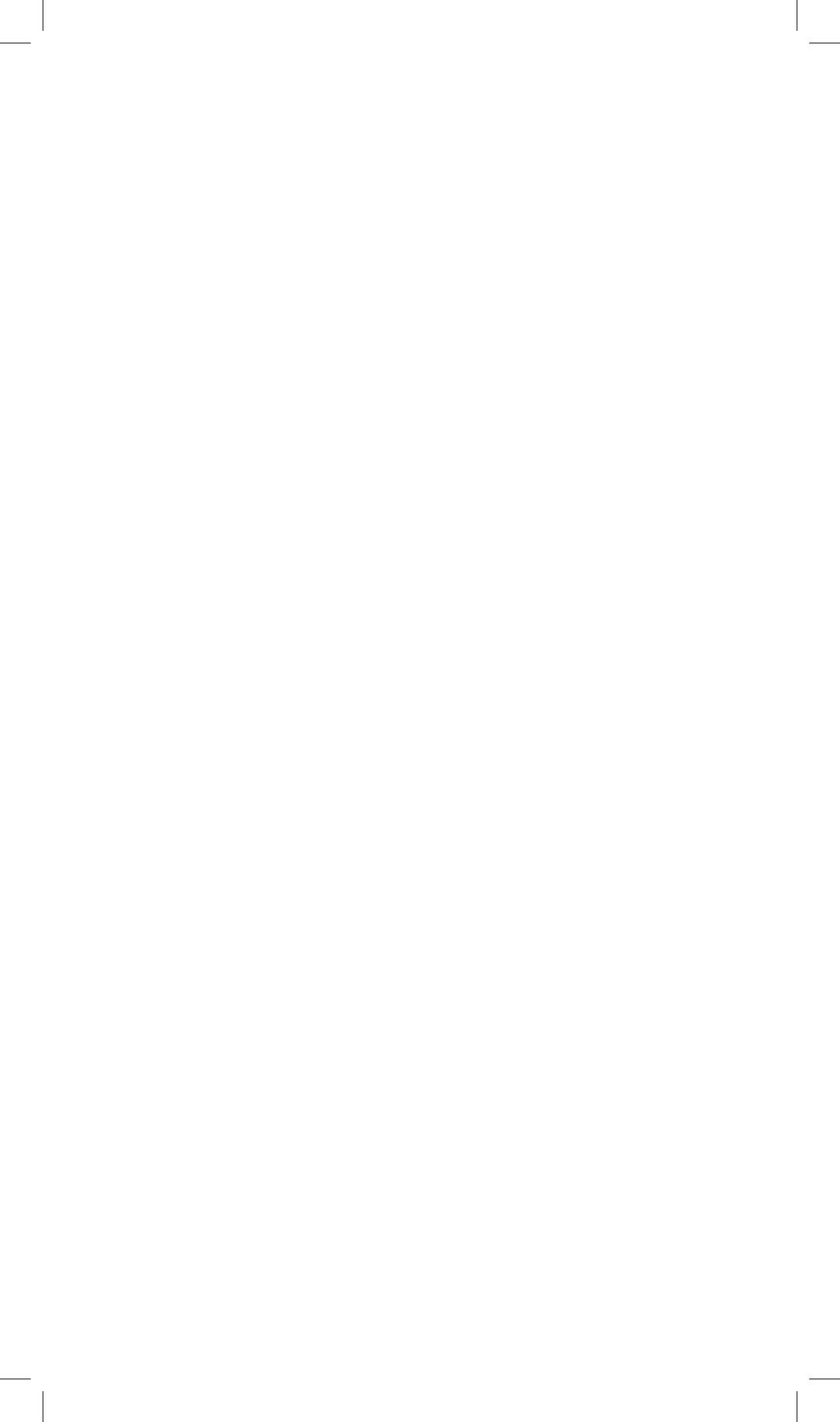
vie principali di trasmissione di italianismi: una via diretta e due vie indirette, prima attraverso il francese, poi attraverso il tedesco. È interessante vedere come viene visto come segno di integrazione del prestito non l'adattamento (*spagetti, ser mozzarella*) bensì il ripristino della forma italiana (*spaghetti, mozzarella*). Infine, Franco Pierno (*Gli italianismi nell'inglese di Toronto e nel francese di Montréal. Stato delle ricerche e progetti sugli italianismi in Canada*) segnala la "utilità relativa" del DIFIT per quanto riguarda l'inglese e il francese del Canada e propone un programma di ricerca per rilevare gli italianismi in queste due lingue.

«Al fine di porre le basi per la decisione sulle concrete strutture e sulle future prospettive del progetto OIM, Gesine Seymer e Raffaella Setti (*Ipotesi di cooperazioni e finanziamenti europei per l'OIM*), vagliando i punti di contatto del progetto OIM con attuali bandi, come quello del programma quadro dell'UE per l'innovazione e la ricerca "Horizon 2020", hanno avanzato delle proposte per la preparazione di una richiesta di finanziamento. In seguito a un'analisi degli obiettivi espressi dai molti bandi, quello dedicato a "Reflective Societies: Cultural Heritage and European Identities" è stato individuato il più idoneo. Il progetto consentirebbe l'avanzamento delle ricerche e una larga diffusione dei materiali e dei risultati raggiunti sia all'interno della comunità scientifica, sia verso il largo pubblico dei cittadini europei (e oltre). Il contatto linguistico tra l'italiano e le altre lingue europee sarebbe il nucleo fondante, con la finalità di stimolare la ricerca sulla migrazione lessicale storica e moderna in Europa e sul prestito intereuropeo, inteso come eredità linguistica e culturale condivisa. Partendo dalla versione in linea del DIFIT, il progetto coinvolgerebbe la maggior parte delle lingue ufficiali della UE, oltre ad alcune lingue storicamente rilevanti come il russo. Con capofila l'Università degli Studi di Firenze, il progetto, coordinato dallo studioso fiorentino Massimo Fanfani, potrebbe avvalersi di cooperazioni già esistenti dell'Accademia della Crusca e dell'Ateneo fiorentino con università e istituti

di ricerca all'estero. L'impianto del progetto dovrebbe inoltre prevedere misure specifiche per coinvolgere e sensibilizzare i cittadini europei su temi e fenomeni linguistici condivisibili e per tutti, in qualche misura, identitari: la continuità e la profondità dei contatti linguistici tra i paesi europei; la conoscenza delle radici comuni del lessico europeo; le enormi risorse offerte da un repertorio comune che può continuamente arricchirsi e rinnovarsi. La relazione è stata completata da considerazioni strategiche, organizzative e amministrative rilevanti per accedere ai fondi europei». Dopo questa domanda è stata avviata un'iniziativa per un progetto PRIN e facendo seguito a una domanda di finanziamento al MIUR sono intanto stati stanziati fondi per il progetto OIM che permettono l'ampliamento del corpus lessicografico, integrando gli italianismi dello spagnolo, portoghese e catalano e possibilmente di altre lingue.

Non vorrei concludere senza ringraziare l'Accademia della Crusca per aver accolto questi atti tra le sue prestigiose pubblicazioni.

HARRO STAMMERJOHANN
Francoforte sul Meno



DAL DIFIT ALL'OIM:
SFIDE LESSICOGRAFICHE
E PROSPETTIVE DI IMPLEMENTAZIONE*

1. INTRODUZIONE

Rossinissimo è il titolo di un festival di musica, organizzato da Cecilia Bartoli, che ebbe luogo nell'estate 2014 a Salisburgo (cfr. Fig. 1), ed è con gli attributi *Vielfaltissimo*, *Schickissimo*, e *Verwöhnissimo* che una torrefazione di caffè amburghese pubblicizza i suoi prodotti, tra cui una macchina per fare il caffè "all'italiana" con le capsule¹. In questi esempi le basi della derivazione, ted. *Vielfalt* (sost.), *schick* (agg.), *verwöhn(en, v.)* significano rispettivamente 'varietà', 'chic' e 'viziare', il suffisso, combinabile con classi di parole diverse, aggiunge una sfumatura di intensificazione e un riferimento, per quanto vago, a cose italiane. Queste formazioni non sono da ritenere lessicizzate, ma dimostrano come un elemento italianizzante quale il suffisso *-issimo* goda di grande popolarità nel vocabolario

* Ringrazio Harro Stammerjohann, Anne-Kathrin Gärtig e Elenora Doria per attente riletture di versioni precedenti di questo testo; rimane mia naturalmente la responsabilità per eventuali errori e sviste.

¹ Così almeno nelle sue campagne pubblicitarie in Svizzera e in Austria dal 2014. Gli esempi del tipo *Vielfaltissimo* (derivazione denominale), *Perfektissimo* (derivazione deaggettivale), *Verwöhnissimo* (derivazione deverbale) si riferiscono al marchio commerciale *Cafissimo* (in uso per capsule e apparecchio) dalla formazione ibrida (radice *Caf-* con la consonante scempia + *-issimo*). Un altro caso (segnalato da H. Stammerjohann), di formazione meno trasparente, è ted. *Profissimo*, nome di una linea di prodotti casalinghi di una catena di negozi tedesca, la cui base lessicale sembra essere il ted. *Profi* 'professionista'; trattandosi di una denominazione inventata a scopi commerciali possono aver prevalso considerazioni di natura puramente fonetica.

tedesco, perlomeno quello della pubblicità, e sembra fungere di per sé come indice di “italianità” degli oggetti, eventi o prodotti cui si abbina². A proposito di tale suffisso, il *Dizionario degli italianismi in francese, inglese, tedesco* (DIFIT) menziona, oltre alla sua funzione in italiano di «suffisso con cui si forma il superlativo degli aggettivi e, raramente, di un sostantivo», il suo uso in tedesco come «suffisso italianizzante aggiunto talora a parole tedesche» (DIFIT, s.v. ‘-issimo’), dove in genere la funzione di indicare qualcosa di tipico proveniente dall’Italia prevale su quella originale di superlativo.



Fig. 1. Uso del suffisso italianizzante *-issimo*
(locandina per un festival musicale a Salisburgo, Austria, giugno 2014)

Quanto segue si accinge a dare un breve abbozzo delle sfide lessicografiche e delle prospettive di implementazione che si sono venute delineando intorno all’impresa del DIFIT (il dizionario, uscito prima in versione cartacea nel 2008, iniziato

² La sequenza [‘is:imo] presenta tra l’altro il vantaggio del susseguirsi di strutture sillabiche ben riproducibili e mediamente più semplici dei nessi sillabici del tedesco.

e curato da Harro Stammerjohann con un gruppo di lavoro internazionale) e alla sua realizzazione digitale in seno al progetto dell'*Osservatorio degli italianismi nel mondo* (OIM). Si è già accennato al DIFIT come strumento di ricerca di cui si può giovare la messa a punto di prestiti del genere, sarà in seguito da precisare lo scopo del progettato osservatorio. Perciò si vorrebbe delineare la struttura del progetto lessicografico e possibili applicazioni del DIFIT, che nella presenza web dell'OIM si trova trasformato in uno strumento di ricerca digitale, adatto per ricerche di vario tipo, che oltre alle consulte lessicografiche e filologiche, per così dire classiche, possono vertere su diversi campi della linguistica³. Come indica il suo titolo, il DIFIT, concepito come risorsa lessicografica specializzata, si focalizza sul contatto linguistico tra l'italiano e tre delle maggiori lingue europee, ovvero il francese, l'inglese e il tedesco. I risultati lessicali di tale contatto sono presentati in prospettiva sia sincronica che diacronica. Il metalinguaggio lessicografico del DIFIT è l'italiano, ma il dizionario stesso è plurilingue, in quanto documenta il passaggio di elementi dalla lingua donatrice, l'italiano, alle lingue riceventi, registrando i prestiti e lemmatizzandoli sotto il rispettivo etimo italiano.

Le sigle DIFIT e OIM non sono intercambiabili, poiché si riferiscono a piani diversi di un'impresa dal proposito comune⁴. La sigla DIFIT sta dapprima per l'opera lessicografica,

³ Come strumento di ricerca che combina le potenzialità delle nuove tecnologie con i risultati del lavoro filologico-umanistico su materiali linguistici concreti, rappresenta un caso esemplare di nuovi approcci radunati sotto la rubrica delle cosiddette *digital humanities* (vi accenna, nella sua conclusione, anche Meyer, in questo volume).

⁴ Sono da attribuire denominazioni meno ambigue assegnando il nome di DIFIT *elettronico* alla versione del DIFIT già presente in rete (all'indirizzo www.italianismi.org) e chiamando OIMBase o similmente il motore di ricerca con la banca dati soggiacente in via di sviluppo come supporto per il riallestimento e ampliamento del corpus lessicografico, al fine di mantenere la sigla OIM come designazione del progetto complessivo (si tratta di uno dei progetti strategici dell'Accademia della Crusca coordinato da Luca Serianni e Matthias Heinz che attualmente raccoglie un gruppo di ricerca internazionale).

uscita in forma cartacea nel 2008, ma verrà usata in questa sede soprattutto per designare i dati lessicografici ivi raccolti in qualsiasi forma, consultabili ormai preferibilmente nella loro versione digitale, che sarebbe da denominare più precisamente DIFIT *elettronico* o DIFIT *in linea*⁵. Difatti, la consultazione del DIFIT è agevolata per la saggia elaborazione del corpus lessicografico esistente, trasferito al formato elettronico e adattato ai fini dell'utente, grazie al disegno del potente database concepito da Gesine Seymer in cooperazione con Marco Biffi, Giovanni Salucci e Maurizio Rago⁶. Il DIFIT costituisce dunque il nucleo del progetto OIM, di più ampio respiro rispetto al dizionario originale, dato che l'osservatorio mira a integrare molte più lingue riceventi.

2. DIFIT E OIM: DOCUMENTAZIONE DEL CONTATTO LINGUISTICO E IMPOSTAZIONE METALESSICOGRAFICA

L'opera lessicografica DIFIT, nella sua impostazione sin dall'inizio plurilingue e comparativa, attinge a una realtà linguistica ben nota: l'italiano (letterario) come lingua di cultura, insieme alle sue varietà dialettali, rappresenta una fonte estremamente ricca per il prestito linguistico. Mentre il suo influsso lessicale sulle altre tre lingue europee durante i diversi periodi storici degli ultimi sei secoli è stato studiato in tutta una serie di lavori, la documentazione lessicografica, finora piuttosto sparsa, con il DIFIT è stata raccolta per la prima volta in un'opera di riferimento che include più di una lingua di arrivo.

Per prestito (ingl. *borrowing*, ted. *Entlehnung*) intendiamo qui il risultato dell'imitazione di un modello linguistico al-

⁵ Il continuo processo di emendamento, di revisione e di allargamento dei dati del DIFIT dovrà infine sfociare in un DIFIT2, che grazie al formato elettronico si potrà pubblicare in varie fasi, sostituendo via via la versione precedente.

⁶ Ambientato in seno ad attività più vaste dell'Accademia della Crusca la banca dati dell'OIM si inserisce nel portale VIVIT (cfr. il riassunto di Marco Biffi nella prefazione di questo volume, il repository digitale VIVIT è consultabile online: <http://www.viv-it.org>).

loglotto da parte di una comunità linguistica⁷. Tale concetto include elementi come i seguenti:

- (1) it. *ciao* (→ fr., ingl., ted.)
- (2) it. *dolce far niente* (→ ingl., ted.)
- (3) it. suffisso *-issimo* (→ p.es. fr. *affairissimo*)
- (4) it. *eppur si muove!* (proverbi, fraseologismi ecc.)
(→ fr., ingl.; in ted. calco: *Und sie bewegt sich doch!*)
(DIFIT-OIM, s.vv.)

Gli elementi di prestito nel lessico di una lingua possono essere parole (monomorfemiche) come *ciao* (1), espressioni polirematiche alquanto complesse come (2) *dolce far niente*, o talvolta formativi (si tratta del fenomeno dell'«induzione di morfemi» di cui parla Gusmani 1986, 155) come il suffisso *-issimo*, che al significato di superlativo o perlomeno di intensificazione aggiunge la sfumatura italianeggiante (del tipo fr. *affairissimo* o ted. *Vielfaltissimo*) di cui sopra. Ma si hanno anche fraseologismi, espressioni proverbiali e idiomatiche, come *eppur si muove* (4), presente in francese e inglese, che in tedesco appare nella veste del calco *Und sie bewegt sich doch!*

Attualmente, il DIFIT registra 8951 italianismi e 4660 etimi italiani, di cui 242 senza attestazione precedente. Dopo un breve riassunto del dizionario e della sua versione digitale, si presenterà qualche esempio di come i dati raccolti possano essere utilizzati per l'analisi linguistica. Da un punto di vista metalessicografico e usando la classificazione proposta da WIEGAND 2001, il DIFIT può essere considerato un «aktives polylaterales Sprachontaktwörterbuch» ('dizionario di contatto linguistico attivo, polilaterale'); attivo in quanto la lemmatizzazione parte dall'etimo, e polilaterale perché registra prestiti di una lingua donatrice nelle diverse lingue riceventi. Inoltre, come si è detto sopra, l'opera è monolingue per

⁷ Cfr. tra gli altri GUSMANI 1986, 10 e *passim*, PINNAVAIA 2001, FANFANI 2011. Per il concetto di "italianismo" si rinvia a STAMMERJOHANN 2010.

quanto riguarda il suo metalinguaggio, l'italiano, mentre si può considerare multilingue per quanto riguarda l'insieme di lingue riceventi (francese, inglese, tedesco). Lo scopo generale del dizionario viene illustrato nell'introduzione:

[I]l presente *Dizionario* [...] vuole offrire più che una semplice somma dei dati finora raccolti e acclarati. Il suo **scopo** è più specifico, è quello di **mettere a confronto l'incidenza dell'italiano sul francese, l'inglese e il tedesco**, le tre lingue che sono al centro dello spazio europeo e sono a più stretto contatto tra loro, con l'intento di **ricostruire le trafe di penetrazione e la diversa sorte delle parole italiane** in questo circuito. (DIFIT: XI; neretto MH)

Prendendo a mo' d'esempio una delle lingue considerate, il caso del francese illustra chiaramente sia l'importanza dell'italiano come lingua modello che la necessità di opere specializzate. Da una ricerca nel maggior dizionario di riferimento per il francese moderno, il *Trésor de la langue française* (TLF), risulta l'incidenza dell'apporto italiano; in quanto lingua di contatto l'idioma di Dante è secondo per numero di unità mutuate solo all'inglese (cfr. Fig. 2)⁸:

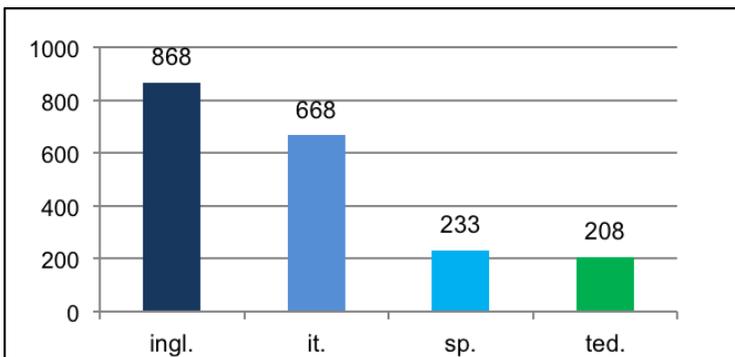


Fig. 2. Prestiti in francese per lingue modello nel TLFi

⁸ Applicando il filtro 'langue empruntée' (lingua ricevente) nella ver-

Nei confronti della pur dettagliata lessicografia generale, si deve ad una risorsa specializzata quale il DIFIT un ragguardevole ampliamento della documentazione di italianismi (cfr. Fig. 3); rispetto ai 668 lemmi del TLFi il numero dei prestiti registrati dal DIFIT (escludendo i casi di attribuzione molteplice o incerta) risulta più che triplicato.

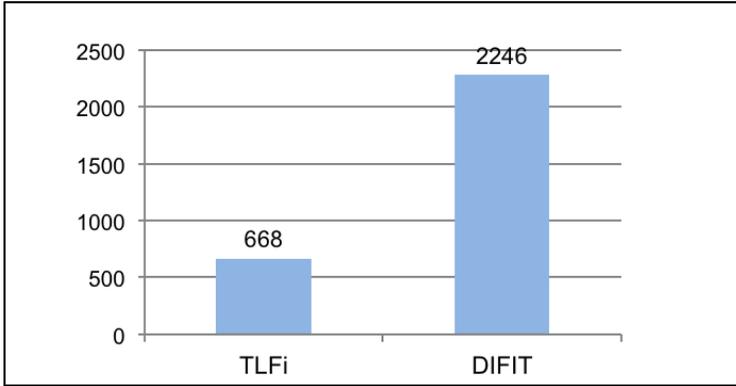


Fig. 3. Italianismi in francese: TLFi vs. DIFIT

Il DIFIT viene attualmente ampliato all'interno del progetto OIM. *L'Osservatorio degli italianismi nel mondo*, progetto collaborativo internazionale che si svolge in seno all'Accademia della Crusca, ha due obiettivi principali:

- 1) rendere accessibili ed editare i dati raccolti nel DIFIT in una banca dati online
- 2) estendere l'inventario delle lingue di contatto oltre le tre grandi lingue europee registrate nel DIFIT, francese, inglese e tedesco, puntando sulla descrizione lessicografica dell'impatto dell'italiano come lingua modello, ad es. in spagnolo, polacco, giapponese ecc.

sione digitalizzata del dizionario (*Trésor de la langue française informatisé*, TLFi) si ottengono i valori per le singole lingue.

L'obiettivo di questa impresa è di creare un osservatorio linguistico sull'impronta che ha lasciato, e che continua a lasciare, l'italiano nelle lingue del mondo, tramandando prestiti a vari livelli dei sistemi linguistici. Prendendo spunto dalla documentazione lessicale, l'OIM abbraccia tuttavia una visione più ampia, inclusiva di fenomeni idiomatici e morfosintattici⁹.

3. IL DIZIONARIO DEL CONTATTO LINGUISTICO COME STRUMENTO PER LE ANALISI LINGUISTICHE

Configurandosi «come una grande banca dati, di cui le voci già ereditate dal DIFIT costituiscono la base» (riprendendo la descrizione di Marco Biffi citata nella presentazione di questo volume), la realizzazione informatica dell'OIM offre una vasta gamma di opzioni di ricerca. A parte la ricerca libera, sono possibili ricerche complesse nell'elenco degli etimi (con l'opzione 'ricerca negli etimi'), applicando e combinando i seguenti criteri:

- etimi italiani (+ varianti)
- categoria grammaticale (etimo)
- anno di prima attestazione (etimo), con l'opzione di delimitazione del lasso di tempo
- arcaico / obsoleto / senza attestazione lessicografica
- uso settoriale / campo lessicale (etimo)
- origine dialettale (etimo)
- registro
- lingua ricevente (per anno o lasso di anni + una, due o tutte e tre le lingue riceventi)

⁹ Cfr. anche l'osservazione conclusiva di Pierno, in questo volume. Alcune prime verifiche su fatti morfosintattici verranno presentati in seguito. In una fase più avanzata del progetto sarà possibile studiare, attraverso il processo dell'adattamento fonico, anche la fonologia dei prestiti (per ora i dati del DIFIT contengono singoli casi di pronuncia marcata).

Le fonti che documentano gli etimi si possono visualizzare in una finestra con l'intera indicazione bibliografica, passando il cursore sulla rispettiva abbreviazione.

L'opzione "ricerca negli italianismi" permette ricerche complesse attraverso i seguenti percorsi di ricerca:

- italianismi (+ varianti)
- categoria grammaticale
- lingua ricevente
- anno di prima attestazione, con l'opzione di delimitazione del lasso di tempo
- arcaico / obsoleto / senza attestazione lessicografica
- uso settoriale / campo lessicale (etimo)
- uso dialettale
- registro
- tipo di prestito:
 - integrale
 - parziale (calco: formale / parziale / semantico)
 - diretto
 - indiretto (con lingua di mediazione)
 - pseudoprestito
- prestiti con origine (italiana) incerta o poco chiara (per esempio informazione etimologica contraddittoria in fonte autorevole)

Anche in questo caso, l'intero riferimento bibliografica delle fonti per gli italianismi si ottiene passando il cursore sull'abbreviazione.

Passando in rassegna tre campi di applicazione esemplari, si daranno in seguito brevi illustrazioni di come i dati del DIFIT possano essere utilizzati per la ricerca linguistica. Sono ricerche portate a termine dapprima sulla base di un computo semimanuale o, in alcuni casi, addirittura manualmente, in base a versioni preliminari dei dati elettronici (ad esempio un documento HTML con tutti i lemmi del DIFIT), i risultati furono poi verificati con l'aiuto del motore di ricerca messo a

disposizione *in toto* nel frattempo.

Ne emergono alcuni primi risultati collegati ad una serie di domande di interesse per la linguistica (soprattutto la linguistica di contatto), nella fattispecie la tipologia dei prestiti, i tipi di formazione delle parole nei lemmi polirematici identificati nel corpus del DIFIT, e, come ulteriore prospettiva, la visualizzazione dei risultati del contatto linguistico con lo strumento OIMap.

3.1. *Tipologia dei prestiti*

Categorizzare e quantificare i diversi tipi esistenti, oltre ai prestiti “classici” – come il calco formale o semantico, i rari casi di pseudo-prestito insieme alle formazioni ibride ecc. – ci dà un’immagine generale dell’impatto dei processi formali e semantici sull’evoluzione diacronica e la stratificazione sincronica del lessico nelle lingue riceventi.

GENERALISSIMO



GENERALISSIMO s. m. [mi./ami](#) [stor.](#) Sec. XVI. Un tempo, comandante supremo di un esercito. (DEL).

F **généralissime** (calco formale) s. m. [mi./ami](#) 1600-12 (TLF).

I **I. generalissimo** s., pl. -os [mi./ami](#) 1621 (OED).

II. generalissima (pseudoprestito) s. f. [mi./ami](#) 1643 (OED).

T **Generalissimus** (< it. × lat. -US) s. m., gen. -, pl. -mi e -musse [mi./ami](#) 1674-78 (DuGW; Hechtenberg).

Fig. 4. Diversi tipi di prestito nelle lingue riceventi (DIFIT-OIM, s.v. ‘generalissimo’)

Come illustrato nel caso di un lemma come *generalissimo*, la formazione del prestito che risulta da un singolo etimo può essere caratterizzata da una varietà di risultati. Qui il francese presenta un calco formale, l’inglese un pseudo-prestito basato su un modello strutturale italiano, mentre il tedesco *Generalissimus* è identificato come combinazione ibrida di una base italiana e un suffisso latino/latineggiante.

Tipo di prestito	%
Prestito diretto	94,5
Calco formale	3,3
Calco parziale	0,8
Calco semantico	1,3
Pseudo-italianismo	0,1

Tab. 1. Distribuzione dei tipi di prestito in DIFIT-OIM

La Tabella 2 indica le percentuali dei diversi tipi di prestito. I prestiti diretti – prestiti in senso stretto – rappresentano nettamente il tipo più comune tra gli italianismi registrati nel DIFIT, mentre il tipo più frequente di calco sono le rese parola per parola di un modello straniero, per es. it. *monte di pietà* > fr. *mont-de-piété*. Seguono i calchi parziali (it. *fare fiasco* > tedesco *Fiasko machen*) e i calchi semantici (it. *futurista* > ingl. *futurist*). Con il solo 0,1% nella documentazione del DIFIT (sono 6 i casi in tutto il dizionario) sono estremamente rari i pseudo-italianismi¹⁰, come l'inglese *generalissima* o il lemma *tutti-frutti*, presente sia in francese che in inglese e tedesco seppure con significati diversi.

Inoltre, il motore di ricerca OIM offre la possibilità di creare indici tematici, basati sui diversi settori del lessico da cui provengono italianismi. Grazie a funzioni che permettono combinazioni di criteri di ricerca è anche possibile elencare i risultati in ordine cronologico e/o a seconda delle lingue d'arrivo¹¹.

¹⁰ Per il *pseudo-prestito* (anche *falso prestito*, ingl. *pseudo-loan*, ted. *Scheinentlehnung*) cfr. WINTER-FROEMEL 2011 (pp. 44-45) che dà l'esempio di ted. *picobello* 'pulito e ordinato in modo impeccabile' per formati che sembrano presi in prestito da un'altra lingua, ma non esistenti nella (presunta) lingua fonte; a proposito del fenomeno dei "false anglicisms" in italiano e delle questioni terminologiche cfr. FURIASSI 2010 e 2015.

¹¹ Si rinvia a STAMMERJOHANN-SEYMER 2007 (p. 51) per una visualizzazione delle vicissitudini dell'influsso lessicale italiano sulle tre lingue prese in considerazione dal DIFIT. Per gli italianismi del solo inglese, Durkin (2014, p. 371) offre un simile quadro sinottico lungo l'asse diacronico in base ai

3.2. Formazione delle parole e prestiti

Dei 4660 lemmi (etimi italiani) del DIFIT, 4161 sono parole semplici, mentre 499 – più del 10% – sono classificati come locuzioni, raccogliendo espressioni polirematiche o espressioni multiparola (ricalcando il termine inglese *multiword expressions*), tra i quali anche unità sintagmatiche complesse o sintagmi come *con spirito* e composti lessicalizzati come *pesce spada*. Dal computo semiautomatico e la susseguente classificazione manuale dei tipi di formazione emergono come strutture preferite, per ordine di frequenza, i sintagmi aggettivali, preposizionali e nominali (più del 80% dei rispettivi lemmi, gli altri tipi sono da ritenere marginali in quanto ripartiti su cifre di di occorrenza esigue):

Struttura	Esempi	%
[N + A]	<i>salto mortale, sinfonia concertante, opera comica</i>	29,9
[A + N]	<i>dolce vita, sacra conversazione, onorata società</i>	8,6
[Prep + N]	<i>a battuta, a conto, con spirito</i>	14,7
[Prep + Art + N]	<i>al dente, alla marinara, dal segno, sul tasto</i>	11,7
[N + Prep + N]	<i>giorno di respiro, lira da gamba, zuppa di pesce</i>	11,5
[N + Prep + Art + N]	<i>commedia dell'arte, commercio all'ingrosso, maestro al cembalo</i>	3,2
[N + N]	<i>pesce spada, stucco lustro, pensione baby</i>	2,2
(...)		

Tab. 2. Distribuzione dei lemmi polirematici in DIFIT-OIM (etimi)

Si osserva che nelle locuzioni la struttura sintagmatica

dati dell'OED. Con le potenzialità del database OIM diventano possibili delle ricerche ancora più raffinate (ad esempio suddividendo ulteriormente gli intervalli storici).

N+A è il formato più frequente in assoluto, rispecchiando la posizione 'default' dell'aggettivo (salvo un gruppo ristretto di aggettivi regolarmente preposti e usati marcati) in italiano e francese, con il 29,9% prevale nettamente sull'ordine inverso A+N, canonico in inglese e in tedesco. È comunque notevole l'incidenza dei sintagmi aggettivali, che prese insieme contano per quasi il 40% delle espressioni multiparola nel corpus. Ciò da prova di una particolare forza predicativa delle sequenze composte da nome e aggettivo la quale sembra renderle *pat-tern* attraenti nel contatto di lingue.

È un caso interessante il composto dalla struttura interna N+N *pensione baby*, ricalcato dall'espressione inglese *baby pensions* (a quanto pare, un occasionalismo, plurale tantum, il riferimento è alle pensioni maturate da persone giovani alla fine di una breve attività lavorativa), il cui secondo elemento è a sua volta mutuato dall'inglese, invertendo la sequenza determinato-determinante della forma modello. Ma lo stesso italiano usa anche l'ordine determinante-determinato in formati del tipo *baby pensionati* (anche *baby-cantante* e altri)¹². In questi casi *baby*, quale primo o secondo elemento nominale della composizione, si usa in funzione di aggettivo (similmente a *calcio-baby*).

In un passo successivo (che richiede tuttora dei passaggi di revisione manuale), tali dati – come, tra l'altro, l'intera banca dati – potranno essere esaminati a seconda delle singole lingue riceventi, cosa che permetterà un confronto tra le dinamiche dei processi di integrazione in ognuna delle lingue in questione.

¹² Nel rispettivo lemma del DIFIT si indica l'etimo *pensione baby* in base alla sua lemmatizzazione nel GRADIT, eppure si trovano in Internet (su pagine in lingua italiana) anche molte attestazioni (seppure non lessicografiche) della formazione *baby pensione* (specie al plurale), menzionata anche da Viviani (2010).

3.3. *Visualizzare la distribuzione areale degli italianismi: OI-Map*

Come ulteriore prospettiva, prima di concludere, vorrei brevemente presentare uno strumento per visualizzare la distribuzione – e la dinamica – areale degli italianismi, ovvero un mapping tool, attualmente in progetto sotto il nome di OIMap. Altri esempi di visualizzazione in questa direzione sono ADIS, diretto da Krefeld e Lücke, e WOLD. Questo strumento è finalizzato alla rappresentazione cartografica delle dinamiche del contatto linguistico rivelate dai dati del DIFIT, ed è in via di elaborazione presso l'università di Salisburgo. In effetti, mentre gli elenchi di unità lessicali prestatati in un dizionario o in una banca dati forniscono informazioni sfaccettate in maniera condensata, le mappe permettono di capire meglio la distribuzione geografica e le tendenze dinamiche delle relazioni di contatto tra le lingue.

OIMap si basa su un software tool libero per la cartografia digitale (<http://batchgeo.com>) e intende delineare la direzione del contatto linguistico. Inoltre, è possibile ottenere l'origine dialettale degli elementi presi in prestito, raffinando le opzioni di ricerca. Pertanto, è possibile individuare varietà dialettali che hanno contribuito un numero relativamente alto di prestiti al lessico di una, due o di tutte e tre le lingue riceventi. Se prendiamo il caso del veneziano (o, con un numero in confronto più basso di unità lessicali, il genovese), emerge un certo quadro delle dinamiche areali (cfr. fig. 5). Nella mappa della figura 2, una freccia indica le direzioni principali dei prestiti. Una maggioranza assoluta (per il genovese: 7 su 9) o relativa (per il veneziano: 15 su 45) di elementi di prestito “migra” esclusivamente verso una sola lingua (rispettivamente il francese e il tedesco).

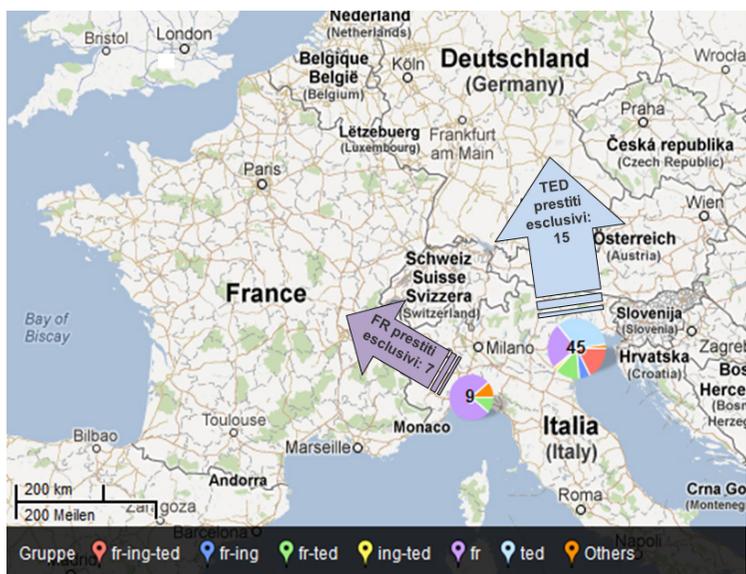


Fig. 5. OIMap – Dinamiche areali di prestiti genovesi e veneziani

Per molte lingue esistono già delle liste elettroniche di italianismi, seppure non tutte pubblicate¹³ e disponibili finora in formati diversi, quindi l'integrazione del materiale esistente nella banca dati OIM dovrebbe presto portare ad un progresso sostanziale – con l'incrementazione della banca dati e l'aumentare del numero delle lingue riceventi diventerà possibile creare un'interfaccia con lo strumento cartografico OIMap. Il suo proposito è quello di creare una mappa delle varie regioni o macroaree linguistiche, indicando la migrazione lessicale che parte dall'italiano e dalle sue varietà.

¹³ Cfr. i lavori nell'ambito di una precedente impresa volta a documentare gli italianismi in un gran numero di lingue, ai quali accenna il contributo di Serianni in questo volume. Grazie al coordinamento di Lucilla Pizzoli, che collaborò già al progetto di allora, e la disponibilità di molti dei redattori coinvolti anteriormente in esso ne potranno essere valorizzati preziosi materiali già raccolti.

4. OSSERVAZIONI CONCLUSIVE

Riassume a pennello le ragioni del DIFIT e del progetto OIM la caratterizzazione che dà dell'italiano il linguista e romanista Georg Bossong nel suo manuale sulle lingue romanze dalla prospettiva comparativa e tipologica:

Das Italienische ist so etwas wie die klassische Sprache Europas. Unter allen romanischen Nationalsprachen steht es dem Lateinischen am nächsten; das antike Erbe war hier stets präsenter als anderswo. Eine Weltsprache im politischen Sinn war es nie, Universalsprache von Kunst und Musik ist es jedoch schon früh geworden. Obgleich es bezüglich Sprecherzahl und Verbreitung nur einen mittleren Rang einnimmt, ist seine historische und kulturelle Bedeutung immens¹⁴ (BOSSONG 2008, p. 197).

Altro che “Weltsprache”, cioè lingua veicolare a livello mondiale, alla lingua italiano è da ascrivere la dignità di una “Kulturweltsprache”, ossia una lingua diffusa a livello mondiale in numerosissimi contesti legati alla cultura, ed è tale potenza di irradiazione culturale che le conferisce il suo particolare, forse unico peso¹⁵ nel concerto delle lingue del mondo. Attualmente si prepara, con la loro messa a giorno e impostazione per il supporto elettronico, l'integrazione di dati

¹⁴ «L'italiano può essere definito come la lingua classica dell'Europa. Tra tutte le lingue romanze diventate lingue nazionali è quella rimasta più vicina al latino; il patrimonio dell'antichità vi è sempre stato più presente che altrove. Non essendo mai stata lingua mondiale in un senso politico essa diventò molto presto lingua universale dell'arte e della musica. Seppure solo di rango medio quanto a numero di parlanti e diffusione, il suo peso storico e culturale è immenso» (Trad. MH).

¹⁵ Più precisamente tale peso risulta non dalla mera somma assoluta dei prestiti presenti in altre lingue, bensì anche dal peso individuale di questi prestiti nelle lingue d'arrivo (per determinare ciò ci si può rifare ai criteri menzionati più avanti da Serianni, nelle conclusioni dell'articolo presente in questo volume).

sulle lingue iberoromanze (spagnolo, portoghese, catalano) nonché il polacco e l'ungherese, dischiudendo a mano a mano il quadro degli intricati contatti di questa *Kulturweltsprache* con le lingue del mondo.

MATTHIAS HEINZ
Universität Salzburg

Bibliografia

- ADIS = Thomas Krefeld - Stephan Lücke, *Atlante linguistico digitale dell'Italia e della Svizzera meridionale (AdIS)*, 2012 (<http://www.adis.gwi.uni-muenchen.de/AIS.php>).
- BOSSONG 2008 = Georg Bossong, *Die romanischen Sprachen*, Hamburg, Buske.
- BRAUDEL 1989 = Fernand Braudel, *Le modèle italien*, Paris, Arthaud.
- DIFIT = Harro Stammerjohann *et al.* (a cura di), *Dizionario di italianismi in francese, inglese, tedesco*, Firenze, Accademia della Crusca, 2008.
- DURKIN 2014 = Philip Durkin, *Borrowed Words. A History of Loanwords in English*, Oxford, O.U.P.
- FANFANI 2011 = Massimo Fanfani, *Prestiti*, in R. Simone (a cura di), *Enciclopedia dell'italiano*, Roma, Treccani ([http://www.treccani.it/enciclopedia/prestiti_\(Enciclopedia-dell'Italiano\)](http://www.treccani.it/enciclopedia/prestiti_(Enciclopedia-dell'Italiano))).
- FURIASSI 2010 = Cristiano Furiassi, *False Anglicisms in Italian*, Monza, Polimetrica.
- FURIASSI 2015 = C. Furiassi, *False borrowings and false Anglicisms: Issues in terminology*, in C. Furiassi - H. Gottlieb (a cura di), *Pseudo-English: false Anglicisms in Italian*, Berlin, De Gruyter Mouton, pp. 253-77.
- GRADIT = *Grande Dizionario Italiano dell'Uso*, ideato e diretto da Tullio De Mauro, 6 voll. Torino: UTET, 2000 (+ *Nuove parole italiane dell'uso* 2003). [La sigla usata nel DIFIT è GDU.]
- GUSMANI 1986 = Roberto Gusmani, *Saggi sull'interferenza linguistica*, Seconda edizione accresciuta, Firenze, Le Lettere.
- HEINZ 2008 = Matthias Heinz, *L'expérience du Dizionario di italia-*

- nismi in francese, inglese, tedesco (DIFIT): objectifs, structure et aspects méthodologiques*, in F. Pierno (a cura di), *Aspects lexicographiques du contact entre les langues dans l'espace roman*, Strasbourg, Université Marc Bloch, pp. 165-80.
- OED = *The Oxford English Dictionary Online*, a cura di J.A. Simpson, Oxford, O.U.P., 2000- (<http://www.oed.com/>).
- OIM = *Osservatorio degli italianismi nel mondo*, progetto dell'Accademia della Crusca, coord. Luca Serianni - Matthias Heinz, con la collaborazione di Marco Biffi - Domenico De Martino - Nicoletta Maraschio - Giovanni Salucci - Gesine Seymer - Harro Stammerjohann et al. [Firenze, Accademia della Crusca, 2014-]: www.italianismi.org [30/11/2015].
- PINNAVAIA 2001 = Laura Pinnaivaia, *The Italian Borrowings in the Oxford English Dictionary: A lexicographical, linguistic and cultural analysis*, Roma, Bulzoni.
- STAMMERJOHANN-SEYMER 2007 = Harro Stammerjohann - G. Seymer, *L'italiano in Europa: italianismi in francese, inglese e tedesco*, in N. Maraschio (a cura di), *Firenze e la lingua italiana fra nazione ed Europa*, Firenze, Firenze University Press, pp. 41-55.
- STAMMERJOHANN 2010 = H. Stammerjohann, *Italianismi*, in *Enciclopedia dell'italiano*, Roma, Treccani ([http://www.treccani.it/enciclopedia/italianismi_\(Enciclopedia-dell'Italiano\)](http://www.treccani.it/enciclopedia/italianismi_(Enciclopedia-dell'Italiano))).
- TLF = [Paul Imbs - Bernard Quemada (sous la direction de)], *Trésor de la langue française. Dictionnaire de la langue du XIX^e et du XX^e siècle (1789-1960)*, vol. 1 (1971)-10 (1983), Paris, Éditions du Centre National de la Recherche Scientifique; voll. 11 (1985)-16 (1994), Paris, Gallimard. Versione online *Trésor de la langue française informatisé (TLFi)*, a cura di J. Dendien: <http://atilf.atilf.fr/tlf.html>.
- VIVIANI 2010 = Andrea Viviani, *Itangliano*, in R. Simone (a cura di), *Enciclopedia dell'italiano*, Roma, Treccani ([http://www.treccani.it/enciclopedia/itangliano_\(Enciclopedia-dell'Italiano\)](http://www.treccani.it/enciclopedia/itangliano_(Enciclopedia-dell'Italiano))).
- WIEGAND 2001 = Herbert Ernst Wiegand, *Sprachkontaktwörterbücher: Typen, Funktionen, Strukturen*, in «Germanistische Linguistik», 161/162, pp. 115-224.
- WINTER-FROEMEL 2011 = Esme Winter-Froemel, *Entlehnung in der Kommunikation und im Sprachwandel*, Berlin/New York, de Gruyter.

WOLD = *World Loanword Database* [a cura di M. Haspelmath - U. Tadmor], [2009-]: <http://wold.clld.org/meaning/4-2#2/24.4/-5.1> [28/01/2015].

Sitografia

<http://batchgeo.com> [28.01.2015]

<http://www.viv-it.org> [20.01.2016]



L'ITALIANO NEL MONDO.
INTENTI E PROPOSITI
DI UN PROGETTO EDITORIALE
SUGLI ITALIANISMI

Il sottotitolo del mio intervento avrebbe potuto essere *Storia di un fallimento*. Tale è stato infatti l'esito del progetto al quale ho posto mano insieme con Lucilla Pizzoli e Leonardo Rossi nel dicembre 2004 per la UTET. L'idea era quella di censire gli italianismi in un numero molto alto di lingue (erano 48 quelle previste inizialmente nel contratto; 78 quelle selezionate in corso d'opera), dando vita a un dizionario, corrispondente al secondo volume. Tra le lingue censite si era dato rilievo alle varietà (francese di Francia e Svizzera, inglese di Gran Bretagna, Stati Uniti, Canada, Australia, spagnolo di Spagna, Messico, Cuba, Colombia, Bolivia, Cile, Ecuador, Perù, Argentina, Uruguay, Paraguay, portoghese di Portogallo e Brasile, tedesco di Germania, Austria e Svizzera): è evidente l'interesse dei contatti linguistici legati all'immigrazione italiana, non solo per le aree di più forte immigrazione (e con più antica tradizione di studi) come l'Argentina. Il primo volume sarebbe stato costituito da saggi divisi in tre sezioni: la prima introduttiva (il saggio iniziale, puntualmente consegnato, era di Harro Stammerjohann ed è fortunatamente leggibile in versione molto più ampia nel suo splendido *La lingua degli angeli*, pubblicato l'anno scorso dalla nostra Accademia); una seconda dedicata alle grandi aree storico-geografiche (con diverso sviluppo a seconda dei casi: ben due contributi di differenti autori per la Francia, in assoluto l'area con la quale l'Italia nelle sue varie realtà storico-culturali ha intrattenuto

i rapporti più intensi e più costanti nel tempo; uno per due piccole nazioni come Svizzera e Albania, diversamente vicine all'Italia non solo per ragioni geografiche; alcuni su macroaree, caratterizzate da rapporti più laschi ed episodici ma anche scarsamente esplorate dalla bibliografia precedente, come l'India e il Sud Est asiatico o l'Africa subsahariana); una terza, infine, che si concentrava sulle aree semantiche salienti: dalla musica alla gastronomia, dall'arte allo spettacolo cinematografico e televisivo, dalla religione allo sport.

In tutto furono individuati più di 60 autori per i saggi del primo volume e più di 80 rilevatori per il censimento linguistico destinato al secondo volume (solo eccezionalmente, come nel caso di Giuseppe Brincat per il maltese, le due figure coincidevano nella stessa persona).

Come avviene in opere che coinvolgono tanti autori diversi, il rischio del ritardo delle consegne era messo in campo dallo stesso editore al momento del contratto (l'art. 4 prevedeva che l'acquisizione dei collaboratori si concludesse entro il giugno 2005 e l'editore dichiarava il proposito di impegnarli «a consegnare i testi *possibilmente* entro il giugno 2006»: l'avverbio usato è ben indicativo). Ma, a distanza di due anni dalla scadenza divisata, le tessere mancanti erano ancora troppe; e, nel frattempo, la casa editrice UTET aveva mutato politica editoriale e non era più interessata a proseguire l'opera oltre la scadenza contrattuale. Nell'ottobre 2008 l'esperienza fu dunque formalmente interrotta.

I ritardi nelle consegne sono abituali in iniziative che coinvolgono più studiosi: qualche volta per negligenza dei singoli, ma più spesso per i tempi stessi della ricerca, che mal si presta a essere confinata entro limiti temporali rigidi e che, di norma, convive con altri interessi e altre attività di studio da parte del singolo ricercatore. Per limitarci all'Accademia della Crusca, possiamo ricordare il convegno *Lingua italiana e scienze* tenutosi nel febbraio 2003, i cui atti sono stati pubblicati solo nel 2012. Ma tant'è: le esigenze di un editore commerciale sono legittimamente diverse da quelle di un'istituzione culturale.

Ci si può chiedere, con la lucidità con cui si possono guardare le cose a distanza di tempo, se, a parte le vicissitudini appena ricordate, il progetto non avesse una sua intrinseca debolezza legata all'ambizione delle sue dimensioni; se non fosse, cioè, il classico gigante dai piedi d'argilla. Non mi si chieda di rispondere a questa domanda e basti aver posto il problema.

I saggi del primo volume permettevano al singolo autore un'esecuzione personale, condizionata dalla storia degli studi, ma anche dalle sue personali esperienze di ricerca; e insomma potevano essere legittimamente orientati più sulla linguistica esterna che non su quella interna o viceversa. Molto più stringente era, com'è naturale, il percorso da compiere per i compilatori del dizionario.

Era previsto un prospetto iniziale che permettesse di raccogliere, per ciascuna lingua indagata, alcune informazioni sistematiche: il nome con cui una lingua X è conosciuta in italiano (*spagnolo* o *castigliano*?); qual è la varietà di volta in volta considerata (il problema si pone, per esempio, nel caso del norvegese); sommarie caratteristiche della lingua X (area in cui viene usata, numero di parlanti, tipologia...); situazione lessicografica ed eventuale disponibilità di studi sugli italianismi o, in generale, sui forestierismi presenti nella lingua X, insieme con indicazioni sui criteri seguiti per la raccolta degli italianismi. Quest'ultima serie di indicazioni è particolarmente delicata: per lingue di grande tradizione culturale, come francese o spagnolo, disponiamo di strumenti sofisticati e siamo in grado di ricostruire con un buon margine di precisione la storia di un possibile italianismo in esse attestato; ma in altri casi le fonti lessicografiche sono precarie. Si può dire che ognuna delle lingue che per pura comodità potremo chiamare "minori" (minori ovviamente in prospettiva italo-centrica) faccia storia a sé. Per il kannada, una lingua dravidica dell'India meridionale parlata da circa 35 milioni di persone, Tiziana Ripèpi ci informa che non esistono né dizionari etimologici, né dizionari monolingui su supporto

informatico, né ricerche su eventuali prestiti dall'italiano (le lingue sondate in proposito sono arabo, persiano, portoghese), ma solo un piccolo dizionario dei prestiti, *Eravalu padakōśa*, che ha permesso di individuare 25 prestiti dall'inglese per i quali l'italiano rappresenta la fonte remota; poco rappresentativi, invece, i termini ricavati da un'enciclopedia, la *Kannaḍa viśvakōśa*, che dedica un articolo abbastanza ampio all'Italia: ma vocaboli come *byārōk* 'barocco' o *ḍalsiṣṭil nyūvo* 'dolce stil novo' non sono che occasionalismi specialistici senza nessuna ricaduta sulla lingua d'arrivo. Nel caso dell'azero, la lingua del gruppo turco parlata in Azerbaigian e altrove da circa 50 milioni di parlanti, Farid Alakbarli si è fondato, oltre che su un dizionarietto di servizio (*Italian-Azeri and Azeri-Italian Dictionary*, Baku 2006), su un ampio repertorio dell'azero letterario, l'*Orthographical Dictionary of Azeri*, Baku 2004, che però, per riprendere le parole del rilevatore, «fornisce unicamente una lista alfabetica dei termini senza alcuna spiegazione quanto all'origine e al significato». Sono emersi 129 prestiti, in gran parte mediati attraverso il russo: interessanti alcune forme derivate, che potremmo considerare "italianismi di secondo grado", ossia forme sviluppatasi all'interno dell'azero, ma partendo da un italianismo; e sono insieme segno della vivacità derivativa della lingua ricevente, ma anche dell'avvenuta acclimazione del prestito (per esempio *Fontanlamaq* 'sgorgare con forza', creato a partire da *Fontan* 'fontana' e *Fleytaçalan* 'flautista', formato da *Fleyta* 'flauto' e da *çalan* 'suonare').

Il prospetto si concludeva, prima delle schede di servizio (sistema usato per l'eventuale traslitterazione, sigle bibliografiche...), con un paragrafo dedicato alla valutazione d'insieme degli italianismi: numero, àmbiti semantici più rappresentati, attuale vitalità dei prestiti ecc.

Questo apparato di informazioni, identico per le varie lingue, aveva l'intento di omogeneizzare ma anche di relativizzare i dati, permettendo di controllare il percorso seguito dal singolo ricercatore: chi abbia la tentazione di esaltare la presenza della letteratura e dell'arte italiane in kannada grazie

a *byārök* o *dalsisṭil nyūwo* può accantonarla alla luce della fonte che documenta quei termini. Che i dati così raccolti mantenessero comunque un margine di aleatorietà è indubbio: ma il vantaggio di disporre di informazioni non facilmente reperibili, anche perché presuppongono specializzazioni molto diverse tra loro, mi sembra altrettanto indiscutibile.

Negli ultimissimi anni la nostra conoscenza in fatto di italianismi nel mondo si è arricchita in modo notevole. Ho già citato la monografia di Stammerjohann, ricchissima di dati e di spunti per le lingue più diverse, ed è inutile ricordare dello stesso autore (con collaboratori) il *Dizionario di italianismi in francese, inglese, tedesco* (DIFIT), pubblicato dalla nostra Accademia nel 2013. Fondamentale è altresì il saggio di Emanuele Banfi (BANFI 2012), che costituisce il secondo volume degli Atti di un convegno della Società di linguistica italiana e che rappresenta un'articolata storia della fortuna dell'italiano in Europa e nel Mediterraneo.

Ma è giusto riconoscere che dal tronco reciso del *Dizionario* progettato per l'UTET sono nati frutti di tutto rispetto, perlopiù in versione generalmente arricchita, e talvolta diversamente orientata, anche perché libera dai vincoli inevitabilmente imposti a un'impresa editoriale a più mani (ed è probabile che qualche titolo mi sia sfuggito).

Ricorderò, dai saggi della prima sezione, ben tre volumi, che si aggiungono alla *Lingua degli angeli* già menzionata: la riflessione d'insieme di Maurizio Dardano (DARDANO 2011), la monografia di Furio Brugnolo (BRUGNOLO 2009) e un saggio di Massimo Palermo e Danilo Poggiogalli che arricchisce in modo significativo il già consistente drappello di studi sulle grammatiche concepite per un apprendente straniero (PALERMO-POGGIOGALLI 2011).

Per la seconda sezione, destinata alle grandi aree storico-geografiche, disponiamo di volumi sulla Svizzera (MARTINONI 2010), l'Ungheria (FÁBIÁN-SZABÓ 2010), l'America latina (PATAT 2012), l'America settentrionale (SENO REED-CERCHI 2010), senza contare una monografia di Stanisław Widłak

apparsa già nel 2006 e quindi parallela al progetto di cui si parla (WIDŁAK 2006); e di articoli dedicati alla Romania (D'ACHILLE 2008), all'Albania (DI GIOVINE 2008), a Malta (BRINCAT 2008), all'Islanda (TANI 2010 e 2011), alla Libia e al Corno d'Africa (RICCI 2009). Dichiaratamente concepito nell'ambito di una ricerca realizzata per l'impresa UTET è un lavoro di Gianguido Manzelli (MANZELLI 2008).

Anche per la terza sezione è stata pubblicata la massima parte dei contributi effettivamente consegnati o comunque progettati in quell'occasione. Mi riferisco ai volumi di Matteo Motolese (2012), Stefano Telve (2012), Martino Marazzi (2011), Paola Manni (2008; vedi anche 2010); e agli articoli di Ilaria Bonomi (2010), Fabio Rossi (2008), Lorenzo Tomasini (2010; vedi anche 2011)¹, Rosa Casapullo (2009), Leonardo Rossi e Robert Wank (2010), Massimo Arcangeli (2007).

I saggi previsti nel primo volume ed effettivamente consegnati sono dunque in gran parte confluiti in pubblicazioni a disposizione della comunità scientifica, anche se c'è almeno un'eccezione: il bel saggio di G. Manzelli sulla Turchia, al quale abbiamo fatto cenno sopra, che mi risulta tuttora inedito nella sua interezza. Molto più precaria la sorte dei regesti di italianismi nelle varie lingue del mondo. Sono due i dizionari apparsi sulla scia dell'impresa UTET, GOMEZ GANE 2012 per il catalano e DASHI 2013 per l'albanese; ad essi si aggiungano una serie di contributi apparsi in rivista e/o in rete per il francese (FANTUZZI 2011), il neogreco (FANCIULLO 2008), il portoghese brasiliano (CARAMORI *et al.* 2013), il saho (BANTI-VERGARI 2008), lo swahili (BERTONCINI-TOSCANO 2006).

L'insieme dei lavori apparsi in questi ultimi anni, promossi dall'impresa UTET o autonomamente progettati e svolti²,

¹ Indipendente dall'impresa UTET era un saggio dello stesso studioso allestito per un convegno della Società di Linguistica italiana svoltosi nel 2005: TOMASINI 2006.

² Si aggiungano anche, per documentare in sincronia, la presenza dell'italiano in Europa, nell'area mediterranea e in Argentina, due volumi

permettono prima di tutto qualche riflessione di metodo; di quelle che possono aspirare a una certa utilità perché fondate non su presupposti meramente teorici, bensì su concrete esperienze di lavoro, rese più significative per la varietà dei contesti indagati e dunque per la necessità di affinare e adeguare di volta in volta la strumentazione messa a punto. Ho già svolto alcune considerazioni in questo senso in un contributo di qualche anno fa (SERIANNI 2008); se fornisco solo il nome del rilevatore, in assenza di indicazioni bibliografiche, s'intende che attingo a materiali allestiti per il progetto UTET.

Un problema centrale che si pone preliminarmente, nel momento del censimento dei prestiti, è quello delle fonti lessicografiche, a cui ho già accennato. Si deve solo alla precaria tradizione lessicografica del bengali se in questa lingua, tra le prime cinque del mondo come numero di parlanti, non è possibile individuare, allo stato, italianismi (Mario Prayer), a differenza non solo dei già citati azero e kannada, ma persino dello hausa – un idioma parlato da alcuni milioni di persone soprattutto in Nigeria ed espressione di una cultura quanto mai distante da quella italiana – nel quale Sergio Baldi ha scovato due italianismi mediati dall'arabo parlato (*sábulú* 'sapone' e *bábúr* 'motocicletta', che risale all'ital. *vapore*). Ma esiste anche un rischio speculare, e paradossale, quello di sovrastimare la precocità degli italianismi in una lingua che disponga di strumenti lessicografici particolarmente sofisticati: è il caso, bene illustrato da Ilaria Bonomi, degli italianismi musicali settecenteschi in francese, documentati in una misura «molto maggiore rispetto ad altre lingue prive di una lessicografia musicale così significativa in questo secolo» (BONOMI 2010, p. 188).

Il problema centrale resta però quello della tipologia del prestito. I prestiti diretti, per l'italiano, sono piuttosto l'ec-

pubblicati dalla Società Dante Alighieri a cura di Paolo Peluffo e di chi scrive (PELUFFO-SERIANNI 2005 e 2006). Importante una ricognizione ad ampio raggio sulla lingua delle comunità emigrate ad opera di Vedovelli (2011).

cezione che la regola, almeno per le lingue più distanti dal punto di vista geo-culturale e non interessate da dinamiche demografiche migratorie; nel giapponese, che rappresenta bene questa fattispecie, l'inglese costituiva, in un'indagine del 1964, non solo l'80% del *gairaigo* – cioè del lessico che non è né indigeno né di origine cinese – ma anche l'abituale tramite di ingresso di altri forestierismi (cfr. LOMBARDI VALLAURI 2005, pp. 150-52). Non c'è dubbio che anche i prestiti indiretti debbano rientrare nell'orizzonte dell'OIM; già in DIFIT, del resto, erano stati inclusi senza restrizioni, ricorrendo a una felice soluzione tipografica, l'uso del semigrassetto, per distinguerli dai prestiti diretti. Ma molte volte siamo in presenza, piuttosto, di un'origine multipla, cioè della difficoltà, o dell'impossibilità, di etichettare un prestito come diretto o indiretto. Stammerjohann menziona l'esempio di rum. *bancă*, che può derivare sia dall'ital. *banca* sia dal francese *banque* (cfr. STAMMERJOHANN 2013, p. 136). Qualche volta è dirimente la fonologia: il rum. *muzică* è certamente un prestito diretto dall'italiano, «come si deduce dall'accentazione sulla terzultima sillaba, che esclude l'intermediazione tedesca o francese»³. In altri casi è decisivo il combinato disposto di fonologia e semantica: il turco *borsa* è l'italiano (o il veneziano) *borsa* in accezione economico-finanziaria, mentre *burs* è il franc. *bourse* 'borsa di studio' e lo stesso avviene per *nota* in accezione musicale rispetto a *not* dal franc. *note* 'voto scolastico' (Gianguido Manzelli). Ricorrente è l'incertezza tra italiano e latino, specie per le parole della sfera religiosa (cfr. SERIANNI 2008, p. 24 n. 13 con la bibliografia ivi indicata).

C'è poi il caso dei nomi propri italiani che diventano nomi comuni in altre lingue con accezioni sconosciute nelle lingue d'origine. Ancora Stammerjohann ricorda il caso di

³ D'ACHILLE 2008, p. 104. È stato proprio uno studioso rumeno nel 1950, Alexandru Graur, a introdurre la nozione di "etimologia multipla", come ricorda ancora D'Achille (ivi, p. 95).

Garibaldi passato a indicare una camicetta da donna (1862), una foggia di cappello (1882), un pesce americano (1885), un biscotto (1896) (STAMMERJOHANN 2013, p. 137). Si può convenire con l'illustre romanista sul fatto che «non si tratta [...] di prestiti perché in italiano questi usi non esistono», ma non avrei dubbi né sul loro interesse linguistico, né tantomeno sul loro significato culturale (e il prestito rappresenta per definizione una dimensione in cui la lingua non può prescindere dalla cultura); dunque un osservatorio sull'italiano nel mondo dovrebbe tenerne conto.

Lo stesso vale, a mio avviso, per gli pseudo-italianismi, come il noto *Tutti frutti* (cfr. STAMMERJOHANN 2013, p. 145) o l'irradiazione di pseudo-suffissi italianeggianti che si ha nello spagnolo rioplatense (come *-eli*, tratto evidentemente dai cognomi italiani in *-elli*, affisso a basi schiettamente ispaniche come in *gordeli* da *gordo* 'grasso' o *locateli* da *loco* 'pazzo', cfr. SERIANNI 2008, p. 28). Si tratta di fenomeni che testimoniano l'influenza di una lingua straniera su un'altra in modo addirittura più efficace di quel che molte volte vale per un singolo prestito isolato. Del resto, la debordante invadenza dell'inglese nell'italiano di oggi si ricava anche dalla presenza di pseudo-anglicismi, sicura marca del prestigio angloamericano: alcuni di essi sono ormai stagionati, come *beauty-case*, *footing*, *slip*, *ticket* 'contributo sanitario'; altri, più recenti come i composti ottenuti con un formante inglese e col rispetto della sequenza determinante-determinato (gli autentici *family-day* e *election-day* hanno germinato *risparmio-day* e persino *Fiaba-day* 'giornata di sensibilizzazione sulla necessità di eliminare le barriere architettoniche urbane', dall'acronimo FIABA 'fondo italiano abbattimento barriere architettoniche'), mentre da composti inglesi come *killer bee* 'ape assassina' sono derivati, questa volta con italianizzazione anche della sequenza dei formanti, i nostrani *afa killer*, *infermiera killer*, *onda killer* (cfr. ADAMO-DELLA VALLE 2005, s. vv.). Neanche forme del genere sarebbero degne di figurare in un dizionario di anglicismi (e molte di esse sono oltretutto occasionalismi); eppure per un

quadro della circolazione dell'inglese e della sua modellizzazione della realtà nell'italiano di oggi si tratta di indizi non eludibili.

Ciò non significa beninteso, nemmeno nella prospettiva dell'OIM, perdersi nel pulviscolo dell'effimero, che può ricevere consistenza solo in un contesto più ampio⁴. Hanno pur sempre importanza centrale per misurare il peso di un prestito quattro indicatori: 1. la presenza del lessico che possiamo considerare se non "fondamentale" *stricto sensu*, almeno di uso largo e comunque non settoriale. Naturalmente non è detto che una parola appartenente a questo strato di lessico in italiano e accolta da un'altra lingua non concorra, in questa, con sinonimi indigeni di più largo uso: per esempio, *koinçidoj* 'coincidere' è un italianismo nell'albanese, ma le forme abituali sono *përkoj* e *përputhet*; *folà* e *folla* sono italianismi del catalano, ma le forme abituali per il significato di 'folla' sono altre (*gentada*, *multitud*, *gernació*); 2. la presenza del lessico astratto; 3. la presenza di modi interiettivi e fraseologici; 4. la presenza di elementi extralessicali: affissi e istituti grafici. Esempifico sommariamente le quattro tipologie:

1. Le parole del lessico di largo uso sono ben rappresentate in albanese, ossia nella lingua che, dopo il maltese, risulta quella con maggiore tasso di italianismi: pensiamo a *brun* e *fatorino* che hanno soppiantato le forme indigene *zeshkan* e *biletashitës*, e ancora a *ma-xhorancë* e *minorancë*, *inJORANT* e *impuls* in accezione scientifico-tecnologica (fisica, elettronica, fisiologia,

⁴ All'indomani della riunificazione tedesca, il politologo Gerhard Lehmbuch «paventava il rischio della trasformazione delle regioni orientali in un nuovo *Mezzogiorno*, usando il termine italiano come espressione paradigmatica di un'ampia area geografica separata dal resto del Paese da un sensibile, e immutato, divario» (CARTOCCI 2006, p. 108). Si tratterà di un uso idioletale per giunta proprio di una persona colta; troppo poco per ipotizzare che questa particolare risemantizzazione di un termine geografico circoli anche oltre i confini italiani.

cfr. DASHI 2013, s.vv.). Notevole anche la presenza in questa sezione del catalano, che ha conosciuto un influsso da parte dell'italiano da antica data (già dal XIII secolo; grazie ai contatti commerciali con Genova e Pisa, al dominio diretto di dinastie catalane in Italia, al prestigio dell'umanesimo e del rinascimento italiano): possiamo ricordare *anxova* 'acciuga', *saquejar* 'saccheggiare' e *majòlica*, che è un "cavallo di ritorno" perché l'ital. *maiolica* deriva da un toponimo catalano (cfr. GOMEZ GANE 2012, pp. 99, 66-67, 114-15). Discorso in qualche modo analogo, almeno quanto all'antichità dei contatti (in questo caso mediati soprattutto dai dialetti settentrionali), vale per l'ungherese, per il quale possiamo citare *part* 'riva' (da porto, con slittamento semantico) e *piac* 'mercato' (da piazza; altre accezioni, più vicine alla base italiana, sono oggi obsolete, cfr. FÁBIÁN-SZABÓ 2010, pp. 141 e 40-41).

2. Al lessico astratto possiamo ascrivere forme come alb. *akuzoj* 'incolpare' e *konsideroj* 'considerare' e cat. *important* 'importante' e *grandios* 'grandioso' (cfr. DASHI 2013, s.vv. e GOMEZ GANE 2012, pp. 119 e 154).
3. Moduli interiettivi e fraseologici (comprese le formule di saluto), a parte i diffusissimi *bravo!* e *ciao*, si ritrovano per esempio nello spagnolo rioplatense *un corno* (*no ver / saber un corno*), *fachamo* (;*qué fachamo?* nei forum di discussione in Internet), *ma que⁵* o nello slov. *ma* 'ma' e nell'avverbio *solo*, che ha inoltre sviluppato un uso particolare come affisso (*solopetje* 'canto solistico') (cfr. OŽBOT 2008). Nel portoghese brasiliano si sono avute creazioni autonome: dall'antico prestito *raça* 'razza' si sono sviluppati modi idiomatici propri come *na raça* 'a tutti i costi' e *ter raça* 'essere dotato di qualità posi-

⁵ Informazioni ricavate da Angela Di Tullio e Gabriela Resnik e citate in SERIANNI 2008, p. 34.

tive, forza, bravura, brio' (cfr. CARAMORI *et al.* 2013). Gianguido Manzelli (2008) considera probabile che l'espressione turca *bir bardak suda firtına* 'una tempesta in un bicchier d'acqua' sia modellato sull'italiano (veneziano), anche se l'unica parola esotica è *firtına* 'fortunale' (venez. *fortuna* [*de mar*]) e se, data la sua diffusione nelle lingue europee, l'espressione potrebbe essere poligenetica.

4. Sui prestiti di carattere morfologico è sufficiente rinviare alle pagine di Stammerjohann (2013, pp. 140 segg.), che illustra la vitalità del suffisso *-esco* o del morfema di superlativo *-issimo*⁶ o anche, nel neogreco, della desinenza verbale *-άρω* che s'incontra in formazioni neologiche ed esotiche, in origine «utilizzata per adattare verbi di provenienza italiana come *σαλπάρω* (it. *salpare*) [...] e poi diventata tipica per adattare tutti i verbi di provenienza straniera, anche non italiana». Esempi di influssi sulla compagine grafica sono offerti dal rumeno (grafie fonetiche, non etimologiche, in *filosofie* ecc.; uso del segno diacritico <h> per rappresentare la pronuncia velare di <c, g> davanti a vocale palatale, cfr. STAMMERJOHANN 2013 p. 143, ma anche D'ACHILLE 2008, p. 101).

Chiudo tornando su un punto appena toccato, che ho esemplificato con albanese e catalano⁷: l'importanza dell'uso reale, come può essere testimoniata solo da un parlante madrelingua. Per l'uso contemporaneo fondarsi soltanto sui dizionari monolingui o bilingui è insufficiente: o per difetto (forme realmente circolanti possono non essere state registrate per i consueti fenomeni di inerzia lessicografica) o, molto più

⁶ Cfr. le osservazioni introduttive di Heinz in questo volume.

⁷ Grazie alla consulenza di B. Dashi e Y. Gomez Gane, parlanti madrelingua.

spesso, per eccesso: l'esistenza di un italianismo in una lingua straniera può essere del tutto marginale e marcata in diacronia o in diafasia (arcaismo, forma di registro molto elevato), senza scalfire minimamente il primato della forma indigena.

LUCA SERIANNI

Università "La Sapienza" di Roma

Bibliografia

- ADAMO-DELLA VALLE 2005 = Giovanni Adamo - Valeria Della Valle, *2006 parole nuove. Un dizionario di neologismi dai giornali*, Milano, Sperling & Kupfer.
- ARCANGELI 2007 = Massimo Arcangeli, *Il lessico sportivo e ricreativo italiano nelle quattro grandi lingue europee (con qualche incursione anche altrove)*, in «Studi di lessicografia italiana», XXIV, pp. 195-248.
- BANFI 2012 = Emanuele Banfi, *La posizione dell'italiano (e di "italiani altri") tra le lingue d'Europa: usi, funzione, prestigio*, in T. Telmon - G. Raimondi - L. Revelli (a cura di), *Coesistenza linguistiche nell'Italia pre- e postunitaria*, Atti del XLV Congresso internazionale di studi della Società di linguistica italiana (SLI) (Aosta/Bard/Torino, 26-28 settembre 2011), Roma, Bulzoni, vol. II, pp. 1-124.
- BANTI-VERGARI 2008 = Giorgio Banti - Moreno Vergari, *Italianismi lessicali in saho*, in «Ethnorêma», IV, pp. 67-93.
- BERTONCINI-TOSCANO 2006 = Elena Bertoncini - Maddalena Toscano, *Swahili, Kiswahili*, preparato per il progetto *UTET - Gli italianismi nel mondo*: http://opar.unior.it/189/1/italianismi_swahili.pdf [14/07/2015].
- BONOMI 2010 = Iliaria Bonomi, *La penetrazione degli italianismi musicali in francese, spagnolo, inglese, tedesco*, in «Studi di lessicografia italiana», XXVII, pp. 185-235.
- BRINCAT 2008 = Giuseppe Brincat, *Sicilianismi, italianismi e anglolatinismi nel vocabolario della lingua maltese*, in F. Pierro (a cura di), *Aspects lexicographiques du contact entre les langues dans l'espace roman*, Strasbourg, PU, pp. 199-214.

- BRUGNOLO 2009 = Furio Brugnolo, *La lingua di cui si vanta Amore. Scrittori stranieri in lingua italiana dal Medioevo al Novecento*, Roma, Carocci.
- CARAMORI *et al.* 2013 = Alessandra Paola Caramori *et al.*, *Gli italianismi nel portoghese brasiliano*: <http://www.accademiadel-lacrusca.it/it/scaffali-digitali/articolo/italianismi-portoghese-brasiliano> [14/07/2015].
- CARTOCCI 2006 = Roberto Cartocci, *L'unificazione fallita: il Sud e il capitale sociale*, in «Vita e Pensiero», LXXXIX, 6, pp. 105-13.
- CASAPULLO 2009 = Rosa Casapullo, *Italiano fuori d'Italia: una panoramica sul lessico italiano della scienza nelle lingue europee*, in «Annali» dell'Università Suor Orsola Benincasa, pp. 637-85.
- D'ACHILLE 2008 = Paolo D'Achille, *Dagli Appennini ai Carpazi. I difficili percorsi degli italianismi nel rumeno*, in *Italianismi e percorsi dell'italiano nelle lingue latine*, Atti del Convegno internazionale (Treviso, 22 settembre 2007), Treviso, Fondazione Cassamarca, pp. 93-113.
- DARDANO 2011 = Maurizio Dardano, *La lingua della Nazione*, Roma-Bari, Laterza.
- DASHI 2013 = Brunilda Dashi, *Italianismi nella lingua albanese*, Roma, Edizioni Nuova Cultura.
- DIFIT = Harro Stammerjohann *et al.* (a cura di), *Dizionario di italianismi in francese, inglese, tedesco*, Firenze, Accademia della Crusca, 2008.
- DI GIOVINE 2008 = Paolo Di Giovine, *Un millennio di storia linguistica albanese: l'influsso lessicale della lingua italiana*, in «L'Italia dialettale», 69, pp. 107-40.
- FÁBIÁN-SZABÓ 2010 = Zsuzsanna Fábíán - Győző Szabó, *Dall'Italia all'Ungheria: parole di origine italiana nella lingua ungherese*, Udine, Forum.
- FANCIULLO 2008 = Franco Fanciullo, *Gli italianismi nel neo-greco*, in «L'Italia dialettale», LXIX, pp. 163-203.
- FANTUZZI 2011 = Marco Fantuzzi, *Italianismi nel francese moderno e contemporaneo*, in «Studi di lessicografia italiana», XXVIII, pp. 285-317.
- GOMEZ GANE 2012 = Yorick Gomez Gane, *Gli italianismi nel catalano. Dizionario storico-etimologico*, Roma, Aracne.
- LOMBARDI VALLAURI 2005 = Edoardo Lombardi Vallauri, *Diversità fonologica e convenzioni grafiche: l'adattamento dei prestiti nel*

- Giappone contemporaneo*, in I. Korzen - P. D'Achille (a cura di), *Tipologia linguistica e società*, Firenze, Cesati, pp. 149-76, alle pp. 150-52.
- MANNI 2008 = Paola Manni, *La lingua italiana nel mondo: commercio e finanza*, Firenze, Centro editoriale toscano.
- MANNI 2010 = P. Manni, *Il "De computis et scripturis" e le origini della moderna terminologia economico-finanziaria*, in E. Giusti - M. Martelli (a cura di), *Pacioli 500 anni dopo*, Perugia, Centro studi Mario Pancrazi, pp. 125-37.
- MANZELLI 2008 = Gianguido Manzelli, *Gli italianismi di mediazione turca nel curdo della Turchia e dell'Iraq*, in R. Lazzeroni et al. (a cura di), *Diachronica et synchronica. Studi in onore di Anna Giacalone Ramat*, Pisa, ETS, pp. 301-26.
- MARAZZI 2011 = Martino Marazzi, *A occhi aperti. Letteratura dell'emigrazione e mito americano*, Milano, Angeli.
- MARTINONI 2010 = Renato Martinoni, *L'Italia in Svizzera. Lingua, cultura, viaggi, letteratura*, Venezia, Marsilio.
- MOTOLESE 2012 = Matteo Motolese, *Italiano lingua delle arti. Un'avventura europea (1250-1650)*, Bologna, il Mulino.
- OŽBOT 2008 = Martina Ožbot, *Alcuni cenni sugli italianismi in sloveno*, in "Linguistica" 58 2008, pp. 159-66.
- PALERMO-POGGIAGALLI 2011 = Massimo Palermo - Danilo Poggiogalli, *Grammatiche di italiano per stranieri dal '500 ad oggi. Profilo storico e antologia*, Pisa, Pacini.
- PATAT 2012 = Alejandro Patat (a cura di), *Vida nueva. La lingua e la cultura italiana in America Latina*, Macerata, Quodlibet.
- PELUFFO-SERIANNI 2005 = Paolo Peluffo - Luca Serianni (a cura di), *Il mondo in italiano. Annuario della Società Dante Alighieri 2005*, Roma, SDA.
- PELUFFO-SERIANNI 2006 = P. Peluffo - L. Serianni (a cura di), *Il mondo in italiano. Annuario della Società Dante Alighieri 2006*, Roma, SDA.
- RICCI 2009 = Laura Ricci, *L'italiano in Africa. Lingua e cultura nelle ex colonie*, in «Carte di viaggio», II, pp. 15-46.
- ROSSI 2008 = Fabio Rossi, *Hollywood Italian: l'italiano all'estero ritratto dal cinema statunitense. Rilievi linguistici*, in S. Taviano (a cura di), *Migrazione e identità culturali*, Messina, Mesogea, pp. 107-21.
- ROSSI-WANK 2010 = Leonardo Rossi - Robert Wank, *La diffusione*

- dell'italiano nel mondo attraverso la religione e la Chiesa cattolica: ricerche e nuove prospettive*, in M. Arcangeli (a cura di), *L'italiano nella Chiesa fra passato e presente*, Torino-Londra-Venezia-New York, Allemandi, pp. 113-71.
- SENO REED-CHERCHI 2010 = Cosetta Seno Reed - Paolo Cherchi, *Gli italiani e l'italiano nell'America del Nord*, Ravenna, Longo.
- SERIANNI 2008 = Luca Serianni, *Gli italianismi nelle altre lingue romanze: prime riflessioni*, in *Italianismi e percorsi dell'italiano nelle lingue latine*, Atti del Convegno internazionale (Treviso, 22 settembre 2007), Treviso, Fondazione Cassamarca, pp. 19-41.
- STAMMERJOHANN 2013 = Harro Stammerjohann, *La lingua degli angeli. Italianismo, italianismi e giudizi sulla lingua italiana*, Firenze, Accademia della Crusca.
- TANI 2010 = Maurizio Tani, *Gli italianismi nella lingua islandese*, in «Italiano LinguaDue», 2, pp. 170-95.
- TANI 2011 = M. Tani, *Per una storia dei rapporti culturali e artistici tra Italia e Islanda*, in «Rivista dell'Istituto di Storia dell'Europa Mediterranea», 7, dicembre 2011, pp. 45-82 (disponibile anche in rete: http://rime.to.cnr.it/2012/RIVISTA/N7/2011/articoli/Maurizio_Tani.pdf [14/07/2015]).
- TELVE 2012 = Stefano Telve, *That's amore. La lingua italiana nella musica leggera straniera*, Bologna, il Mulino.
- TOMASIN 2006 = Lorenzo Tomasin, *Gli italianismi marinareschi nelle "lingue esotiche": problemi ricostruttivi e fonti documentarie*, in E. Banfi - G. Iannàccaro (a cura di), *Lo spazio linguistico italiano e le "lingue esotiche": rapporti e reciproci influssi*, Roma, Bulzoni, pp. 85-96.
- TOMASIN 2010 = L. Tomasin, *Sulla diffusione del lessico marinaresco italiano*, in «Studi linguistici italiani», XXXVI, pp. 161-90.
- TOMASIN 2011 = L. Tomasin, *La lingua della marineria*, in R. Simone (a cura di), *Enciclopedia dell'italiano*, Roma, Istituto dell'Enciclopedia Italiana, pp. 856-58.
- VEDOVELLI 2011 = Massimo Vedovelli (a cura di), *Storia linguistica dell'emigrazione italiana nel mondo*, Roma, Carocci.
- WIDŁAK 2006 = Stanisław Widłak, *Italia e Polonia. Popoli e lingue in contatto*, Kraków, Wydawnictwo Uniwersytetu Jagiellońskiego.

THE LIMITS OF LEXICOGRAPHICAL ABSTRACTION.
SOME STRENGTHS AND PROBLEMS
OF THE DATA ARCHITECTURE
IN THE LEHNWORTPORTAL DEUTSCH

1. INTRODUCTION

The idea of a database-driven ‘inverted’ loanword dictionary (ENGELBERG 2010) that documents borrowings from a given donor language into many recipient languages goes back to at least the 1970s (cf. KARAULOV 1979). Apart from technical considerations, the major obstacle to such an enterprise for any major language is the large number of recipient languages to be considered. So far, the only comprehensive lexicographical resource of this kind available is for Dutch (VAN DER SIJS 2010). The underlying database for this monumental work has been compiled in a huge effort by the author; the data are available through the print publication and as a PDF file of the book, but to date there is no digitized search interface to the database. Apart from the *Osservatorio degli italianismi nel mondo* that is currently in development (OIM 2014, HEINZ-GÄRTIG 2014), the only project of a somewhat similar intended scope seems to be the *Lehnwortportal Deutsch* (<http://lwp.ids-mannheim.de>; cf. MEYER-ENGELBERG 2011) that aims to provide free access to a growing number of dictionaries of German loanwords in other languages. The web portal is maintained and hosted by the Institute for the German language (IDS) and is online in a preliminary version since late 2012. In a first stage of this project, dictionaries on German loans in three Slavic languages have been integrated; an ongoing research project develops and integrates

dictionaries on Polish-mediated German loans in the East Slavic languages Russian, Belorussian and Ukrainian (MEYER 2014a). An increasing number of loanword resources on Germanisms in other language families is currently being acquired or prepared for inclusion.

The overall structure and basic conception of the *Lehnwortportal Deutsch* (henceforth, LWP) has been described in several publications (cf. MEYER 2013a, MEYER 2014b) and will not be reviewed here in depth. Apart from traditional lemma-based access to the individual resources through a uniform web user interface, the web portal features an ‘inverted’ dictionary whose headwords (henceforth, *metalemmata*) are German etyma; the entries of this dictionary mainly consist of pointers to entries in the various loanword dictionaries of the portal, thus allowing users to trace the way of a specific German word into the different recipient languages. The inverted dictionary is not simply the product of an automated rearrangement of the digital or digitized data presented in the loanword dictionaries. The German etyma as they actually appear in the loanword dictionary entries – often in some dialectal and/or diachronic form – are mapped in a manual, lexicographically informed process onto ‘normalized’ *metalemmata*, typically a diasystematically corresponding form in contemporary standard New High German, using a dedicated in-house software tool.

The major innovation of the LWP is the data organization metaphor underlying the various cross-resource search options (cf. MEYER 2013b) offered for advanced purposes. Both the portal-specific *metalemmata* and the actual object-language words (etyma, loanwords including morphophonological variants, derivatives etc. listed in the entries – not only headwords and not necessarily lexemes) as they are presented or ‘recorded’ in the entries of the individual loanword dictionaries are represented as the vertices of a directed acyclic graph (DAG) whose edges represent various relations between words, such as “etymon x is mapped to *metalemma* y”, “loanword x is

2. A SKETCH OF THE PORTAL ARCHITECTURE

The LWP data model has been designed with a number of rather specific requirements in mind. A major objective of the portal was to make a growing range of existing dictionaries of German loanwords in other languages available under a single roof and with a uniform access structure. Apart from dedicated third-party funded projects, the IDS Mannheim has no resources for compiling such dictionaries on its own, so it is only to be expected that the legacy dictionaries to be included in the portal come in a variety of original data formats (XML, plaintext, image-digitized, ...), with very different entry structure and with incompatible formats of disystemic, grammatical and other information. It is important to note that the resources are taken 'as is', although there are plans to add a versioning system to the portal that will allow designated lexicographers to make corrections and additions.

From the outset, the LWP was meant to provide more than a simple compilation or mashup of lexicographical resources. Experts should be able to perform detailed, powerful and fast searches in a steadily growing lexicographical portal database that makes systematic use of *ex post facto* interlinking of legacy resources on the level of individual words, not just entries or documents as a whole. Updating and augmenting entries should be possible and be handled separately from changes in portal-specific cross-referencing information.

To fulfil these requirements, two different representations of the lexicographical data are used side by side. Each entry in one of the loanword dictionaries is a standalone XML document; each loanword dictionary retains its own XML schema. No attempt has been made to create a kind of super-schema that purports to fit all microstructural needs of the individual dictionaries. Even the merely three dictionaries included so far in the LWP display remarkable idiosyncrasies that would have been impossible to foresee in the design process for such a super-schema. Trying to find an all-encompassing XML

schema for all dictionaries would be a hard task and force constant readjustment. Just two examples of particularities: (a) In the Standard Polish loanword dictionary of the LWP (DE VINCENZ-HENTSCHEL 2010), derivatives of a given Polish lemma are often annotated with those word senses of the base lexeme that are semantically ‘compatible’ with the meaning of the derivative. (b) The dictionary of German loanwords in the Cieszyn dialect of Polish (MENZEL-HENTSCHEL 2005) systematically specifies etymologically related *lexical parallels* to the lemma in other languages of the same linguistic area (Polish, Czech, German; both standard languages and dialects), irrespective of the exact and often unclear borrowing history (i.e., which language borrowed from which).

Schema-specific XSL transformations are used to produce HTML views of the entries that are at least close to the original entry presentations, so that, apart from minor design changes, the dictionaries may be consulted in their original appearance.

The XML documents are integrated into the portal database without significant alterations; however, all XML elements corresponding to recorded words are provided with an additional XML attribute, viz. a globally unique ID (UUID), for cross-referencing purposes. The list of metalemmata and different sorts of cross-referencing added for the LWP are coded in separate XML documents that make heavy use of the UUIDs just mentioned. These XML documents are not edited manually, but through a complex in-house GUI application that takes care of keeping the link consistency intact.

In principle, the cross-linked XML documents as such could be used for portal-wide search operations. But searching XML documents with varying schemas and recursively traversing ever longer chains of cross-references between XML documents would soon become too slow as the portal grows larger.

As *relations between recorded words* are the greatest common factor both of different microstructures in loanword dictionaries and of cross-referencing XML documents, a natural

solution to this and other technical problems is to represent the portal's lexicographical content as a network, a *directed acyclic graph*, of relations between the recorded words. Such a graph can be stored and accessed in a scalable and efficient manner in a dedicated NoSQL graph database such as Neo4j (cf. WEBBER-EIFREM-ROBINSON 2013). Currently, the LWP emulates a graph database through a conventional relational database. The graph is derived in a fully automated, dictionary-specific process from the underlying XML resources mentioned above: Structural configurations between XML elements – expressible by relative XPath expressions – are translated into graph edges; word-specific information (di-systemic, grammatical, and other data on recorded words) is converted to node attributes in a unified data format. The graph as such is not an independent resource and must be recreated or adjusted every time an underlying XML resource changes; thus, no complex and error-prone synchronisation between XML documents and the graph structure is needed.

3. STRENGTHS

With its inter-dictionary cross-references and the uniform layout and access structure of the portal and its component dictionaries, the LWP is a standard example of a dictionary net in the sense of ENGELBERG-MÜLLER-SPITZER 2013. Moreover, its graph-based data representation adds a layer of domain-specific *uniform access semantics* for all portal resources, a layer that is created through a process of abstracting from the idiosyncrasies of individual dictionary macro- and microstructures and reframing the lexicographical data in a common graph-based metalanguage that enables users to search for information without having to consider the structural diversity of the various lexicographical sources.

- In comparison to XML-based solutions, the data architecture is highly *scalable* and can handle complex cross-resource searches for a growing number of com-

ponent dictionaries without considerable performance penalty.

- The architecture is *flexible* and accommodates to a wide range of entry microstructures in legacy dictionaries instead of imposing a single lexicographical data model.
- The duality of arbitrarily many interlinked XML documents on the one side and a graph structure automatically extracted from these documents on the other allows the portal-makers to *cleanly separate* dictionary-specific content from portal-added information (by coding different content in different XML documents) without sacrificing a uniform cross-resource view on the data where the boundaries between the different sources of information are virtually invisible.
- The portal-wide part of the data model is *granular*: Since the graph nodes represent individual recorded words, not whole entries, cross-references can and must be formulated at the word level.
- The graph is *extendable*; arbitrarily many additional attributes can be assigned to both nodes (words) and edges (relations). These attributes may contain user-generated plaintext comments; versioning information in the case of updates to the original entries; information on the degree of plausibility of, say, a borrowing relationship as indicated by the source; semantic or ontological classification as a complement to a traditional lexicographical meaning definition; etc. There are no restrictions on how this additional information is encoded and represented in, e.g., separate XML source documents, as long as the graph can automatically be constructed using all available documents. This paves the way for innovative bookkeeping techniques as far as versioning and user-generated additions to the resources are concerned, since all versions of the lexicographical information could easily be co-present in one and the same graph, as long as all nodes and edges are

systematically provided with versioning information.

- The graph structure remains semantically *consistent* even when different lexicographical sources give contradictory information on a particular word, say, specify different gender for a German etymon noun. The simple reason is that words recorded in different resources are, as a matter of principle, mapped onto different vertices of the graph. Any statements on cross-resource identities are added content that might be coded as an identity-expressing edge by the portal staff. German etyma are a special case; their diasystematic identity is coded by mapping them onto the same metalemma. Metalemmata serve as simple pointers to dictionary-specific etyma and do not possess attributes of their own, so no problems with conflicting information arise. Search requests may easily return inconsistent data from different resources, so that an informed evaluation and decision will be left to the user.
- The DAG has a *simple* conceptual structure that can be mapped easily onto description standards such as TEI and onto semantic web standards such as RDF (every graph edge is indeed basically isomorphic to an RDF triple) to ensure interoperability with other applications and the possibility of cross-linking with other data sets; cf. WANDL-VOGT-DECLERCK 2013 for an example of publishing lexicographical resources in the Linked Open Data framework.
- The graph structure is well suited for *visualisation* and is already being used for illustrative purposes in the LWP, as connected subgraphs can be displayed interactively for each German metalemma. However, the simple visual metaphor of graphs also lends itself to exploitation for novel and intuitive kinds of searches, similar to, e.g., the interactive construction of graphs representing search queries in treebanks as described in VOORMANN-LEZIUS 2002.

4. PROBLEMS

Somewhat paradoxically, most of the conceptual, lexicographical, and practical difficulties that the LWP faces ensue from features that have been described above as advantages of the basic conception of the portal.

A major source of trouble that will quickly become obvious even to occasional users of the portal is the *heterogeneity* of the resources included, not only as far as reliability, quality and volume of the provided information is concerned.

- Loanword dictionaries vary in the *criteria of inclusion/exclusion* of etyma, loan translations etc. The two Polish-related dictionaries in the LWP (MENZEL-HENTSCHEL 2005, DE VINCENZ-HENTSCHEL 2010) do not include loanwords whose German etyma are of Latin or Greek origin; many yet-to-be-included resources beg to differ on this point. If a user queries the portal for languages that have borrowed such etyma, she or he will probably get a thoroughly misleading picture that largely depends on the particular decisions of the various dictionary editors. There is, of course, no proper general solution for this in a context where legacy resources have to be taken “as is” and might, in addition, differ considerable as to their age and intended comprehensiveness.
- Both systematic and occasional *lacunae* in the lexicographical data pose another major problem. One of the currently three component dictionaries of the LWP (MENZEL-HENTSCHEL 2005) does not indicate part of speech or any other grammatical information on etyma and loans. This makes the design of advanced searches difficult: At present, the mere choice of including part of speech or gender as a criterion in a portal-wide search implies that no results from MENZEL-HENTSCHEL 2005 will be displayed. Two possible workarounds come to mind, neither of which is

entirely satisfactory: The user might add a disjunctive criterion (“POS is either noun *or unspecified*”; such an option is not implemented yet) or such a disjunctive criterion is added implicitly by the search engine. Adopting the latter solution would only be sensible if search results with unspecified POS are somehow marked as not completely conforming to the criteria specified.

An example for an occasional type of lacuna would be missing information on the meaning of recipient-language derivatives from loanwords in cases where well-known productive derivational patterns are involved. From the point of view of print lexicography, this is a reasonable way of saving space and avoiding redundancy, but it is a problem for anyone using meaning paraphrases as a search criterion in the portal. A possible solution for this particular problem could be to simply copy the meaning(s) associated with the base of the derivative and flag this via an additional attribute of the derivative. Such artefacts can quickly make the search system unwieldy for end users, however.

Sometimes textual condensation routinely used in print lexicography produces lacunae in a digitized version, for example in cases where a previous word sense definition is referred to later in the entry using words such as ‘idem’ or its German equivalent ‘dss.’ Usually, the exact reference cannot be established algorithmically, such that the only way out is to fill in the full definition manually.

- As briefly explained above, the German *metalemmata* constitute the backbone for cross-resource links in the graph. Operationalizing the process of individuating and assigning *metalemmata* as well as defining their interrelations has turned out to be very difficult, mainly because it is often unclear whether it is appropriate to ‘identify’ two German etyma given in different re-

sources. It is clear that the main criteria for this must rely on linguistic reconstruction in historical linguistics, but often details remain unclear or seem to be irrelevant to the majority of portal users. To give an example, the Standard Polish dictionary in the LWP (DE VINCENZ-HENTSCHEL 2010) gives the following New High German etyma for Polish *sztuciec* ‘case for storage’ (<http://lwp.ids-mannheim.de/art/wdlp/lemma/sztuciec>): *Stutz*, *Stutzen*, *Stütze* ‘stump, vessel’. Of these three forms, only *Stutzen* remains in active use in Standard German (though there is an unrelated homonym *Stütze* ‘support’), and in addition it is difficult to pinpoint the precise morphological and historical-phonological relationships between these variants. Should all three forms or just *Stutzen* be included as metalemmata? Should dialectal variants that come to be integrated into the standard language in specialised meanings be represented as separate metalemmata, or, more generally, how do we determine the maximal admissible ‘diasystematic distance’ between variant forms mapped onto the same metalemma? Comparable problems are of course faced by all projects that try to integrate heterogeneous lexicographical resources into a dictionary net through some concept of meta-index, cf. HEROLD-GEYKEN-LEMNITZER 2012.

Moulding entry data into the graph structure of the LWP is a complex process of interpretation of both explicit and implicit lexicographical information. Once again, users must be aware that the advanced search options operate on a formalized, artificially homogenized view of the original data.

- Most obvious is the case of *dictionaries without a systematic microstructure* whose entries are written in a rather narrative or even discursive style, as in the Slovene dictionary of the LWP (STRIEDTER-TEMPS 1963) that has been included into the LWP through image-digit-

zation of the entries. For the construction of the graph, additional manual extraction of lexicographical information into a fixed XML structure was necessary. This process was difficult and error-prone as many details, especially regarding the exact delimitation of different entry sections, determination of the scope of qualifiers, and the interpretation of implicit references within the entry required careful human inspection and often enough remained unclear even to trained philologists. It goes without saying that such issues are intrinsic to any attempt at retrodigitizing unstructured text, but they are much more apparent in a setting where each piece of information must be assigned to precisely one word (vertex) or word-word-relation (edge).

- As stated above, the *set of attributes and their values is uniform* across the portal, which entails the mapping of all dictionary-specific grammatical and diasystematic classifications and categories onto portal-wide ‘ontologies’. Thus, although denominations for dialects and older forms of German such as ‘New High German’ (NHG), associated with slightly differing, yet not always explicitly specified, temporal and areal boundaries by different authors, must be equated for the portal’s purposes. Users searching for loans from NHG etyma will therefore not get words borrowed in a specific period in the history of High German, but words borrowed from *whatever was classified as a ‘New High German’ etymon* in the various source dictionaries, including dialectal specifications such as ‘Bavarian-Austrian’ that, strictly speaking, leave the temporal component unspecified.
- Relations between words are often *left unspecified* in the source dictionaries when they are supposed to be obvious to the intended audience. A good example is the entry on the Polish noun *kształt* ‘shape etc.’ in DE VINCENZ-HENTSCHEL 2010 (<http://lwp.ids-mann->

heim.de/art/wdlp/lemma/ksztalt) that notes two variant forms of the word, viz. *ksztalt* and *kstalt*. There is a productive adjective in *-owny* derived from the loanword, also listed in two variant forms, *ksztaltowny* and *kstaltowny*. The precise relations, namely that *ksztaltowny* is derived from *ksztalt* and *kstaltowny* from *kstalt*, are left unspecified, i.e. they are not recoverable from the XML markup of the entry. This enforces a rather contrived algorithmic translation from XML to DAG, e.g. by encoding both adjectives as derived from both nouns, or (this is the current solution in the LWP) encoding both adjectives as derived from the ‘canonical’ lemma variant. In every conceivable solution to this problem, the interpretation of the relation (edge type) “x is a derivative of y” becomes blurred and opaque. – A similar case where information was left unspecified because it was unknown to the dictionary compiler can be found in the Slovene dictionary (STRIEDTER-TEMPS 1963): Often enough, it is not clear whether a word is a Slovene derivative of a loanword borrowed from a German etymon *e* or a Slovene loanword borrowed from a German derivative of *e*, cf. Slovene *drukati/druk/drukar* ‘to print/(a) print/printer’ vs. German *drucken/Druck/Drucker* (<http://lwp.ids-mannheim.de/art/st/lemma/drukati>). The two different borrowing pathways – which may even occur simultaneously, with Slovene speakers alternatively borrowing the German derivative or forming a derivative on their own according to a productive Slovene pattern – necessitate a rather crude operationalization. In this specific case, a borrowed derivative has been assumed whenever the German derivative is mentioned in the entry.

As a general principle, the direction of graph edges in the LWP is intended to convey a uniform semantics: In a relation $x \rightarrow y$ the target node word y may to be thought of

as being ‘dependent on’ and usually temporally subsequent to the source word node *x*. The following relation types have been implemented so far in the LWP data model: metalemma → German etymon; German etymon → loanword in target language; lexeme → derivative lexeme; lexeme → compound lexeme; word → variant form (orthographical, phonological variant); lexeme → lexical parallel (see above for explanation). The relation type ‘metalemma → German etymon’ is special as it does not literally indicate a temporal sequence, even though it might be read as typically implying a certain linguistic provenance that is, however, not made explicit in the metalemma, because the latter is, as stated above, a commonly used Standard NHG lexeme, wherever possible. The main advantage of this approach, apart from technical considerations, is that this semantics transitively extends to arbitrarily long chains (‘paths’) of connected unidirectional edges as they arise, inter alia, with borrowing chains. In some cases, however, this semantics is compromised when the edge direction is either unclear (there is not enough evidence nor any suitable operationalization to decide on the ‘correct’ direction) or irrelevant (semantically, there is no directedness). Such a situation will lead to search results that are misleading at best because the graph structure seems to be erroneous:

- The *lexical parallels* adduced in MENZEL-HENTSCHEL 2005 leave the actual borrowing pathway(s) unspecified, as stated above, mostly because in many cases they cannot be reconstructed with certainty on the basis of the evidence available. Hypotheses on plausible borrowing histories are left to a free text commentary. To take an example, Cieszyn Polish (*h*)*ajbisz* has German *Eibisch* ‘marshmallow (the plant, *Althaea officinalis*)’ as its etymon and Czech *ibišek*, *ajbiš* as lexical parallels. The DAG subgraph for this entry (<http://lwp.ids-mannheim.de/art/meta/lemma/Eibisch>) has edges such as “German etymon *Eibisch* → Cieszyn Polish loanword *ajbisz*” and “Cieszyn Polish loanword *ajbisz* → Czech

parallel form *ajbiś*”, suggesting a borrowing chain *Eibisch* → *ajbisz* → *ajbiś*. The free text commentary reveals that the Cieszyn Polish form has “obviously” been borrowed from Czech, so the chain should rather be *Eibisch* → *ajbiś* → *ajbisz*. This is a simple and clear case; for a really complicated example see <http://lwp.ids-mannheim.de/art/wdlt/lemma/bachórz>. Things become even more difficult where multiple borrowings, possibly along different pathways and concerning only certain derivatives or even word senses of derivatives, are concerned. In all of these cases, the simple graph metaphor cannot but fail on its semantic promise, all the more since the ‘correct’ relations cannot be reconstructed from the digital source because they are specified, if at all, in the commentary. The only clean way of dealing with this problem is to exclude lexical parallels from normal portal-wide search queries.

- The relation of being an *orthographical or phonological variant* is, generally speaking, a symmetric one and therefore clashes with the directedness requirement of the DAG. In a group of multiple variants, the only way of connecting each node to every other one in the group would be to introduce arbitrary edge directions that avoid cycles. For the time being, the LWP always uses an operationalization of electing one member of the group as the ‘canonical’ variant and introducing only variance-edges from this node to all other group members. These ‘non-canonical’ variants are not linked to any other nodes; any other edges linking the ‘canonical’ variant to other nodes outside the group must then be interpreted as potentially also or only ‘applicable’ to all or some of the ‘non-canonical’ variants. For example, in the case of Polish *kształt* vs. *kształt* and *kształtowny* vs. *kształtowny* discussed above, the only edges introduced are: *kształt* → variant *kształt*; *kształt* → derivative *kształtowny*; *kształt* → derivative *kształtowny*. That the

last edge mentioned is really only ‘applicable’ to the ‘non-canonical’ variant *kstatt* remains unexpressed and would not be deducible from the source data anyway. One remedy in these cases would be to form a non-directed, strictly local ‘hyperedge’ that spans all variants and combines them to a sort of ‘hypernode’ that all other edges attach to. Such an approach is possible with modern graph databases.

- A further problem arises with the *directedness of morphological derivation*. Often enough, a decision on the ‘correct’ direction is notoriously difficult and theory-dependent, for example with conversion (zero derivation). This is particularly relevant for establishing the direction of derivational edges between metalemmata, as the source loanword dictionaries will usually define a direction for derivation relations within the recipient language.

Sometimes, the fundamental decision to take recorded words as graph nodes and, therefore, as basic building blocks of the data model, leads to a representation that is, surprisingly, not granular enough and makes it impossible to code certain relations adequately. The most important case is that of ‘multiple borrowing’ where “one and the same etymon” has been borrowed at different times or, more generally, in different diasystematically related forms, into a target language. This leads to (a) word senses of a loanword that are attributable to only one of these forms and (b) to derivative formations that also relate to only one of these forms. By way of example, the LWP’s Polish dictionary DE VINCENZ-HENTSCHEL 2010 features an entry on Standard Polish *waga* ‘scales’ where the NHG and the Middle High German form of the etymon are separately associated with some of the 24(!) word senses of the loanword, albeit in a non-exhaustive and overlapping fashion. The same kind of association with word senses is also coded for the derivatives of *waga*. At least in principle,

this can lead to a state where a certain loanword derivative is ‘compatible’ with only one or at least not with all given forms of the etymon. This is not the case with *waga*, where all derivatives are associated with at least one word sense that is itself connected to both etyma. In another case, however, the entry on Polish *lump* ‘wretch’ (<http://lwp.ids-mannheim.de/art/wdpl/910>) lists two related German etyma, viz. *Lump* ‘wretch’ and the compound noun *Lumpenzucker* ‘bad sugar’; all derivatives only relate to a word sense that itself is associated only with *Lump*, not with *Lumpenzucker*. In this case, a path in the graph such as *Lumpenzucker* → borrowed as *lump* → derivative adjective *lumpowski* leads to the wrong conclusion, viz. that of *lumpowski* being indirectly related, amongst other things, to the etymon *Lumpenzucker*. Note that splitting loanwords into homonyms is not an option for DE VINCENZ-HENTSCHEL 2010 since word senses, as we saw in the case of *waga*, can often be linked to multiple etyma.

5. CONCLUSION

Interlinking a growing number of digital resources has become a major topic of research in the digital humanities. The present paper has demonstrated some of the conceptual, philosophical, and practical difficulties that arise even in the context of manually cross-referencing resources from a rather narrow content domain and providing a common access structure for them. As a matter of fact, the main strength of the LWP and its main problems come by the same name, viz. abstraction.

For most of the issues sketched in this paper, there is no ready-made solution. Three strategies come to mind that can help to live with what may be perceived as a thoroughly unsatisfactory situation. First, user-generated content and the possibility for experts to provide additions and corrections can help to overcome some of the graph’s shortcomings, if only in the long run. Second, users must be made aware at every turn of the pitfalls of taking search results at face value. The

graph-based abstraction used in the LWP is ‘leaky’ in a sense well-known to computer scientists: in many cases, thorough knowledge of the abstraction process and a careful inspection of the resources underlying the abstraction is necessary to interpret query results adequately. And third, the search engine should take a maximally generic approach unless otherwise specified – it is better to have too many results (and let the human user sort out the false positives) than to use restrictive search strategies that possibly sort out relevant hits. As far as the LWP is concerned, the strategy adopted for the advanced HTML form based search is roughly as follows: the user may specify search criteria for a German etymon and a loanword related to this etymon. Now the search engine does not search for matching etymon-loanword pairs that are connected through a path in the graph but instead looks for German metalemmata that are connected to both the etymon and the loanword – the simple rationale for this setup being that every node in the graph is linked to at least one metalemma. This approach neutralizes the problems with the semantics of graph relations discussed above: for example, not every variant of an etymon is connected to every variant of a loanword that is etymologically ‘connected’ to it; but all variants of the loanword and all variants of the etymon are connected to the same German metalemma, because otherwise there would be no etymological ‘connection’. In many cases, this strategy yields too many hits, but this is preferable over a situation where careful consideration of the DAG-construction algorithm would be needed in order to formulate a maximally exhaustive search query.

PETER MEYER

Institut für Deutsche Sprache, Mannheim

References

- DE VINCENZ-HENTSCHEL 2010 = Andrzej de Vincenz - Gerd Hentschel, *Wörterbuch der deutschen Lehnwörter in der polnischen Schrift- und Standardsprache. Von den Anfängen des polnischen Schrifttums bis in die Mitte des 20. Jahrhunderts*, Oldenburg, BIS-Verlag ("Studia slavica Oldenburgensia", 20) (<http://diglib.bis.uni-oldenburg.de/bis-verlag/wdlp> [14/07/2015]).
- ENGELBERG 2010 = Stefan Engelberg, *An inverted loanword dictionary of German loanwords in the languages of the South Pacific*, in A. Dykstra - T. Schoonheim (eds.), *Proceedings of the XIV Euralex International Congress*, Leeuwarden, 6-10 July 2010, Leeuwarden, Fryske Akademy, pp. 639-47.
- ENGELBERG-MÜLLER-SPITZER 2013 = Stefan Engelberg - Carolin Müller-Spitzer, *Dictionary Portals*, in R.H. Gouws et al. (eds.), *Wörterbücher | Dictionaries | Dictionnaires. Ein internationales Handbuch zur Lexikographie | An International Encyclopedia of Lexicography | Encyclopédie internationale de lexicographie*, Supplementary Volume: *Recent Developments with Focus on Electronic and Computational Lexicography*, Berlin-Boston, de Gruyter, pp. 1023-35.
- HEROLD-GEYKEN-LEMNITZER 2012 = Axel Herold - Alexander Geyken - Lothar Lemnitzer, *Integrating lexical resources through an aligned lemma list*, in Ch. Chiarcos - S. Nordhoff - S. Hellmann (eds.), *Linked data in linguistics*, Berlin/Heidelberg, Springer, pp. 35-44.
- HEINZ-GÄRTIG 2014 = Matthias Heinz - Anne-Kathrin Gärtig, *What a multilingual loanword dictionary can be used for: Searching the Dizionario di italianismi in francese, inglese, tedesco (DIFIT)*, in A. Abel - C. Vettori - N. Ralli (eds.), *Proceedings of the XVI EURALEX International Congress. The User in Focus*, 15-19 July 2014, Bolzano/Bozen, Bolzano, EURAC research, pp. 1099-107 (http://www.euralex.org/elx_proceedings/Euralex2014/euralex_2014_085_p_1099.pdf [14/07/2015]).
- KARAULOV 1979 = Jurij Nikolaevič Karaulov, *Obratnyj slovar' zaimstvovaniĭ kak sposob isučeniĭa lingvoëkologii*, in «Izvestija Akademii Nauk SSSR», *Seriĭa Literatury i Jazyka*, 38/6, pp. 552-62.
- MENZEL-HENTSCHEL 2005 = Thomas Menzel - Gerd Hentschel,

- Wörterbuch der deutschen Lehnwörter im Teschener Dialekt des Polnischen*, 2., ergänzte und korrigierte elektronische Ausgabe (http://www.bkge.de/Publikationen/Online/Woerterbuecher/Deutsche_Lehnwoerter_im_Teschener_Dialekt [14/07/2015]).
- MEYER 2014a = *Graph-based representation of borrowing chains in a web portal for loanword dictionaries*, in A. Abel - C. Vettori - N. Ralli (eds.), *Proceedings of the XVI EURALEX International Congress. The User in Focus*, 15-19 July 2014, Bolzano/Bozen, Bolzano, EURAC research, pp. 1135-45 (http://www.euralex.org/elx_proceedings/Euralex2014/euralex_2014_088_p_1135.pdf [14/07/2015]).
- MEYER 2014b = Peter Meyer, *Von XML zum DAG. Der lexikographische Prozess bei der Erstellung eines graphenbasierten Wörterbuchportals*, in M.J. Domínguez Vázquez - F. Mollica - M. Nied Curcio (eds.), *Zweisprachige Lexikographie zwischen Translation und Didaktik*, Berlin/Boston, de Gruyter, pp. 303-21.
- MEYER 2013a = Peter Meyer, *Ein Internetportal für deutsche Lehnwörter in slavischen Sprachen. Zugriffsstrukturen und Datenrepräsentation*, in S. Kempgen et al. (eds.), *Deutsche Beiträge zum 15. Internationalen Slavistenkongress*, Minsk 2013, München, Sagner ("Die Welt der Slaven", Sammelbände, Band 50), pp. 233-42.
- MEYER 2013b = Peter Meyer, *Advanced graph-based searches in an Internet dictionary portal*, in I. Kosem et al. (eds.), *Electronic lexicography in the 21st century: thinking outside the paper. Proceedings of the eLex 2013 conference*, 17-19 October 2013, Tallinn, Estonia, Ljubljana-Tallinn, Trojina, Institute for Applied Slovene Studies - Eesti Keele Instituut, pp. 488-502 (http://eki.ee/elex2013/proceedings/eLex_2013_34_Meyer.pdf [14/07/2015]).
- MEYER-ENGELBERG 2011 = Peter Meyer - Stefan Engelberg, *Ein umgekehrtes Lehnwörterbuch als Internetportal und elektronische Ressource: Lexikographische und technische Grundlagen*, in H. Hedeland - T. Schmidt - K. Wörner (eds.), *Multilingual Resources and Multilingual Applications* (= "Arbeiten zur Mehrsprachigkeit/Working Papers in Multilingualism", Folge B, Nr. 96), Hamburg, Universität Hamburg, pp. 169-74.
- OIM = *Osservatorio degli italianismi nel mondo*, progetto dell'Accademia della Crusca, coord. Luca Serianni - Matthias Heinz,

- con la collaborazione di Marco Biffi - Domenico De Martino - Nicoletta Maraschio - Giovanni Salucci - Gesine Seymer - Harro Stammerjohann et al. [Firenze, Accademia della Crusca, 2014-]: www.italianismi.org [30/11/2015].
- SPOHR 2012 = Dennis Spohr, *Towards a Multifunctional Lexical Resource. Design and Implementation of a Graph-based Lexicon Model*, Berlin-Boston, de Gruyter, ("Lexicographica", Series Maior, 141).
- STRIEDTER-TEMPS 1963 = Hildegard Striedter-Temps, *Deutsche Lehnwörter im Slovenischen*, Wiesbaden, Harrassowitz.
- VAN DER SIJS 2010 = Nicoline van der Sijs, *Nederlandse woorden wereldwijd*, Den Haag, SDU Uitgever (https://pure.knaw.nl/portal/files/458170/Nww_compleet_archief.pdf [14/07/2015]).
- VOORMANN-LEZIUS 2002 = Holger Voormann - Wolfgang Lezius, *TIGERin - Grafische Eingabe von Benutzeranfragen für ein Baumbank-Anfragewerkzeug*, in S. Busemann (ed.), *KONVENS 2002*. 6. Konferenz zur Verarbeitung natürlicher Sprache. Proceedings. DFKI Document D-02-01, Saarbrücken, Deutsches Forschungszentrum für Künstliche Intelligenz (<http://konvens2002.dfki.de/cd/pdf/04P-Voormann-Lezius.pdf> [14/07/2015]).
- WANDL-VOGT-DECLERCK 2013 = Eveline Wandl-Vogt - Thierry Declerck, *Mapping a Traditional Dialectal Dictionary with Linked Open Data*, in I. Kosem et al. (eds.), *Electronic lexicography in the 21st century: thinking outside the paper, Proceedings of the eLex 2013 conference*, 17-19 October 2013, Tallinn, Estonia, Ljubljana-Tallinn, Trojina, Institute for Applied Slovene Studies - Eesti Keele Instituut, pp. 460-71 (http://eki.ee/elex2013/proceedings/eLex2013_32_Wandl-Vogt+Declerck.pdf [14/07/2015]).
- WEBBER-EIFREM-ROBINSON 2013 = Jim Webber - Emil Eifrem - Ian Robinson, *Graph Databases*, Sebastopol, CA, O'Reilly.



GLI ITALIANISMI NELLA LINGUA MALTESE: NON SOLO PRESTITI

1. INTRODUZIONE

La lessicografia maltese inizia con un nome illustre, quello di Hieronymus Megiser, pioniere del collezionismo linguistico e del comparatismo. Desideroso di vedere con i propri occhi la piccola isola che aveva resistito per quattro lunghi mesi nell'estate del 1565 all'assedio ottomano, vi si recò nel 1588. L'assedio attirò l'attenzione di molti intellettuali europei, prima di tutto perché fu un atto veramente eroico: 600 cavalieri e 9.000 soldati maltesi e stranieri, sotto la guida del Gran Maestro Jean Parisot De (La) Valette, difesero le proprie postazioni contro 40.000 uomini inviati in 200 navi da Solimano il Magnifico. Il secondo motivo fu che l'isola era governata dall'Ordine Ospedaliero Militare di San Giovanni di Gerusalemme e Rodi, che era composto di nobili provenienti da tutta l'Europa. Le perdite furono ingenti: se i turchi persero il celebre ammiraglio Dragut e 20.000 uomini, l'Ordine sacrificò 250 cavalieri e 8.000 uomini, e questo significa che molte famiglie aristocratiche di Francia, Spagna, Portogallo, Germania, ecc. furono toccate da vicino. Per questa ragione l'isola divenne meta di parenti e curiosi, e l'evento ispirò centinaia di opere letterarie che, propagando l'evento, fecero scoprire l'isola che era governata da un'istituzione internazionale dedita alla difesa della cristianità e degli interessi marittimi degli stati menzionati sopra, e con la quale tutti i nobili e gli intellettuali sentivano una certa affinità. Si capisce, dunque, perché Hieronymus Megiser, che nel 1588-89 studiava legge a Padova, intraprese un viaggio culturale per l'Italia e decise

di fermarsi a Malta.

L'interesse di Megiser alle lingue fu precoce, e seguiva la moda collezionista di quel periodo: dalla raccolta di versioni del *Pater Noster* in 40, 50 e 52 lingue diverse arrivò al *Thesaurus Polyglottus* (1603). A Malta fu incuriosito dalla situazione linguistica dell'isola, governata da un ordine plurilingue che adottava come lingua ufficiale il latino, ma stava adoperando sempre più l'italiano, che era anche la lingua culturale degli isolani, i quali però, benché cristiani, parlavano una lingua semitica. Di conseguenza non solo incluse il maltese tra le 400 lingue del *Thesaurus* (rappresentato da 9 voci soltanto), ma gli dedicò una monografia, il *Propugnaculum Europae*, che fu ristampata tre volte tra il 1606 e il 1611. In quest'opera lo *Specimen* comprendeva 121 parole, trascritte impressionisticamente in alfabeto tedesco (cfr. MEGISER 1606, pp. 9-14 in FRIGGIERI-FRELLER 1998, pp. 135-38). Domenico Magri riportò una cinquantina di voci maltesi nella sua enciclopedia religiosa, *Notitia de' vocaboli ecclesiastici* (1644), Skippon nel racconto dei suoi viaggi (1664-1680) elencò 381 voci maltesi e il cavaliere Thezan produsse un glossario bilingue che registrò 3.110 parole nella sezione maltese-italiano e 3.925 nella sezione italiano-maltese del ms. *Regole per la lingua maltese* (tardo sec. XVII o inizio XVIII; CASSOLA 1992). Di un altro glossario settecentesco si è scoperta soltanto la seconda metà (da Ghedira a Zur) che consiste di 1.909 voci, il che permette d'ipotizzare un volume di circa 4000 lemmi in tutto (cfr. CASSOLA 1996). La filigrana indica che fu prodotto tra il 1742 e il 1775, e l'analisi di alcune bozze sembra attribuire il lavoro a Padre Pelagio, il quale avrebbe eseguito una revisione del glossario di Thezan eliminando i dieci grafemi arabi adoperati nell'originale e convertendo dunque l'ordine dei lemmi nel solo alfabeto latino (cfr. BRINCAT 2004, pp. 236-37). Verso la fine del Settecento furono compilati due veri e propri vocabolari da Gian Francesco Agius De Soldanis e da Michel Anton Vassalli. L'ambizioso vocabolario di De Soldanis, redatto tra il 1755 e il 1759 in quattro parti, non fu pubblicato, ma era

di gran lunga superiore ai tentativi precedenti perché nella sezione maltese-italiano-latino comprendeva 12.000 lemmi. Michel Anton Vassalli pubblicò il suo *Lexicon* a Roma nel 1796 e accolse un maggior numero di lemmi, ben 18.000. Nel corso dell'Ottocento si assisté a uno sforzo consapevole di codificare la lingua maltese, e furono redatte numerose grammatiche e vocabolari bilingui e trilingui, perché nel frattempo si era introdotto l'inglese come lingua alta accanto all'italiano. Alcuni si ponevano scopi pratici, ma il livello scientifico fu raggiunto da Dun Karm Psaila con un *English-Maltese Dictionary* in tre volumi (1936-1955) basato sul *Concise Oxford Dictionary*, dunque a senso unico, da Erin Serracino Inglott il cui *Miklem Malti* presenta 30.000 lemmi in 9 volumi (1975-1989) e offre importanti considerazioni etimologiche, e da Joseph Aquilina che dedicò gli ultimi decenni della sua vita alla compilazione dell'esauritivo *Maltese-English Dictionary* in due volumi (1987-1990), seguito dallo *English-Maltese Dictionary* in 4 volumi (1999-2000). Il suo vocabolario maltese contiene 41.016 lemmi, ridotti nella versione *Concise* (2006) a 22.649.

2. GLI ITALIANISMI NEI VOCABOLARI MALTESI

Nell'elenco di Megiser, che comprende solo le voci d'uso più frequente, solo tre sono di origine non araba (latina, siciliana o italiana): *Kattùs* ('gatto'), *Summara* (< *sciumara* 'fiume') e *Turco* (*sic*). Skippon accoglie 55 termini siculo-italiani su 355 voci raccolte "sul campo" (il 14,5%), e un numero maggiore di termini romanzi ricorre nei vocabolari di De Soldanis e Vassalli, malgrado le tendenze puristiche degli autori, entrambi fautori dell'ipotesi delle origini fenicie del maltese. Gian Francesco Agius De Soldanis nella sezione maltese-italiano-latino, composto tra il 1755 e il 1759, registrò 12.000 lemmi, di cui 2.000 circa sono di origine siciliana o italiana (il 17%), mentre Michel Anton Vassalli nel suo *Lexicon* (1796) portò i lemmi a 18.000, tra cui incluse le voci di origine siciliana o italiana d'uso più frequente, al contrario di Antonio Emanuele

Caruana che cent'anni dopo nel suo *Vocabolario della lingua maltese* (1903), ne eliminò gran parte, riducendo il lemmario ad appena 9.947 voci.

La base più affidabile per gli italianismi nella lingua maltese odierna è dunque il vocabolario compilato da Aquilina (1987-1990), l'analisi degli etimi del quale permette di presentare la seguente tabella:

<i>Maltese-English Dictionary (1987-1990)</i>			<i>Concise (2006)</i>	
totale lemmi	41.016	100%	22.649	100%
etimologia:				
arabo	13.293	32,41%	5.080	22,42%
siciliano	1.746	4,26%	977	4,31%
italiano	19.735	48,12%	12.980	57,30%
altri dialetti italiani	38			
latino	191		37	
francese	215		94	
inglese	2.511	6,12%	1.914	8,45%
altre lingue	269		46	
etimo sconosciuto	1.527	3,72%	273	1,21%
coniazioni locali	1.491	3,63%	1.248	5,51%

Tab. 1. La composizione del lessico maltese. Un confronto tra la versione piena e la versione minore *Maltese - English Dictionary* di Joseph Aquilina

Risulta chiara la preponderanza di voci di origine non araba nel maltese di oggi che complessivamente arriva al 76,78%. Tuttavia il maltese resta riconoscibile come lingua di origine araba per causa della morfologia, che resta in gran parte semitica, benché sostanzialmente semplificata, per le parole grammaticali e per il lessico fondamentale che comprende ovviamente le voci d'uso più frequente.

Si osserva infatti che nella traduzione in maltese dell'elenco delle parole fondamentali di Swadesh e Lees solo 6 su 100 sono di origine siciliana o italiana, però gli stessi concetti sul piano della terminologia avanzata, cioè dotta, tecnica, specializzata, astratta, o metaforica, sono spesso rappresentati da ter-

mini adottati dall'italiano. Se si tiene conto di questi derivati la cifra degli italianismi sale a 63 su 100. Qualche esempio: il sapere è *għerf*, la testa è *ras*, vedere è *tara*, la luna è *qamar*, grande è *kbir*, collo è *għonq*, dente è *sinna*, naso è *mnieħer*, ma poi si ha *sapjenza*, *testiera*, *veduta*, *lunari*, *grandjùz*, *kullana*, *dentatura*, *nažali*. Le parole per foglia, albero, coda e lingua sono rispettivamente *werqa*, *sigra*, *denb*, *ilsien* e sono rimaste nell'uso vivo per i significati di base, però *folja*, *arblu*, *kuda*, *lingwa* sono stati adottati per i loro significati estesi: 'foglia di un libro', 'albero della bandiera o della vela', 'coda dell'abito, a strascico', 'lingua, distinta da dialetto'. Le voci italiane indicano soprattutto concetti e oggetti che non si conoscevano (erano conosciuti) intorno all'anno mille, quando fu introdotto l'arabo a Malta: *akkwadott*, *dormitorju*, *incineratiur*, *terrapien*, *karnivàl*, *luviera* (portauovo), *manuskritt* arricchiscono il lessico perché affiancano le parole di base *ilma* 'acqua', *orqod* 'dormi', *rmied* 'cenere', *ħamrija* 'terra' e *mimli* 'pieno', *laħam* 'carne', *bajda* 'uovo'.

Tutto sommato il progresso materiale e culturale viene registrato nel lessico che cresce in modo esponenziale e questo processo viene confermato da Albert Borg e Manwel Mifsud che nel 1997 hanno pubblicato la versione maltese del lessico del livello soglia del Consiglio d'Europa: il loro elenco delle 1.585 parole fondamentali per la competenza comunicativa minima comprende 748 voci di origine semitica, 641 di origine romanza e 196 di origine inglese. Anche per un campionario così contenuto l'elemento non semitico raggiunge il 52,81% (MIFSUD-BORG 1997). Per quanto riguarda le parti del discorso, è interessante notare che la maggior parte degli italianismi registrati in AQUILINA 1987-1990 appartiene alla classe dei sostantivi (13.420), il che sembra ovvio, ma sono numerosi anche gli aggettivi (3.898), i verbi (2.010) e gli avverbi (597).

3. SICILIANISMI O ITALIANISMI

La tabella riprodotta sopra evidenzia che lo strato siciliano è di appena il 4,26% del lessico totale. Questa cifra, però, è basata soltanto sui sicilianismi “sicuri”, cioè che palesano tratti consonanti tipici del siciliano, il che può sorprendere se si tiene in mente che il dialetto siciliano è stato il primo contatto del maltese con una parlata romanza, e che è durato dal 1150 circa al 1550 circa sia come acroletto sia come basiletto, proseguendo anche fin a tutto l'Ottocento come interazione a livello orale tra mercanti, marinai, pescatori, notai, religiosi, insegnanti e amministratori. Il divario tra sicilianismi e italianismi è dovuto alla difficoltà di distinguere tra gli uni e gli altri, perché in effetti, tutte le voci italiane adottate in maltese vengono sottoposte al cambiamento delle vocali secondo l'uso siciliano, che conosce solo tre vocali in fin di parola (-a, -i, -u), e che trasforma E ed O lunghe in -i- e -u-. Questa corrispondenza è ancora produttiva in maltese, come anche alcuni raddoppiamenti consonantici, specialmente dei suffissi *-abbli*, *-attiva*, *-azzjòn*, *-azzjoni*, *-izzjoni*, *-izzja*, *-izzju* (*probabbli*, *indikattiva*, *nazzjòn*, *informazzjoni*, *pozzizzjoni*, *gustizzja*, *vizzju*). In questo modo tutti gli italianismi assumono una forma sicilianeggiante e pertanto questi criteri non li distinguerebbero. Il criterio più sicuro mi è sembrato quello delle corrispondenze non produttive, che sono tutte di tipo consonantico: S > [ʃ], SS > [ʃʃ], SK > [ʃk], ST > [ʃt], PL > [pl], PL > [tʃ], CL > [tʃ], FL > [fl], BI > [dʒ] (cfr. BRINCAT 2004, p. 125). Ammetto che in questo modo molti sicilianismi autentici si mascherano da italianismi, anche perché in realtà coincidono con termini italiani (per es. *komuni* < sic. *comuni* o < ital. *comune*).

4. CONSERVAZIONE E INNOVAZIONE

È interessante notare che il maltese conserva alcune voci antiche siciliane che sono cadute in disuso in Sicilia, come *flixxùn*

< sic. *flascuni*, ‘bottiglia’; *missier* < sic. *mio sire*, ‘padre, genitore’; *linka* < *inca* ‘inchiostro’. In alcuni casi non è chiaro se si tratti di un termine che è stato conservato soltanto a Malta, o di una coniazione locale su base siciliana: per esempio *ġelu* per ‘glassa’, *brazzi* per ‘appliques’, *bozza* per ‘lampadina’. Inoltre, alcuni termini siciliani e italiani subiscono un ampio processo di derivazione: la parola *ćuc* dal siciliano *ciucciu* (VS ‘somaro’ e ‘uomo ignorante’) è stata adottata soltanto col significato metaforico di ‘stupido’, ‘cretino’, perché l’animale ha conservato il termine arabo *ħmar*, ma *ćuc* ha generato sei derivati: i sostantivi *ćucùn* ‘ignorantone’ (**ciuccone*) e *ćucata* ‘fesseria’, ‘cretinata’ (**ciucata*); l’aggettivo *ćucèsk* ‘stupido’ (**ciuchesco*) e l’aggettivo comparativo *ićweć* ‘più stupido (di)’ (lett. ‘più ciuco di’); e anche i verbi *ćewweć* ‘far fare la figura di stupido a qualcuno’ (**ciucare*) e *iććewweć* (‘ha fatto una fesseria’; lett. **ciucarsi*). D’altro canto la voce italiana *pipa* ha generato il sostantivo *tipjip* ‘il fumo’ e il verbo *pejjep* (**pipare*) ‘fumare’, che si è esteso a qualsiasi tipo di attività del fumatore, non solo al fumare la pipa. Le neoformazioni del genere non sono rare (BRINCAT-MIFSUD 2016).

5. PSEUDO-ITALIANISMI

Come si sa, a Malta l’inglese è lingua ufficiale dal 1814, penetrando gradualmente nella scuola e nella società in concorrenza con l’italiano fino alla vigilia della seconda guerra mondiale, e irrompendo con l’istituzione della scuola primaria obbligatoria dal 1946 in poi. La maggiore esposizione all’inglese ha causato l’adozione di un numero sempre maggiore di termini inglesi e alla coniazione di voci dall’aspetto italiano ma che sono inesistenti nella lingua della vicina penisola (cfr. BRINCAT 2012). Alcuni esempi denotano nuovi uffici e ruoli sociali come *żviluppaturi* per ingl. ‘developers’ sulla base del verbo *żviluppa* (< *sviluppare* 3 p.s.presente) + il suffisso *-tùr* (< *-tore*), per analogia con *inventùr* < *inventore*, ecc. Allo stesso modo si è coniato il sostantivo *pubblikatùr* (< *ippubblika* +

tùr) per tradurre ingl. ‘publisher’, e sulla base anglo-latina di ingl. ‘prosecutor’ si è coniato il termine legale *prosekutùr* per ‘pubblico ministero’. Alcuni verbi e i loro relativi participi passati aggettivali hanno prodotto forme sconosciute all’italiano: *komméss* (< ingl. *committed*) ‘impegnato’, *relatàt* (< ingl. *related*) ‘connesso, attinente’, *irtiràt* (ingl. *retired*) ‘in pensione’, così via. Dal punto di vista morfologico i verbi adottati dall’inglese formano il participio passato con il suffisso *-at* (< *-ato*) o *-ut* (< *uto*; p. es. *spellùt* < *to spell* ‘scrivere una lettera dopo l’altra’) e l’infinito sostantivato con *-ar* (< *-are*): così il verbo ingl. *to assess* ‘valutare’ produce le forme *assessjàt* e *l-assessjàr*, e il verbo ingl. *to affect* ‘influenzare, riguardare, valere per’ ha il participio passato *affettwàt*, *l-affettwàr*.

Gli anglolatinismi vengono facilmente adattati al modello sicilianeggiante, e la maggioranza dei maltesi crede che siano parole di origine italiana: *evalwazzjoni* < ingl. *evaluation* ‘valutazione’, *akkomodazzjoni* < ingl. *accomodation* ‘alloggio (specie negli alberghi)’, *ritaljazzjoni* < ingl. *retaliation* ‘reazione (violenta, nel calcio o in battaglia)’, *plattitudni* < ingl. *platitude* ‘banalità’ (sul modello ingl. *altitude* = ital. *altitudine*), *okkupanza* < ingl. *occupancy* ‘la presenza negli alberghi’, *amenitajiet* < ingl. *amenities* ‘strutture e servizi, specie alberghieri’, *inforzabbli* < ingl. *enforceable* ‘applicabile, di legge o regolamento che si fa rispettare’, e ovviamente non si hanno problemi con il neologismo ingl. *governance* che diventa subito malt. *governanza*. Lo “strano” italiano dei maltesi lascia perplessi gli italofoeni che non sanno l’inglese, specialmente a causa di calchi polirematici: *ittra reġistrata* < ingl. *registered letter* ‘lettera raccomandata’, *sugġetti relatati* < ingl. *related subjects* ‘materie scolastiche associate, connesse’, *armamenti kemikali* < ingl. *chemical armaments* ‘armi chimiche’, *azzjoni industrjali* < ingl. *industrial action* ‘sciopero, vertenza sindacale’, *kumpanija reputabbli* < ingl. *a reputable company* ‘una società di buona reputazione’. Definisco strano l’italiano di Malta per il semplice fatto che in una conversazione con italiani i maltesi tendono a usare in modo disinvolto termini come **evaluazione*,

**accomodazione*, **ritaliazione*, **occupanza*, ecc., convinti che siano parole italiane. Fatto sta che, mentre una volta erano gli scrittori a fare la lingua, oggi giorno l'innovazione linguistica è in mano ai giornalisti, i quali chiamano sempre in causa la fretta con cui devono fare il loro lavoro e allora si sentono costretti a improvvisare. Risulta che spesso questi traducono "a orecchio", cioè sulla base della somiglianza fonetica, e raramente badano ai "falsi amici". Forse l'esempio più lampante è quello della *facilità korrettiva* per designare il carcere, che traduce inglese *correctional facility*, che non è altro che un eufemismo politicamente corretto, appunto, per 'carcere'. Un altro caso è *deputàt mexxej*, adoperato in alternanza con *viçi kap* (< *vice capo*), dove *deputàt* traduce l'inglese *deputy* 'vice', senza badare al fatto che *deputàt* (sostantivo) è corrente per 'membro eletto al parlamento'. Il neologismo viene ripetuto spesso nei giornali, alla radio e nei telegiornali e poi si dichiara che ormai è stato accolto dal pubblico generale.

È impossibile negare che la lingua deve crescere e adattarsi ai tempi ma oggi sembra che la rapidità e l'imponenza dell'adattamento di termini inglesi possano essere in qualche modo nocive. Una parte della responsabilità è attribuibile al mistilinguismo dei parlanti su scala nazionale. Avendo studiato la metà delle materie scolastiche in lingua inglese (e l'altra metà in maltese), il parlante medio tende a pronunciare la prima parola che gli viene in mente, in modo naturale e quasi inconsapevole, perché tanto questo è il parlato informale. Il pericolo, se pericolo è, consiste nel fatto che nella commistione di codice ogni parola inglese (e l'*Oxford English Dictionary* ne registra 616.000) è disponibile e spesso anche modificabile mentre, come abbiamo visto, il maltese standard com'è rappresentato nel vocabolario di Aquilina consta di 41.000 lemmi. La sproporzione è lampante, ma chi si preoccupa della conservazione del registro standard si conforta con il fatto che l'uso scritto, anche giornalistico, è più formale del parlato e pertanto richiede una certa sorveglianza. A parte questo, nel maltese scritto esiste il serio problema della grafia delle

parole inglesi di origine germanica che sono comunemente usate a Malta e che non sono adattate foneticamente. Alcuni giornalisti e scrittori sono a favore della loro trascrizione fonetica secondo le norme dell'alfabeto maltese, mentre altri obiettano che la veste inconsueta stona e preferiscono conservare la grafia originale. Tuttavia la questione non interessa in questa sede e ci limitiamo a osservare che tutti gli italianismi vengono facilmente e felicemente adattati ai modelli fonetici siculo-maltesi.

L'aspetto meno felice per quanto riguarda gli italianismi è che oggi l'esposizione all'italiano si è ridotta rispetto al passato. La televisione italiana, che dagli anni Cinquanta ha fatto riscoprire l'Italia e riaccessò il desiderio di apprendere l'italiano sia formalmente nelle scuole sia spontaneamente tramite l'ascolto regolare, oggi ha una audience giornaliera del solo 20% in prima serata rispetto al 52% di vent'anni fa. Pertanto come fonte d'innovazione lessicale l'italiano sta giocando un ruolo minore e, effettivamente, sono poche le parole adottate recentemente dall'italiano, e si limitano alla terminologia del giornalismo (*servizz, klandestini*) e dello sport (*kampjonàt, vertikalizazzjoni, konkluzjoni*).

GIUSEPPE BRINCAT

University of Malta / L-Università ta' Malta

Bibliografia

- AQUILINA 1987-1990 = Joseph Aquilina, *Maltese-English Dictionary*, Malta, Midsea Books.
- BRINCAT 2004 = Giuseppe Brincat, *Malta. Una storia linguistica*, Centro Internazionale sul plurilinguismo Università di Udine, Genova, Le Mani.
- BRINCAT 2012 = G. Brincat, *La produttività del superstrato sull'adstrato nell'integrazione di elementi esogeni*, in V. Orioles (a cura di), *Per Roberto Gusmani. Linguistica storica e teorica*, Udine,

Forum, vol. 2, t. 2, pp. 61-71.

BRINCAT-MIFSUD 2016 = Joseph Brincat - Manwel Mifsud, *Maltese*, in P.O. Müller *et al.* (a cura di), *Word-Formation. An International Handbook of the Languages of Europe*, vol. 4, Berlin, De Gruyter (in corso di stampa).

CASSOLA 1992 = Arnold Cassola, *The Biblioteca Vallicelliana 'Regole per la Lingua Maltese'*, Malta, Said International.

CASSOLA 1996 = A. Cassola, *Il Mezzo Vocabolario maltese-italiano del '700*, Malta, Said International.

FRIGGIERI-FRELLER 1998 = Albert Friggieri - Thomas Freller, *Malta, the Bulwark of Europe*, Malta, A. Friggieri.

MIFSUD-BORG 1997 = Manwel Mifsud - Albert Borg, *Fuq l-Għatba tal-Malti. Deskrizzjoni komunikattiva tal-Ilsien Malti fuq il-mudell tat-Threshold Level 1990*, Council of Europe.



GLI ITALIANISMI IN POLACCO: STORIA E REALTÀ ATTUALE

L'importanza del contributo italiano all'arricchimento del lessico polacco è confermata da numerosi studi, svolti innanzitutto da polonisti, sia come monografie a sé stanti che come parte di lavori storico-lessicali riguardanti vari campi semantici. Al fine di motivare la prospettiva di studio che si intende adottare nel progetto OIM, vorremmo proporre una rassegna sommaria delle principali opere e contributi dedicati al prestito dell'italiano in polacco che costituisca la base sulla quale formulare una proposta in merito.

Il prestito linguistico, sin dal famoso e sempre attuale studio di Weinreich (1963), viene considerato come frutto di interferenze risultanti dall'esposizione al contatto con una lingua straniera e con un'altra nazione, fortemente presente nella cultura di quella di arrivo. Al fine di valutare fenomeni di interferenza linguistica tra italiano e polacco occorre tracciare un breve panorama della storia delle relazioni tra le due nazioni, dei contatti non solo culturali, ma anche economici e politici, ampiamente documentati da storici e storici dell'economia soprattutto per il periodo di massima intensità corrispondente ai secoli XV-XVII.

1. LO SFONDO STORICO

Agli albori della presenza italiana nelle terre polacche stanno gli ecclesiastici, che arrivarono nell'XI secolo, su invito del re Bolesław Chrobry, desideroso di controbilanciare la presenza del clero tedesco (PTAŚNIK 1922, p. 3), anche se, come osser-

va giustamente Tygielski (2005, p. 166), essi furono rappresentanti della cultura latina e della chiesa cattolica piuttosto che italiani. Una presenza di stranieri¹ più sentita a livello economico inizia nella Polonia meridionale a partire del Trecento, in un paese ormai diventato potente sotto il regno di Kazimierz Wielki (Casimiro il Grande) e dei suoi successori Jagelloni. I nuovi arrivati, senz'altro dinamici e desiderosi di fare soldi e carriera, trovarono rapidamente il loro spazio in un paese vasto e influente. Già nel Trecento viene evidenziato il ruolo dei genovesi nella gestione delle miniere di salgemma (Bochnia e Wieliczka nei pressi di Cracovia) e dei metalli pregiati (argento, piombo, rame)², funzione senz'altro di prima importanza per il re e l'economia dell'epoca; nel corso del Quattro e Cinquecento seguirono banchieri, coniatori di monete, commercianti in articoli di lusso (seta ed altri tessuti pregiati, spezie), orafi ed altri artigiani, architetti e costruttori, medici e farmacisti, e finalmente artisti, musicisti e uomini di cultura che si stabilirono a Cracovia e in altre città della regione Małopolska (Piccola Polonia), ma anche a Lvov, Vilnius e Danzica. Non è irrilevante seguire le professioni dei nuovi arrivati né le cariche, spesso di riguardo, che occupavano, dato che esse testimoniano il ruolo svolto dagli italiani presso la corte reale, ancora prima che la principessa Bona Sforza diventasse moglie di Sigismondo (il Vecchio) e regina della Polonia (1518); inoltre si spiegano meglio alcuni atteggiamenti di insofferenza da parte dei polacchi che dovettero subire gli effetti della concorrenza straniera³.

¹ All'epoca, gli stranieri venuti dall'Europa occidentale e meridionale venivano chiamati *Gallici*, a prescindere dal luogo di provenienza; la denominazione *Italicus* si afferma nel XV secolo (PTAŚNIK 1922, pp. 7-8).

² Si considerino i capitoli II, V, VIII di PTAŚNIK 1922, nonché TYGIELSKI 2005, pp. 166-77.

³ Gli atteggiamenti, sia positivi che negativi, dei polacchi nei confronti degli italiani, la formazione su questa base dello stereotipo dell'italiano e l'ingresso dell'italianità sono discussi nei capitoli V-VI di TYGIELSKI 2005.

La letteratura riguardante le relazioni italo-polacche è ricca e racchiude studi di stampo storico, economico e culturale. Fonte alquanto preziosa di informazioni in merito è la monografia di Tygielski⁴ (2005) che va oltre ai fatti prettamente storici verso un'analisi sociologica del fenomeno della presenza italiana in terre polacche, studiando non solo i campi di attività, ma anche le motivazioni ed i meccanismi di inserimento dei nuovi arrivati nella società, la loro posizione sociale, le emozioni che suscitavano. Basata su un ventaglio molto largo di testi italiani e polacchi, documenti storici, ma anche corrispondenza, diari, relazioni di viaggio, scritti di umanisti e letterati, essa offre una panoramica larga ed esaustiva della presenza stessa degli italiani nel nostro paese, delle cause che li hanno spinti a lasciare la Penisola, nonché dell'influsso che l'Italia, gli italiani e l'italiano hanno esercitato sulla cultura polacca nel senso lato del termine.

Per lo stesso periodo, nell'ottica storico-economica di grande interesse per l'argomento della presenza degli italiani in Polonia in quell'epoca, vanno citati gli studi di Mazzei (1983 e 1999) dedicati alla circolazione di uomini d'affari e beni nell'Europa centro-orientale, alle grandi famiglie italiane insediate in Polonia (Soderini, Montelupi), il ruolo politico che hanno svolto (1999, pp. 72-93), agli itinerari e agli interessi dei mercanti. Allo stesso ambito appartiene l'indagine sul commercio italiano di seta in Polonia di Manikowski (1983), che, attraverso le sorti delle grandi famiglie fiorentine a Cracovia, mette in rilievo il meccanismo del rapido arricchimento dei commercianti in stoffe di lusso, nonché i fattori che lo hanno incitato.

Particolarmente fruttuosi per i contatti linguistici sono i

⁴ L'autore, professore di storia dell'Università di Varsavia, diplomatico e direttore dell'Istituto Polacco a Roma (1993-1995), ha posto al centro dei suoi interessi le relazioni italo-polacche tra il Cinque ed il Seicento. La vasta ed eterogenea bibliografia citata nell'opera permette svariati approfondimenti su temi e argomenti di interesse.

viaggi e le relazioni che ne hanno lasciato i protagonisti, sia italiani che polacchi. Gli italiani che si recano in Polonia, ne descrivono le città, gli abitanti ed i costumi contribuendo a diffondere la conoscenza di questo paese lontano, sono generalmente diplomatici, ambasciatori e segretari di uomini politici; la lettura delle loro memorie e relazioni di viaggio, spesso divertente, appare altrettanto istruttiva in quanto costituisce un riflesso dell'interesse per il diverso⁵. Invece del vasto movimento che spinge i polacchi a recarsi in Italia, chi per studio e istruzione presso le università italiane⁶, chi per diletto e perfezionamento personale, nonché dell'ispirazione che i polacchi traevano dalla cultura italiana, rendono conto numerose fonti polacche, delle quali ci limitiamo a segnalare le ormai classiche monografie di Barycz (1965 e 1969), Sajkowski (1973), Ulewicz (1999), che ne evocano svariati aspetti in un ampio contesto culturale e storico, posizionando nel contempo la Polonia rispetto all'Europa occidentale⁷.

Ne emerge che i rapporti italo-polacchi hanno una lunga tradizione e non stupisce che i diversi campi – economico, letterario, artistico, scientifico – in cui sono avvenuti i contatti tra le nazioni e le culture abbiano lasciato delle tracce a livello linguistico. Dai secoli remoti qui rievocati che, coincidendo con il tempo di gloria dell'italiano in tutta l'Europa costituiscono il punto di partenza per chi volesse seguire il percorso complesso dell'azione esercitata dagli italiani e dall'italiano sui polacchi e la loro lingua, si passa al Sette-Ottocento, periodo

⁵ Tra le relazioni di viaggio, particolarmente interessante è il testo di Fantuzzi (1998), che presenta la Polonia sullo sfondo di altre terre dell'Europa settentrionale. In questo contesto si veda inoltre WOLF 1949.

⁶ Occorre segnalare i contributi raccolti nel volume *Relazioni tra Padova e la Polonia. Studi in onore dell'Università di Carcovia nel VI centenario della sua fondazione*, Padova, Antenore, 1964, in quanto rendono conto, seppur in modo frammentario, del numero di polacchi – studenti, laureati e dottori – presso l'università patavina tra il Cinque e il Settecento.

⁷ Cfr. gli studi di Kot, raccolti da Barycz (1987) sulle ispirazioni francesi e italiane di intellettuali polacchi.

politicamente cupo, poiché corrisponde alla perdita dell'indipendenza conseguente agli smembramenti successivi della Polonia. Sebbene a partire dal Secolo dei Lumi il centro d'interesse si sposti verso la Francia, l'italiano continua ad essere percepito come lingua di sicura importanza, sia nell'ambito dell'arte che per i contatti regolari degli ecclesiastici polacchi con Roma⁸. Pur non facendo parte del curriculum scolastico, esso viene tuttavia insegnato in alcuni istituti di élite⁹.

Nel corso dell'Ottocento cresce l'importanza dell'insegnamento delle lingue straniere moderne, a svantaggio di quelle classiche, in quanto strumenti di conoscenza e mezzo di comunicazione: in Polonia, che a quest'epoca non ha esistenza come stato autonomo e subisce l'acculturazione forzata alla Germania e alla Russia, viene insegnato, oltre a queste due lingue, essenzialmente il francese e in modo più contenuto, anche l'inglese. Per quanto riguarda l'italiano, dai manuali pubblicati nel corso dell'800 risulta che esso manteneva il suo status di lingua di prestigio presso le classi colte pur non facendo parte di insegnamento istituzionale, condizione che perdura sia nel periodo 1918-1939 che nel dopoguerra¹⁰, testimoniando l'interesse continuo per l'italiano, che fosse dettato da motivi culturali, economici o religiosi. La posizione di lingua di prestigio in vari campi, dall'arte alle finanze e la gastronomia, congiunta a un atteggiamento di "simpatia storica" dei polacchi nei confronti dell'Italia ha indubbiamente facilitato la permeabilità ad influssi lessicali.

⁸ Cfr. CIEŚLA 1974, pp. 126-27.

⁹ Come i collegi dei padri teatini o il prestigioso "Collegium Nobilium" di Varsavia.

¹⁰ Per ragioni di spazio ci limitiamo a rinviare ad alcuni approfondimenti: SOSNOWSKI 2005 (pp. 174-76) per la posizione dell'italiano nella Polonia dell'Ottocento, PALMARINI 2013 per l'insegnamento dell'italiano nel periodo successivo all'indipendenza del 1918, JAMROZIK 2014 (pp. 27-30) per una panoramica globale dell'insegnamento dell'italiano in Polonia, compreso il periodo fino al 1989. Per indicazioni bibliografiche riguardanti l'insegnamento dell'italiano nel dopoguerra, si veda anche MIKOŁAJEWKI-BEDNARZ 1991.

2. GLI ITALIANISMI IN POLACCO: QUESTIONI DI METODOLOGIA

I contatti con l'Italia portano, attraverso i secoli, una notevole quantità di parole che si inseriscono in polacco con un grado diverso di assimilazione formale e semantica. Lo studio di questi prestiti¹¹ è tuttavia lungi dall'essere semplice per varie ragioni, oltre a quelle di mera tipologia: il problema più importante consiste nello stabilirne la fonte originaria; infatti, i prestiti dall'italiano hanno seguito varie vie di penetrazione nella lingua polacca, di cui il passaggio diretto dall'italiano al polacco è solo una delle possibilità. Le altre due grandi fonti di trasmissione sono, per ordine di frequenza decrescente, il francese, lingua di cultura che soppianta l'italiano a partire dal Seicento, e il tedesco, che ha dato al polacco il maggior numero dei prestiti in assoluto, senza dimenticare il ceco, che talvolta funge da tramite a parole germaniche. Il cammino tortuoso delle parole italiane verso il lessico polacco e il rapporto cronologico complesso delle etimologie viene segnalato da Witaszek-Samborska (1992, pp. 27-36), Giermak-Zielińska (1997, p. 367), Walsleben (1997, pp. 16-20), mentre Borejszo (2007, p. 12) limita chiaramente il suo campo d'interesse ai prestiti giunti direttamente dall'italiano¹². Nonostante la diversità di opinioni in merito, ci sembra che uno studio approfondito dell'impatto che ha avuto l'italiano sul lessico polacco debba prendere in considerazione tutti i casi in cui

¹¹ Occorrerebbe a questo punto riportare la complessa discussione sulla definizione e sulla tipologia del prestito, cfr. WEINREICH 1964, DERROY 1956 e GUIRAUD 1971, riassunta da WITASZEK-SAMBORSKA 1992 (pp. 14-16). Alla tipologia dei prestiti è inoltre dedicato lo studio di Cyran (1974) e in parte quello di Rybicka (1976). Si vedano inoltre i riferimenti bibliografici citati in BOREJSZO 2007 (pp. 12-14).

¹² Vanno segnalati inoltre vari articoli di Walczak che esamina da un punto di vista più teorico il problema di quella che chiama "profondità" dei prestiti polacchi; cfr. WALCZAK 1982 e 1997. Lo studio sintetico svolto da Biniak (2011) su un corpus di anglicismi contenuti in 10 dizionari polacchi del Novecento testimonia la variazione dei risultati nel corso della classificazione.

è possibile risalire all'etimologia italiana, a prescindere dalla via d'importazione. In questo senso l'esame degli italianismi in polacco andrebbe poggiato sullo sfondo di altri studi più ampi dedicati agli italianismi, quali HOPE 1971, SERIANNI 2008, DIFIT (2008), STAMMERJOHANN 2013.

Un'altra questione di interesse metodologico sorge con le fonti da prendere in esame nella costituzione del corpus. La ricca monografia di Borejszo (2007), la quale senz'altro, sia per l'eshaustività che per gli scopi prefissi di documentazione e analisi dei meccanismi di adattamento¹³, costituisce un modello per la presente ricerca, si concentra sugli italianismi nel polacco contemporaneo e si basa su fonti essenzialmente lessicografiche della seconda metà del Novecento fino ai primi anni del secolo in corso. Il corpus di Borejszo raccoglie 1096 lessemi di origine italiana, usati tutt'ora nel polacco, sebbene provengano da periodi cronologici diversi, dalla lingua antica fino ai primi del nostro secolo¹⁴: come afferma esplicitamente l'autrice, dei prestiti antichi sono stati presi in considerazione solo quelli che sussistono fino ai nostri giorni (p. 95). Sono state quindi deliberatamente scartate parole entrate nel polacco nei secoli passati, attestate sia nei dizionari che nei testi del

¹³ «Celem tej pracy będzie zatem zebranie i uporządkowanie tego, co w polszczyźnie 2. poł. XX wieku i przełomu tysiącleci można uznać za ślad oddziaływania języka włoskiego na polską leksykę. Opis obejmuje podstawowe zagadnienia dotyczące procesów adaptacyjnych zapożyczanych wyrazów [...] Zebrany materiał pochodzi z wydanych w 2. poł. XX wieku w Polsce leksykonów wyrazów obcych, słowników języka polskiego oraz innych opracowań grupujących słownictwo badanego okresu» (pp. 12-13). [“Lo scopo del presente lavoro sarà quindi raccogliere e sistematizzare quello che nel lessico polacco, nella seconda metà del Novecento e a cavallo dei secoli, può considerarsi traccia dell'influsso della lingua italiana. La descrizione comprende le principali questioni riguardanti i processi di adattamento delle parole prese in prestito [...] Il materiale raccolto proviene dalle opere pubblicate in Polonia nella seconda metà del Novecento: dizionari di parole straniere, vocabolari di lingua polacca ed altri lavori contenenti il lessico del periodo considerato”].

¹⁴ Per la cronologia dei prestiti raccolti si vedano le pp. 21-29.

periodo XVI-XIX secolo, ma ormai uscite dall'uso. Sembra tuttavia che in un lavoro di lessicografia elettronica, che non subirà i limiti di volume imposti necessariamente al cartaceo, tali lessemi dovrebbero trovare il dovuto posto, dato che sono la palese testimonianza di contatti culturali tra i paesi. Inoltre, si tratta spesso di terminologie appartenenti ai campi in cui l'Italia primeggiava, cadute in disuso con la scomparsa dei denotati, ma, ciononostante, di indubbio interesse culturale e storico non solo per gli italianisti, ma anche per i polonisti stessi. Una ricerca diacronica del genere fu svolta da Zawadzka (1976) su base di uno dei primi dizionari pubblicati in Polonia, il *Thesaurus polono-latino-graecus* di Grzegorz Knapski (Gregorius Knapius) del 1621, tuttora considerato una delle fonti più interessanti per la lingua polacca dell'epoca¹⁵. Nonostante le difficoltà riscontrate nello stabilire le etimologie, l'autrice ha rilevato 124 prestiti di origine italiana, la cui struttura semantica riflette i campi in cui l'italiano dominava all'epoca: quasi un quarto (28) riguarda il cibo, seguono i termini scientifici (botanica, zoologia, medicina – 10 in tutto), quelli del campo della moda (9) e della vita sociale, ovvero vita di corte (9)¹⁶. Tale struttura non stupisce se proiettata sullo sfondo del ruolo degli italiani alla corte di Cracovia. Inoltre, i prestiti raccolti da Zawadzka sono una testimonianza preziosa per l'evoluzione del polacco stesso: *antypast* (< *antipasto*), come *foryster* (< *forestiero*) sono caduti in disuso, sostituiti da formazioni indigene *przystawka*, *cudzoziemiec*; *ciarlatan* ha subito un ulteriore adattamento sotto l'influsso del francese, per cui la parola odierna è *szarlatan*, *cytryn* è diventato femminile – *cytryna*. Siamo convinti che un'indagine sistematica svolta sui prestiti contenuti nelle fonti antiche possa portare alla luce informazioni di rilievo per la storia sia dell'italiano che del polacco.

¹⁵ Si veda a proposito PUZYNNINA 1961.

¹⁶ ZAWADZKA 1974, p. 124.

Lo conferma lo studio di Walsleben (1997) dedicato ai prestiti dalle lingue romanze nel polacco del Seicento: basandosi non solo su fonti lessicografiche, ma anche su un corpus di testi di circa 1,63 milioni di parole corrispondenti a 35.000 lemmi, l'autore stabilisce una lista di prestiti dall'italiano e dal francese in polacco, confrontati in seguito con il tedesco (pp. 369-79); ogni prestito viene corredato da esempi tratti dai testi studiati, il che confluisce in una ricca e documentata panoramica degli usi effettivi dei prestiti romanzi nel polacco del Seicento¹⁷. Limitandoci a riportare un solo esempio della ricca raccolta di Walsleben (p. 83), la parola (tuttora viva, seppur letteraria) *awantaż*, oggi considerata un francesismo, nei testi seicenteschi appare anche in una grafia che ne segnala l'origine italiana (*awantagia* – adattamento grafico; oppure il plurale *awantadzi*, adattamento a partire della pronuncia): probabilmente nel corso del Seicento l'uso oscillava tra l'influsso italiano e quello francese, dopodiché la forma francese ha dominato. Lo spoglio dei testi dà inoltre la possibilità di fissare i limiti cronologici dei prestiti, di seguire l'evoluzione semantica verso usi traslati, le variazioni morfologiche e grafiche (*awizy/awizyjel/avisy*), le strutture sintattiche (*awizować* o *czymś*) e la combinabilità lessicale (*awizyje domowe, awizy dochodzą*)¹⁸. Informazioni di questo tipo danno l'opportunità di estendere la ricerca in altre direzioni, di andare al di là del mero studio dell'adattamento e arricchiscono la sincronia di una prospettiva diacronica indubbiamente preziosa per ogni studioso della lingua.

¹⁷ Una metodologia analoga, seppur necessariamente più ristretta perché non ancora elettronica, è stata adottata da Rybicka-Nowacka (1973) nello studio di parole di origine latina nei testi memorialistici del XVII secolo.

¹⁸ Esempi tratti da WALSLEBEN 1997, pp. 84-85.

3. GLI ITALIANISMI IN POLACCO: STATO ATTUALE E PROSPETTIVE DI RICERCA

Volendo dare un primo sguardo al corpus studiato da Borejszo, la struttura semantica degli italianismi nel polacco odierno rispecchia gli ambiti in cui era ovvia la supremazia dell'Italia nei secoli passati: dei 1096 lessemi raccolti, quasi la metà (488), riguardano le belle arti (musica, danza, balletto, letteratura, pittura, teatro), 171 la gastronomia; seguono varie terminologie tra le quali il gruppo più sostanzioso (54) spetta all'economia¹⁹. Per citarne alcuni:

- musica: *akompaniament, akord, adadżio, arpedżio, concerto grosso, kwartet, kwintet, madrygał, serenada, trąba, wirtuoz*
- pittura: *akwarela, akwatinta, fajans, graffiti, karykatura, majolika, terakota*
- letteratura e teatro: *bufon, groteska, impresario, impreza, pajac, sonet*
- gastronomia: *brokuły, cukinia, kapusta, karczoch, makaron, musztarda, por, pulpet, sałata, szparagi, szpinak, tort*
- economia: *awizo, brutto, bank, bankier, kasa, konto, kredyt, lombard, saldo, taryfa, waluta*

Tra le varie terminologie riportate da Borejszo, possiamo segnalare i seguenti campi semantici:

- arte militare (32 termini): *armata, arsenał, bastion, bomba, forteca, gwardia, kapral, karabinier, kawaleria, szpada, sztylet*
- architettura (28): *altana, balkon, belweder, fronton, galleria, kopuła, lodżia, pałac, parapet, sufity*

¹⁹ Vedi il capitolo III (*Kategorie semantyczne zapożyczeń włoskich*), pp. 30-45, specialmente la tabella dei dati numerici (p. 45).

- marina (23): *barka, busola, galera, korsarz, molo*
- moda (15): *krawat, koral, mankiet, ecc.*

Sarebbe senz'altro fruttuoso effettuare il confronto qualitativo e quantitativo tra la ripartizione semantica del materiale contemporaneo dei prestiti e quella, ad esempio, del Seicento, in modo da valutare la stabilità delle terminologie.

Sulla base di un primo accertamento che abbiamo svolto, indipendentemente dai dati raccolti da Borejszo, sui prestiti italiani del polacco attuale stanno *Słownik wyrazów obcych* (Dizionario dei forestierismi) della casa editrice PWN e *Słownik Języka Polskiego* della stessa casa editrice (consultabile al sito <http://sjp.pwn.pl/>). Al confronto del corpus di Borejszo, gli italianismi citati nell'ultima edizione del dizionario *Słownik wyrazów obcych* sono più numerosi: su circa 40.000 unità il dizionario evidenzia 1229 italianismi (segnati con *wł.* = włoski), pari al 3,5%; in questo numero sono tuttavia inclusi anche termini arcaici e letterari, testimonianze di vecchi costumi o maccheronismi di un'altra epoca che Borejszo ha deliberatamente scartato, come:

<i>antypast</i> < <i>antipasto</i>	<i>kares</i> < <i>carezza</i>
<i>lazaret</i> < <i>lazzaretto</i>	<i>oblig</i> < <i>obbligo</i>

Questi dati andrebbero comunque completati da altre opere lessicografiche rilevanti per il polacco moderno: *Słownik współczesnego języka polskiego* in 2 volumi di Dunaj (editore Wilga, 1995), *Uniwersalny słownik języka polskiego* in 4 volumi di Dubisz (editore PWN, 2005) oppure, specificamente per i prestiti, *Wielki słownik wyrazów obcych* di Bańko (editore PWN, 2005). Andrebbe inoltre considerati, specie per gli usi attuali, il corpus di riferimento del polacco *Narodowy Korpus Języka Polskiego* (<http://nkjp.pl/>) che costituisce la più ampia

fonte di dati del polacco contemporaneo²⁰. Occorre inoltre annoverare gli italianismi inclusi nel *Słownik Języka Polskiego* di Doroszewski (1958-1968, 10 volumi con supplemento), considerato il dizionario storico del polacco del Novecento, illustrato con esempi d'autore provenienti dalla letteratura polacca dell'Ottocento e del primo Novecento²¹.

In seguito gli italianismi raccolti da fonti lessicografiche dovrebbero essere sottoposti ad ulteriori accertamenti, specialmente per quanto riguarda i contesti d'uso, in base al sopramenzionato corpus di riferimento del polacco *Narodowy Korpus Języka Polskiego* che permette di seguire le tendenze di impiego degli italianismi nella lingua odierna.

Il campione rilevato nel *Słownik wyrazów obcych* e nel *Słownik Języka Polskiego* è stato incrociato con i dati statistici riportati dal lessico di frequenza *Słownik frekwencyjny współczesnej polszczyzny* (1990). Questo lessico, seppur per certi aspetti datato, in quanto basato sul materiale linguistico degli anni '80, dà un riferimento più saldo rispetto alle varie liste di frequenza riscontrate in rete²². Il campione preso

²⁰ *Narodowy Korpus Języka Polskiego* (Corpus Nazionale di Lingua Polacca) è stato elaborato sotto la direzione dell'Instytut Podstaw Informatyki PAN (Istituto dell'Informatica presso l'Accademia Polacca delle Scienze) in collaborazione con l'Instytut Języka Polskiego PAN (Istituto della Lingua Polacca presso l'Accademia Polacca delle Scienze), la casa editrice PWN e Zakład Językoznawstwa Komputerowego i Korpusowego (Dipartimento di Linguistica Computazionale e dei Corpora) presso l'Università di Łódź, enti ormai noti per gli studi svolti sui corpora. Il *Narodowy Korpus Języka Polskiego*, per le sue dimensioni e la metodologia degli spogli, viene considerato il corpus di riferimento per la lingua polacca.

²¹ Gli italianismi racchiusi in questo dizionario sono state oggetto di tre tesi di laurea dirette da D. Zawadzka nel 1987 e discusse presso la Cattedra di Lingua e Letteratura Italiana dell'Università di Varsavia, dedicate rispettivamente alle terminologie di: a) arte, letteratura e vita sociale (I. Nikisz); b) commercio, marina, vita militare e sportiva e scienze naturali (J. Orłowska); c) musica (M. Poławski). Questi tre corpora, composti in tutto di 474 unità, ricoprono la parte storicamente prevedibile dei prestiti dall'italiano.

²² Per le questioni riguardanti la frequenza segnaliamo inoltre le opere di Witaszek-Samborska (1993) e Zgółkowa (1983).

in considerazione nella formazione del lessico comprendeva 500.000 occorrenze, suddivise in 5 sottocampioni di 100.000 parole grafiche ciascuna, corrispondenti a 5 tipi di testi (divulgazione scientifica, articoli di informazione, articoli di fondo, prosa artistica, teatro), corredati da un campione di testi parlati. Nella lista di frequenze comprendente 10.355 lemmi, gli italianismi sono 638, il che corrisponde a 6,1% costituendo una frequenza relativamente alta. Nei singoli sotto-corpora del lessico la frequenza degli italianismi si ripartisce come segue:

T1 (divulgazione scientifica)	98	occorrenze	=	2,1%
T2 (stampa informazioni)	141		=	3,7%
T3 (stampa critica)	81		=	2,0%
T4 (prosa artistica)	147		=	3,7%
T5 (teatro)	91		=	3,3%
T6 (parlato)	80		=	3,2%

Spiccano per la frequenza i termini seguenti: *akompaniament* (nel significato giornalistico di ‘avvenimento simultaneo’, non come termine musicale), *awaria*, *final*, *gazeta*, *kampania (wyborcza)*, *kredyt*, *poczta*, *rakieta*. La stragrande maggioranza dei prestiti dall’italiano inclusi nella lista di frequenza riguarda la vita quotidiana ed è entrata in polacco in un’epoca relativamente recente (fine Ottocento, Novecento), confermando le constatazioni di Borejszo.

Per finire, ci siamo proposti di allargare questo campione con parole italiane penetrate nel polacco negli ultimi tempi, di cui alcune non sono ancora pienamente integrate e di conseguenza sono soggette ad oscillazioni morfologiche, grafiche e fonetiche. A questa ricerca, svolta su vari materiali linguistici (testi della stampa e pubblicità, inserti di cucina, menu dei ristoranti), abbiamo integrato i dati di un lavoro effettuato con un gruppo di studenti nel 1996 sui termini di cucina che all’epoca entravano timidamente nel mercato po-

lacco²³. Il confronto tra il materiale attuale con quello di quasi quindici anni fa dimostra che varie parole di origine italiana, specialmente nel settore di cucina, hanno ormai trovato un posto fisso nel lessico polacco: ne testimonia la diminuzione di varianti grafiche (o, se si vuole, di grafie erronee), come nel caso di *spaghetti* o *pizza margherita*, scritti ormai quasi regolarmente con la *h* (al posto della grafia polacca *spagetti*, *margerita*). In vari casi si nota anche la rinuncia all'iperonimia (*ser mozzarella*, ovvero 'formaggio mozzarella'): ormai è dato per scontato che l'utente polacco sappia che *mozzarella* o *parmigiano* siano formaggi. Segno di inserimento di una parola nel sistema linguistico è la creazione di derivati o alterati; il polacco forma facilmente diminutivi, per cui appare *ciabatka*, per una *ciabatta* più piccola. Nel linguaggio della stampa appaiono dei calchi da termini italiani, come *operacja czyste ręce* ('operazione mani pulite'), *skruszony* ('pentito'), *niebiescy* ('azzurri'), *ośmiornica* ('la piovra'); inoltre, nel caso di vocaboli poco trasparenti e propri per la realtà italiana, quindi difficilmente traducibili, si ricorre alla spiegazione che accompagna il termine italiano citato generalmente in corsivo o tra virgolette: *afera łapówkarska tangentopoli*, *styl gry catenaccio*.

Ne consegue che l'italiano risulta sempre presente nel polacco di oggi, seppur il grado di adattamento dei prestiti degli ultimi anni sia sicuramente minore, riducendosi sempre più spesso all'inserimento della forma italiana nel paradigma della declinazione (ad es. nom. *pizza*, gen. *pizzy*, acc. *pizze*; nom. pl. *pizze*). Ovviamente, dal punto di vista quantitativo l'italiano, in quanto fonte di prestiti, non può competere con l'inglese²⁴, tuttavia esso continua ad alimentare il lessico polacco, specialmente nel settore della cucina, testimoniando

²³ I risultati della prima ricerca sono stati in parte presentati in JAMROZIK 2000.

²⁴ Sarebbe tuttavia proficuo estendere la ricerca confrontando l'apporto quantitativo dell'italiano nel polacco di oggi con quello del francese, considerato tradizionalmente lingua di cultura.

della continuità dei rapporti che attraversano la storia della cultura polacca sin dal primo Rinascimento.

4. CONCLUSIONI PER IL PROGETTO DI RICERCA SUGLI ITALIANISMI IN POLACCO

In base a quanto esposto sopra, al fine di tracciare un quadro possibilmente completo della presenza dell'italiano nel polacco proponiamo di integrare le ricerche sulla lingua contemporanea con quelle di stampo storico. Per gli italianismi contemporanei, ovvero i termini che, entrati nel polacco in varie epoche storiche, vi permangono tutt'ora, occorrerebbe corredare regolarmente la lista esauriente stabilita da BOREJSZO 2007 con esempi d'uso tratti soprattutto dal *Narodowy Korpus Języka Polskiego* che fornirebbero attestazioni di strutture sintattiche, collocazioni e contesti di occorrenza, nonché le tendenze di evoluzione semantica delle parole italiane in ambiente polacco.

Inoltre, sempre sulla base del *Narodowy Korpus Języka Polskiego*, completato con fonti attuali, tratti principalmente dalla stampa che costituisce la testimonianza più attendibile di novità lessicali, andrebbero rilevati i "nuovi arrivati" dall'italiano. Questi vanno comunque considerati con la dovuta cautela, al fine di separare l'occasionalismo politico dal lessema che, oltrepassando lo stadio dell'effimero, ha certe probabilità di attecchire nella lingua di arrivo.

Dal punto di vista storico invece, di cui ci pare di aver dimostrato l'utilità per la presente ricerca, occorre innanzitutto passare al vaglio le fonti lessicografiche (di cui una prima lista provvisoria fa appendice alla bibliografia) che racchiudono dei dati stabili, separando nel contempo le unità lessicali tuttora vive da quelle ormai cadute in disuso. Per fonti lessicografiche intendiamo non solo dizionari, ma anche raccolte di termini o studi tematici sui settori in cui il polacco si avvaleva del prestito, come, ad esempio, BOCHNAKOWA 1984 per i termini

gastronomici tra il XV e il XVIII secolo²⁵ o BOREJSZO 1990 per i nomi del vestiario fino al 1600²⁶. Occorre annoverare anche gli studi sui dizionari stessi, che, non di rado, indagano la parte dei forestierismi nel materiale lessicografico (KURASZKIEWICZ 1988; MATUSZCZYK 2006).

In un secondo tempo, seguendo la strada tracciata da Walsleben per il Seicento, sarebbe opportuno svolgere una ricerca su vari tipi di testi considerati rappresentativi per ogni epoca, cercando di rinvenire, laddove possibile, i contesti d'uso per ottenere, anche nel caso dei prestiti arcaici, informazioni sulla loro combinatoria lessicale e sintattica. Tale ricerca, seppur laboriosa, porterebbe a un panorama più completo degli italianismi in polacco: andando oltre al semplice inventario, si potrebbero seguirne le etimologie, le varianti grafiche e morfologiche, le combinazioni lessicali. Il contesto storico, infatti, è indispensabile per comprendere appieno i fenomeni che intercorrono nella lingua di oggi, specialmente se così strettamente correlati alla cultura come i prestiti lessicali.

ELŻBIETA JAMROZIK
Uniwersytet Warszawski

²⁵ L'autrice studia l'apporto dalle lingue romanze in dizionari, testi letterari, libri di cucina e vi rinviene 119 prestiti, tra i quali 58 dall'italiano, 51 dal francese, 7 dallo spagnolo, 2 dal rumeno e 1 dal portoghese, notandone le varianti grafiche (*brochul*, *brochul*, *brokoli*; *fegatelle*, *figa tele*) e l'origine (*bianka* < *bianco mangiare*).

²⁶ Gli italianismi (tutti entrati nel corso del Cinquecento) sono 17 *vs* 81 prestiti dal tedesco, 56 dal latino, 29 dal ceco, 23 dalle lingue orientali, 11 dall'ungherese e 9 dal francese, riflettendo i vari influssi stranieri sulla cultura polacca. A questo studio segue Borejszo (2001) dedicato alle denominazioni del vestiario nel polacco contemporaneo: vi domina il francese, seguito dall'inglese, tedesco e russo; le altre lingue (tra cui l'italiano) vi sono rappresentate in numero nettamente inferiore.

Bibliografia

Studi

- BARYCZ 1965 = Henryk Barycz, *Spojrzenia w przeszłość polsko-włoską*, Wrocław, Ossolineum.
- BARYCZ 1969 = H. Barycz, *Z dziejów polskich wędrówek naukowych za granicę*, Wrocław-Warszawa-Kraków, Zakład Narodowy im. Ossolińskich.
- BINIEK 2011 = Kamila Biniek, *Etymology: who is right? Aspects of etymological research on the source languages of borrowings in selected Polish dictionaries*, in P. Chruszczewski - Z. Wąsik (a cura di), *Languages in Contact*, Wrocław, Wyd. Wyższej Szkoły Filologicznej, pp. 237-46.
- BOCHNAKOWA 1984 = Anna Bochnakowa, *Terminy kulinarne romańskiego pochodzenia w języku polskim do końca XVIII w.*, Kraków, Wyd. UJ.
- BOREJSZO 1990 = Maria Borejszo, *Nazwy ubiorów w języku polskim do roku 1600*, Poznań, UAM.
- BOREJSZO 2001 = M. Borejszo, *Nazwy ubiorów e współczesnej polszczyźnie*, Poznań, Wyd. WiS.
- BOREJSZO 2007 = M. Borejszo, *Zapożyczenia włoskie we współczesnej polszczyźnie*, Poznań, Wyd. UAM.
- BRÄHMER 1980 = Mieczysław Brahmer, *Powinowactwa polsko-włoskie*, Warszawa, PIW.
- CIEŚLA 1974 = Michał Cieśla, *Dzieje nauki języków obcych w zarysie*, Warszawa, PWN.
- CYRAN 1974 = Władysław Cyran, *Mechanizm zapożyczania wyrazów w języku polskim*, in «Rozprawy Komisji Językowej ŁTN», XX, pp. 23-37.
- DIFIT = Harro Stammerjohann *et al.*, *Dizionario di italianismi in francese, inglese, tedesco*, Firenze, Accademia della Crusca, 2008.
- FANTUZZI 1998 = Giacomo Fantuzzi, *Diario del viaggio europeo (1652) con Istruzione et avvertimenti per far viaggi lunghi*, testo a cura di P. Salwa e W. Tygielski ; introduzioni e note di W. Tygielski, Varsavia-Roma, Accademia Polacca delle Scienze.
- GIERMAK-ZIELIŃSKA 1997 = Teresa Giermak-Zielińska, *Les mots d'origine italienne en polonais contemporain*, in H. Stammerjohann (a cura di), *Italiano: lingua di cultura europea*, Tübingen, Günter Narr Verlag, pp. 363-73.

- HOPE 1971 = Thomas E. Hope, *Lexical borrowing in the Romance languages. A critical study of Italianisms in French and Gallicisms in Italian from 1100 to 1900*, Oxford, Basil Blackwell.
- JAMROZIK 2000 = Elżbieta Jamrozik, *L'italiano in Polonia ieri e oggi*, in S. Vanvolsem *et al.* (a cura di), *L'italiano oltre frontiera*, vol. II, Firenze, Cesati, pp. 319-28.
- JAMROZIK 2013 = E. Jamrozik, *La lingua italiana in Polonia: tradizione, stato attuale e prospettive*, in N. Maraschio - D. Di Martino - G. Stanchina (a cura di), *L'italiano dei vocabolari*, Firenze, Accademia della Crusca, pp. 109-27.
- JAMROZIK 2014 = Elżbieta Jamrozik, *Zarys historii nauczania języka włoskiego w Polsce*, in «Języki Obce w Szkole», 3, pp. 23-30.
- KOT 1987 = Jan Kot, *Polska złotego wieku a Europa. Studia i szkice*, a cura di H. Barycz, Warszawa, PIW.
- KURASZKIEWICZ 1988 = Władysław Kuraszkiewicz, *Wyrazy obce w polskim materiale Leksykonu Jana Maczyńskiego*, in M. Basaj *et al.* (a cura di), *Wokół języka: rozprawy i studia poświęcone pamięci prof. M. Szymczaka*, Wrocław, Zakład Narodowy im. Ossolińskich, pp. 251-57.
- MANIKOWSKI 1983 = Adam Manikowski, *Il commercio italiano dei tessuti di seta in Polonia nella seconda metà del XVII secolo*, Warszawa, Wyd. Uniwersytetu Warszawskiego.
- MAZZEI 1983 = Rita Mazzei, *Traffici e uomini d'affari italiani in Polonia nel Seicento*, Milano, Franco Angeli.
- MAZZEI 1999 = R. Mazzei, *Itinera Mercatorum. Circolazione di uomini e beni nell'Europa centro-orientale. 1550-1650*, Lucca, Pacini Fazzi.
- MATUSZCZYK 2006 = Bożena Matuszczyk, *Słownik języka polskiego S.B. Lindego: warsztat leksykografa*, Lublin, Wyd. KUL.
- MIKOŁAJEWSKI-BEDNARZ 1991 = Jarosław Mikołajewski - Irena Bednarz, *L'insegnamento dell'italiano in Polonia*, in I. Baldelli - B.M. Da Rif (a cura di), *Lingua e letteratura italiana nel mondo oggi*, Firenze, Olschki, pp. 511-17.
- PALMARINI 2013 = Luca Palmarini, *Un italiano "fortunato" a Cracovia. Biografia e produzione linguistico-letteraria di Fortunato Giannini tra Italia e Polonia*, in «Romanica Cracoviensia», 13, pp. 217-33.
- POCIECHA 1949 = Władysław Pociecha, *Z dziejów stosunków kulturalnych polsko-włoskich*,

- in H. Barycz - J. Hulewicz (a cura di), *Studia z dziejów kultury polskiej*, Warszawa, Gebethner i Wolf, pp. 179-208.
- PTAŚNIK 1922 = Jan Ptaśnik, *Kultura włoska wieków średnich w Polsce*, Warszawa, Instytut Wydawniczy „Bibloteka Polska”.
- PUZYNINA 1961 = Jadwiga Puzynina, *“Thesaurus” Grzegorza Knapiusza: siedemnastowieczny warsztat pracy nad językiem polskim*, Wrocław, Zakład Narodowy im. Ossolińskich.
- RYBICKA-NOWACKA 1973 = Halina Rybicka-Nowacka, *Rzeczowniki zapożyczone z łaciny w języku polskim XVII w (na materiale literatury pamiętnikarskiej)*, Wrocław, Zakład Narodowy im. Ossolińskich.
- RYBICKA 1976 = H. Rybicka, *Losy wyrazów obcych w języku polskim*, Warszawa, PWN.
- SAJKOWSKI 1973 = Alojzy Sajkowski, *Włoskie przygody Polaków. Wiek XVI-XVIII*, Warszawa, PIW.
- SERIANNI 2008 = Luca Serianni, *Gli italianismi nelle altre lingue romanze: prime riflessioni*, in *Italianismi e percorsi dell'italiano nelle lingue latine*, Atti del Convegno di Treviso, 28 settembre 2007, Ca' dei Carraresi, Treviso/Paris, Fondazione Cassamarca/Unione Latina, pp. 19-41.
- SOSNOWSKI 2005 = Roman Sosnowski, *Il dizionario italiano-polacco di Erazm Rykaczewski. Alcuni cenni sulla presenza dell'italiano in Polonia nel XIX secolo*, in «Revista di Italianística», 9, pp. 173-83.
- STAMMERJOHANN 2013 = Harro Stammerjohann, *La lingua degli angeli. Italianismo, italianismi e giudizi sulla lingua italiana*, Firenze, Accademia della Crusca.
- TYGIELSKI 2005 = Wojciech Tygielski, *Włosi w Polsce XVI-XVII w.*, Warszawa, Biblioteka Więzi.
- ULEWICZ 1999 = Tadeusz Ulewicz, *Iter Romano-Italicum Polonorum, czyli o związkach umysłowo-kulturalnych Polski z Włochami w wiekach średnich i renesansie*, Kraków, Universitas.
- WALCZAK 1982 = Bogdan Walczak, *Z zagadnień etymologizacji zapożyczeń romańskich w języku polskim*, in J. Tokarski (a cura di), *Język. Teoria – dydaktyka. Materiały V Konferencji Młodych Językoznawców Dydaktyków*, Kielce, WPS, pp. 172-94.
- WALCZAK 1997 = B. Walczak, *Słownictwo obcego pochodzenia na warsztacie badacza: problem granic „głębokości” opisu genetycznego*, in «Prace Slawistyczne», 105, pp. 269-79.

- WEINREICH 1964 = Uriel Weinreich, *Languages in contact: findings and problems*, The Hague, Mouton.
- WIDŁAK 2006 = Stanisław Widłak, *Italia e Polonia. Popoli e lingue in contatto*, Kraków, Wyd. Uniwersytetu Jagiellońskiego.
- WALSLEBEN 1997 = Albrecht Walsleben, *Romanische Lehnwörter in polnischen Texten des 17. Jahrhunderts*, München, Verlag Otto Sagner.
- WITASZEK-SAMBORSKA 1992 = Małgorzata Witaszek-Samborska, *Wyrazy obcego pochodzenia we współczesnej polszczyźnie*, Poznań, Wydawnictwo Nakom.
- WITASZEK-SAMBORSKA 1993 = M. Witaszek-Samborska, *Zapóżylenia z różnych języków we współczesnej polszczyźnie (na podstawie słowników frekwencyjnych)*, Poznań, Poznańskie Towarzystwo Przyjaciół Nauk.
- WOLF 1949 = Zygmunt Wolf, *Podróżnicy włoscy o Polsce XVII wieku*, in H. Barycz - J. Hulewicz (a cura di), *Studia z dziejów kultury polskiej*, Warszawa, Gebethner i Wolf, pp. 281-92.
- ZAWADZKA 1976 = Daniela Zawadzka, *Zapóżyżenia włoskie w języku polskim XVI wieku*, in «Kultura i społeczeństwo», XX, 1, pp. 117-26.
- ZGÓŁKOWA 1983 = Halina Zgółkowska, *Słownictwo współczesnej polszczyzny mówionej. Lista frekwencyjna i rangowa*, Poznań, UAM.

Fonti lessicografiche

Dizionari di lingua polacca fino al 1939

Lexicon Latino Polonicum ex optimis Latinae linguae Scriptoribus concinnatum, a cura di Jan Mączyński (Ioannes Maczinsky), Regiomonti Borussiae [Konigsberg], Typographus Ioannes Daubmannus, 1564 (versione online <http://www.wbc.poznan.pl> [14/07/2015]).

Thesaurus polono-latino-graecus seu promptuarium lingua Latinae et Graecae, a cura di Grzegorz Knapiesz (Knapski; Gregorius Cnapius), Cracoviae [Kraków], F. Caesario, 1621 (versione online: <http://www.mimuw.edu.pl> [14.07.2015]).

Nowy dykcjonarz, to jest mownik polsko-niemiecko-francuski, a cura di Michał Abraham Trotz, Leipzig, Johann Friedrich Gleditsch, 1764 (versione online: <http://ebuw.uw.edu.pl/dlibra/doccontent?id=802&dirids=1> [14/07/2015]).

- Słownik języka polskiego*, a cura di Samuel Bogusław Linde, 6 voll., Warszawa, Drukarnia XX. Pijarów, 1807-1814 (versione online: <http://kpbc.umk.pl/publication/8173> [14/07/2015]).
- Słownik języka polskiego*, a cura di Aleksander Zdanowicz *et al.*, 2 voll., Wilno, Maurycy Orgelbrand, 1861 (versione online: <http://eswil.ijp-pan.krakow.pl> [14/07/2015]).
- Słownik języka polskiego*, a cura di Jana Karłowicz - Adama A. Kryński - Władysława Niedźwiedzki, 8 voll., vol. I: Warszawa, Drukarnia E. Lubowskiego i S-k; voll. II-VI: Warszawa, Drukarnia Gazety Handlowej; voll. VII-VIII: Warszawa, Drukarnia Współczesna, 1900-1927 (versione online: <http://ebuw.uw.edu.pl/dlibra/docmetadata?id=236&dirids=1&tab=3> [14/07/2015]).

Dizionari contemporanei di lingua polacca

- Słownik języka polskiego*, a cura di Witold Doroszewski, 11 voll., vol. I-IV: Warszawa, Wydawnictwo „Wiedza Powszechna”, 1958-1962; voll. V-XI: Warszawa, Państwowe Wydawnictwo Naukowe, 1963-1969 (versione online: <http://doroszewski.pwn.pl> [14/07/2014]).
- Słownik języka polskiego*, a cura di Mieczysław Szymczak, 3 voll., Warszawa, Państwowe Wydawnictwo Naukowe, 1978-1981.
- Słownik współczesnego języka polskiego*, a cura di Bogusław Dunaj, Warszawa, Wydawnictwo Wilga, 1996, versione digitale su CD.
- Inny słownik języka polskiego*, a cura di M. Bańko, 2 voll., Warszawa, Wydawnictwo Naukowe PWN, 2000, versione digitale su CD.
- Uniwersalny słownik języka polskiego*, a cura di Stanisław Dubisz, 4 voll., Warszawa, Wydawnictwo Naukowe PWN, 2003 (versione digitale: <http://usjp.pwn.pl> [14/07/2014]).
- Wielki słownik języka polskiego PAN*, a cura di Piotr Źmigrodzki, Kraków, Instytut Języka Polskiego Polskiej Akademii Nauk, in fase di elaborazione dal 2007 (<http://www.wsjp.pl> [14/07/2014]).

Dizionari del lessico storico

- Słownik polszczyzny XVI wieku*, voll. 1-34 a cura di Maria Renata Mayenowa - Franciszek Peplowski, Wrocław, Ossolineum, 1966-1994; voll. 35- a cura di Krzysztof Mrowcewicz - Patrycja Potoniec, Warszawa, Instytut Badań Literackich Polskiej Akade-

mii Nauk, 1995- (versione online: <http://kpbc.umk.pl/dlibra/publication?id=17781> [14/07/2014]).

Słownik języka Jana Chryzostoma Paska, a cura di H. Koneczna - W. Doroszewski, 2 voll., Ossolineum, Wrocław-Warszawa-Kraków, Ossolineum, 1965-1973.

Słownik języka polskiego XVII i 1. połowy XVIII wieku, curato dal gruppo di ricerca presso l'Instytut Języka Polskiego Polskiej Akademii Nauk sotto la direzione di W. Gruszczyński, Kraków, Instytut Języka Polskiego Polskiej Akademii Nauk, 2004 (versione online: <http://sxvii.pl> [14/07/2014]).

Dizionari di parole straniere

Dykcjonarz zawierający wyrazy i wyrażenia z obcych języków polskiemu przyswojone, a cura di Michał Amszejewicz, Warszawa, Aleksander Gins, 1859 (<http://www.sbc.org.pl/> [14/07/2014]).

Słownik ilustrowany języka polskiego, a cura di Michał Arct; Warszawa, Wydawnictwo M. Arcta, 1916 (versione online: http://mbc.cyfrowemazowsze.pl/publication/5612_[14/07/2014]).

Trzaski, Everta i Michalskiego Encyklopedyczny słownik wyrazów obcych: pochodzenie wyrazów, wymowa, objaśnienia pojęć, skróty, przysłowia, cytaty, a cura di W. Trzaska - L. Evert - J. Michalski, Everta i Michalskiego, Księgarnia Wydawnicza Trzaski, 1939 (<http://mbc.malopolska.pl/> [14/07/2014]).

Wielki słownik wyrazów obcych, a cura di Mirosław Bańko, Warszawa, Wydawnictwo Naukowe PWN, 2005 (versione online: <http://sjp.pwn.pl/slowniki/Wielki-s-%C5%82ownik-wyraz%C3%B3w-obcych.html> [14/07/2014]).

Dizionario di frequenza

Słownik frekwencyjny współczesnej polszczyzny, Ida Kurcz *et al.*, Kraków, Wydawnictwo Uniwersytetu Jagiellońskiego, 1990.

GLI ITALIANISMI NELL'INGLESE DI TORONTO
E NEL FRANCESE DI MONTREAL.
STATO DELLE RICERCHE E DEL PROGETTO
SUGLI ITALIANISMI IN CANADA

1. INTRODUZIONE

Toronto e Montreal possono essere considerate le capitali culturali (e linguistiche), rispettivamente del Canada anglofono e del Québec; in entrambe, infatti, si trovano le sedi dei maggiori media nazionali (quotidiani e canali televisivi)¹, nonché delle università più importanti del paese². In termini di presenza italiana, culturale e sociale, queste due città sono accomunate da diversi elementi: si tratta di grandi realtà urbane in cui si è stabilito un alto numero d'italiani, dando luogo al formarsi di due consistenti "Little Italy", culturalmente e geograficamente ancora riconoscibili³ e con a disposizione, ormai

¹ A Toronto si trovano la CBC (il più importante canale televisivo canadese), la CTV Television, la TVO (il più importante canale televisivo dell'Ontario) e altri canali regionali o di diffusione cittadina. Per quello che riguarda i giornali: il maggior quotidiano nazionale, «The Globe and Mail», e il «National Post»; senza dimenticare il «Toronto Star», il quotidiano canadese con il ampio numero di copie in circolazione (vi lavorò anche Ernest Hemingway). Montreal ospita TVA e Télé-Québec. La sede di Radio Canada (versione francofona della CBC), per ovvie ragioni di natura politico-amministrativa, è a Ottawa. A Montreal, poi, risiedono i maggiori giornali della francofonia canadese: «La Presse», il «Journal de Montréal»; ma anche un quotidiano in lingua inglese di ampia diffusione, «The Gazette».

² La University of Toronto (prima università canadese) e, a Montreal, la McGill University (anglofona).

³ A Toronto le zone storiche dell'insediamento italiano si collocano in diversi punti; tra le principali, quelle situate tra le grandi arterie urbane di

da diversi anni, di propri mezzi d'informazione⁴. Entrambe le città vantano poi la presenza di un Istituto italiano di cultura e di una consistente presenza dell'italianistica accademica⁵. Sia la comunità italiana di Toronto sia quella di Montreal hanno espresso scrittori di rilievo, il cui successo editoriale ha permesso loro di distinguersi non solo in ambito "italo-canadese", ma anche a livello nazionale⁶. Detto questo, la ricerca non può e non potrà esclusivamente concentrarsi su questi contesti, ma anche e, in alcuni casi, soprattutto, su due entità geo-linguistiche più ampie e difficilmente circoscrivibili, ossia il Canada in cui si parla il *Canadian English* (CE) e il Canada dove si parla il *Français du Québec* (FdQ; ovviamente meno vago da un punto di vista territoriale, ma non così scontato).

College street e Queen street, e, più a nord, quella del Dufferin-Davenport District; negli ultimi decenni, tuttavia, si è verificata una forte emorragia di famiglie verso i sobborghi residenziali nella zona a nord di Toronto (Woodbridge, Vaughan, Markham). Per Montreal, un luogo storico della immigrazione italiana è quello situato nel quartiere Rosemont-La-Petite-Patrie (la "Petite Italie"), a cui si è in seguito aggiunto quello di Saint-Léonard (di una qualche utilità, soprattutto per un approccio storico di queste immigrazioni urbane, gli studi di ZUCCHI 1998 e MINGARELLI 1957; ben documentato e più interessante da un punto di vista sociolinguistico è invece l'articolo di POGGI 2008).

⁴ A Toronto, si trova il quartier generale di Omni TV, sezione italiana, canale televisivo con informazioni locali e nazionali, e altri programmi, e, per la carta stampata, il «Corriere canadese». A Montreal, la sede del «Corriere italiano». Vi sono poi diversi canali radiofonici che trasmettono in lingua italiana.

⁵ Nel caso di Toronto occorre inoltre registrare, dagli anni settanta in poi, una nutrita produzione critica, di provenienza soprattutto accademica e dedicata a caratteristiche ed evoluzione del cosiddetto "italiese", da alcuni ritenuto addirittura una sorta di "nuovo dialetto", nato dall'incontro tra le lingue degli immigrati italiani e l'inglese locale; a questo proposito, per avere un'idea, si vedano i primi contributi sull'argomento: PIETROPAOLO 1974; CLIVIO 1975; CLIVIO 1985; CLIVIO 1986.

⁶ Basti qui ricordare Nino Ricci, scrittore pluripremiato, autore del best-seller *Life of Saints*, cresciuto a Toronto; e, per Montreal, il celebrato drammaturgo Mario Micone, emigrato, da bambino, dal Molise natale in Québec.

2. QUALCHE STATISTICA

I dati della presenza linguistica italiana a Toronto e provenienti dall'ultimo censimento realizzato dal governo canadese (*Statistics Canada* del 2011)⁷ descrivono una situazione ancora in buona salute. I recensiti che dichiarano l'italiano come madre-lingua ammontano infatti a quasi 167.000⁸, situando così il loro idioma materno al quarto posto dopo inglese, francese e cantonese. Diverse le cifre riguardanti le lingue parlate a casa, con una riduzione drastica di circa 100.000 locutori, facilmente spiegabile in una realtà in cui sono numerose le famiglie eterogenee o con divari linguistici tra le generazioni. Si tratta a ogni modo di una presenza consistente, ma occorre tener conto che, secondo le stime del 2006, l'italiano era la prima lingua, geograficamente distribuita nelle zone urbane ed extraurbane sopra menzionate⁹.

Per quello che riguarda la presenza linguistica italiana a Montreal, l'ultimo censimento governativo colloca ancora l'italiano madre-lingua al quarto posto, ma, questa volta, dopo le due lingue nazionali, e dopo l'*Arabic* (un'ampia categoria designante i dialetti arabi, soprattutto magrebini)¹⁰. Le cifre riguardanti le lingue parlate a casa, con significanti riduzioni come nel caso di Toronto, mantengono comunque l'italiano in una comoda posizione.

3. QUALI STRUMENTI E QUALI METODI?

Il censimento degli italianismi sembra sempre barcamenarsi tra due poli principali: la necessità di affidarsi a grandi corpora

⁷ I dati del *Census Profile* 2011 del governo canadese sono disponibili online al seguente indirizzo web: (www12.statcan.gc.ca/census-recensement/ [14/07/2015]).

⁸ Maschi: 80.980, femmine: 85.440, totale: 166.415.

⁹ I dati del *Census Profile* 2006 del governo canadese sono disponibili online al seguente indirizzo web: (www12.statcan.gc.ca/census-recensement/ [14/07/2015]).

¹⁰ Maschi: 54.285, femmine: 55.835, totale: 110.125.

lessicografici da una parte e il controllo sui parlanti dall'altra. Nonostante i limiti e i difetti della prima scelta metodologica, la verifica sistematica degli italianismi registrati dai dizionari di altre lingue resta un punto di partenza imprescindibile. Certo, un'inchiesta condotta sul terreno, considerabile una cartina al tornasole della registrazione lessicografica, è auspicabile, sebbene non sempre materialmente applicabile. Nel campo della ricerca della diffusione degli anglicismi nelle lingue europee, per esempio, è la strada metodologica seguita da Manfred Görlach per il suo *Dictionary of European Anglicisms* (GÖRLACH 2001), il quale, partendo da opere di riferimento, si è poi affidato all'interrogazione di «native speakers» per ottenere conferme. L'impatto effettivo e reale di un italianismo è un dato forse più interessante di quello offerto da meri elenchi di termini tecnici, ormai probabilmente desueti o, al meglio, appartenenti a nicchie settoriali.

Quanto al contesto geografico, cominciando dai corpora lessicografici e dagli studi settoriali, per ampiezza della ricerca, completezza e novità dell'informazione il *Dizionario degli italianismi in francese, inglese e tedesco*, diretto da Harro Stammerjohann (DIFIT), rappresenta una solida base. Il DIFIT, tuttavia, ha un'utilità relativa per quello che riguarda il *CE* e il *FdQ*. Entrambe le lingue, soprattutto la seconda, non rientrano, del resto, negli obiettivi linguistici di Stammerjohann e dei suoi collaboratori, anche se alcune voci del *CE* sono registrabili grazie a spogli di studi e raccolte lessicali dedicati al *North American Slang*¹¹ e, almeno in un caso, attingendo direttamente alla realtà linguistica canadese: si tratta della voce *mangiakekka*¹². Il *FdQ*, invece, salvo errore da parte mia, non

¹¹ Soprattutto il volume di CHAPMAN-KIPFER 1998.

¹² Si tratta di un nomignolo affibbiato negli anni novanta dagli italo-canadesi ai canadesi di origine inglese, per via di alcune loro abitudini alimentari (cfr. DIFIT, s.v. *mangia-*), probabilmente calcolato su un termine già esistente nello slang americano, *cake-eating*, designante (soprattutto nel gergo studentesco d'inizio Novecento) «an effeminate fellow» (cfr. LIGHTER

è rappresentato, come del resto nemmeno quelli che sono stati denominati “français regionaux” (di Svizzera, Belgio, etc.). Data la presenza di voci provenienti dalla variante anglo-americana (anglo-australiana, anglo-canadese) dell'inglese, questo appare come una sorta di sbilanciamento e quasi un pegno da pagare nei confronti della lingua mondialmente dominante.

A parte il DIFIT, l'italianistica in generale (soprattutto quella nordamericana) non offre studi scientificamente validi sulla presenza degli italianismi nel *CE*, mentre è più generosa per quello che riguarda l'influenza esercitata dall'italiano sull'anglo-americano statunitense¹³. Quest'ultima produzione è pur sempre utile, dato che il *CE* appartiene, per ovvie ragioni storiche e geo-politiche, a una più ampia area, quella dell'inglese nordamericano (*General American*), nella quale *CE* e *American English* sarebbero varianti, con distinzioni di tipo fonetico e lessicale.

A questi studi si possono affiancare alcuni contributi più recenti che si concentrano sulla produzione d'italianismi definibile “endogena”, legata soprattutto (ma non esclusivamente) alla catena commerciale *Starbucks* (americana, ma ben presente anche sul territorio canadese) che ha introdotto falsi (o pseudo-) italianismi non solo semantici¹⁴, ma anche (e soprattutto) formali, sfruttando lo pseudo-suffisso *-ccino* (derivato da *cappuccino*): *frappuccino* o *freddoccino* o di recente, addirittura *tipoccino* (da *tip*, “mancia”), con uno slittamento considerevole dal campo semantico della gastronomia a quello

1994-1997, s.v.).

¹³ A tal proposito, uno spoglio di studi settoriali sugli italianismi negli Stati Uniti è già stato compiuto dal DIFIT, nella cui bibliografia compaiono gli studi pionieristici di Joseph Siracusa e Hermann Haller (cfr. DIFIT, pp. xxxi e xxxvii).

¹⁴ Come, per esempio, *venti*, *trenta*, indicanti le once con cui si misura una bevanda; o *grande*, altra categoria di misurazione, sempre per consumazioni liquide (cfr., a questo proposito, LANZILOTTA 2014).

più pragmatico del compenso economico¹⁵.

Questi studi, pur importanti, offrono spesso uno sguardo semanticamente e geograficamente limitato, più che altro in relazione con la gastronomia e geograficamente limitato all'area statunitense. Come vedremo, tuttavia, la terminologia studiata da questi contributi sarà inclusa a pieno titolo nell'ampia categoria degli italianismi che mi propongo di recensire.

Un passo successivo consiste nell'affidarsi direttamente ai corpora lessicografici prodotti dalla linguistica canadese. Questo comporta, per il ricercatore di formazione prettamente europea, un confronto con metodologie e cammini meta-lessicografici (e meta-linguistici) radicalmente differenti. Come ha scritto Valeria Della Valle, riferendosi al contesto italiano, la lessicografia offre un'immagine riflessa della società intellettuale «nelle sue varietà, nella sua mobilità e nel suo legame indissolubile con il passato» (DELLA VALLE 1993, p. 91). Questa affermazione appare ancor più stringente quando ci si riferisce al Canada, la cui storia culturale e politica è relativamente recente, linguisticamente sdoppiata, caratterizzata dalla ricerca di un'identità nazionale (e, al suo interno, di diverse identità regionali o locali) soprattutto attraverso l'auto-riconoscimento linguistico. A tutto ciò si associa e sovrappone, come vedremo, un particolare atteggiamento della linguistica accademica.

In quest'articolato processo socio-culturale uno spartiacque cronologico è, a mio parere, il 1977, anno del *Bill 101* o, meglio, della *Charte de la langue française*, una legge che definisce i diritti linguistici dei cittadini canadesi del Québec e stabilisce il francese come lingua ufficiale della stessa provincia. La legge, nella sua prima versione, prevedeva, tra le varie disposizioni, l'obbligo delle insegne commerciali (e non solo) monolingui (disposizione poi invalidata dalla corte superiore del Québec con conferma della corte costituzionale del Canada), ma, soprattutto, l'obbligo dell'istruzione in francese,

¹⁵ Il caso *freddoccino* è stato studiato da VEDOVELLI 2005.

tranne che per gli anglofoni (o meglio, per coloro i cui genitori avevano ricevuto un'istruzione anglofona in Québec; decisione, questa, poi emendata dal *Constitution Act* del 1982). È già percepibile in queste disposizioni, soprattutto la seconda, quali fossero le ripercussioni per le minoranze linguistiche (soprattutto gli italiani) che, invece, aspiravano a un'educazione anglofona per i propri figli e non tardarono a opporsi in modo politicamente agguerrito (cfr. POGGI 2008, pp. 329-30).

Lo spartiacque rappresentato dalla *Charte* appare valido non solo per le questioni politico-culturali, ma anche per alcuni snodi fondamentali nella storia della linguistica canadese.

Prima della *Charte* la rivendicazione accademica di un'identità linguistica canadese sembra riassumersi nella preparazione e conseguente pubblicazione (dopo alcuni testi preparatori, a destinazione prevalentemente didattica) del *Dictionary of Canadianisms on Historical Principles* (DCHP 1967). Il DCHP recensiva i "canadesismi", ossia i termini propri alla situazione linguistica canadese, tanto nel *CE* quanto nel "francese canadese" (quindi, non solo il *FdQ*) ed era il frutto del lavoro di un'équipe di accademici coordinati da Walter S. Avis, docente universitario di Linguistica a capo di un centro di ricerca lessicografico collegato all'Università di Victoria (situata nell'ovest del paese). Concepito seguendo il modello (di circa quindici anni più vecchio) rappresentato dal *Dictionary of Americanisms on Historical Principles* (MITFORD 1951), ma, soprattutto, dell'*Oxford English Dictionary*, il DCHP era basato su una meticolosa ricerca, condotta su una larga documentazione scritta, letteraria, ma anche amministrativa, trattatistica, giornalistica, etc.

Il *canadianism* era dunque un elemento lessicale o francese o inglese, e, data anche la scelta di lavorare solo su materiale documentario scritto, non era concepibile che un elemento lessicale precipuamente canadese potesse provenire da fonti esogene (o endogene, per il tramite della presenza delle grandi comunità di immigrati). Scorrendo il ponderoso volume del DCHP, salvo errore da parte mia, il solo *canadianism* non

“indigeno” è *grocceteria*, termine designante un piccolo negozio di alimentari (talvolta, dotato di self-service)¹⁶.

Negli anni in cui si preparava e si dava alle stampe il DHCP, la linguistica nordamericana preparava una svolta decisiva (e ancora duratura) verso studi perlopiù dedicati a materiali sincronici e orali, sulla scia dei lavori di William Labov (risalenti alla fine degli anni sessanta), ma anche grazie ai sentieri aperti da linguisti come Edward Sapir¹⁷. La *Charte*, nella mia ricostruzione storico-culturale, si sovrappone a questa svolta e, tutto sommato, l'accostamento non è anodino, dato che le nuove tendenze linguistiche comportavano ripercussioni scientifiche (ed extra-scientifiche) promuoventi e valorizzanti le autonomie e le identità linguistiche delle minoranze (basti pensare alla mole di lavori dedicati alle lingue amerindiane).

Dopo l'esperienza del DCHP, in Canada il lavoro lessicografico si è andato progressivamente slegando dai contesti accademici, pienamente convertiti alla ricerca sociolinguistica nordamericana. Emblematico, del resto, il fatto che il principale curatore del *Canadian Oxford Dictionary* (COD), da più di un decennio la fonte per eccellenza per la conoscenza del CE (non fosse altro per il peso economico-culturale della casa editrice) è una studiosa indipendente dall'accademia, Katherine Barber, conosciuta mondanamente come «Canada Word's Lady». La stessa équipe di collaboratori conta un solo docente universitario in veste di consulente, ossia un sociolinguista dell'Università di Toronto, Jack Chambers, specialista della variazione e del cambio fonetico (con le ovvie implicazioni sociali). Detto questo, il COD presenta peculiarità non trascurabili, soprattutto rispetto ai corpora lessicografici relativi

¹⁶ Cfr. DCHP, s.v. Penso si tratti di uno pseudo-ispansimo, costruito secondo il modello di *cafeteria*, o *washeteria* (termine, quest'ultimo, in voga nel Sud degli Stati Uniti).

¹⁷ Un'efficace sintesi di questa svolta periodale, con conseguente cambiamento di attenzione dallo scritto al parlato, è offerta in MARAZZINI 2014.

al *FdQ* (che vedremo tra poco): il COD proviene da un ambiente di lavoro più che altro “torontonian”, quindi da un ambiente multietnico e multilingue non marcato da esigenze identitarie; il COD, inoltre, disponibile anche online nella sua ultima edizione (COD 2005), si rivela molto più sensibile alla registrazione d’italianismi, non solo a quelli ereditati dall’inglese britannico (e, quindi, trasmessi attraverso i consueti canali esogeni della musica, dell’arte, della gastronomia classica) o a quelli provenienti dalla società italo-americana, ma anche agli italianismi che risultano frutto dell’immigrazione italiana in Canada.

Questa attenzione sembra essere mantenuta tra un’edizione e l’altra, con micro-aggiornamenti consistenti. È il caso, per esempio, della voce *calabrese* per la quale, nel passaggio dalla prima alla seconda edizione, si aggiunge una nuova accezione che rivela esistenza di un italianismo canadese:

COD 1998:

calabrese /,kalə'brɪ:z, -'breɪzeɪ/ *n. Brit.* a large variety of broccoli. [Italian, = Calabrian]

COD 2004 (e COD 2005):

calabrese /,kalə'breɪzeɪ/ *noun* 1 *esp. Cdn* a round white crusty Italian bread. 2. *n. Brit.* a large variety of broccoli. [Italian = Calabrian]

I micro-aggiornamenti possono anche comportare l’introduzione di nuovi lemmi, come nel caso di *friulano*, assente nella prima edizione del 1998, ma invece presente nell’edizione seguente (COD 2004 e COD 2005):

friulano /fri:u:'lano:/ *noun Cdn* a mild, pale yellow firm cow’s milk cheese of a kind made orig. in Friuli. ORIGIN Italian, = ‘Friulan’.

Friulano è il nome di un formaggio venduto dalla Saputo, impresa fondata da italiani a Montreal e, ormai, solidamente diffusa in tutto il Canada; si tratta di uno dei casi in cui un italianismo proviene da un nome commerciale, percepito come nome comune dalla lessicografia ufficiale e, a mio avviso, includibile come italianismo per l'OIM.

Pur pochi, questi esempi danno l'impressione che la redazione del COD sia in grado di recuperare italianismi locali, probabilmente interrogando informatori anglofoni: la stessa Barber si è detta disponibile a fornire liste di eventuali italianismi non ancora registrati¹⁸.

Durante gli anni del DCHP la storia accademica del Québec (almeno per quello che riguarda la linguistica francese/francofona e romanza in generale) era ancora in formazione, ma anche in piena effervescenza. Come ha ricordato il linguista e lessicografo Claude Poirier, in un articolo pubblicato quindici anni fa (POIRIER 2000), gli stessi allievi dell'allora migliore università quebecchese (Laval), per continuare i propri studi e completarli con un dottorato, partivano alla volta di Strasburgo, punto di riferimento francofono per la linguistica romanza (grazie soprattutto all'insegnamento di Straka e Baldinger e alla presenza della *Revue de Linguistique Romane*); negli anni settanta cominciavano a rientrare e gettare le basi per una nuova linguistica romanza e francese in patria. Proprio in quegli anni, infatti, maturava a Laval l'idea della creazione di un *Trésor de la langue française au Québec* (TLFQ) e la necessità di creare le infrastrutture scientifiche per una ricerca sul *FdQ*.

Alle origini del TLFQ vi erano proprio quei linguisti, appena ricordati, che si erano formati alla scuola strasbur-

¹⁸ Nel momento in cui l'ho contattata, la curatrice del COD era in Europa e senza documentazione a disposizione; a memoria, tramite un messaggio di posta elettronica (inviato in data 16 maggio 2014) mi ha fornito l'esempio di *gino*: nome comune per indicare il tipico italiano (con il corrispondente femminile *gina*).

ghese. I risultati del lavoro dell'equipe del TLFQ sono per ora disponibili online tramite il *Fichier Lexical du TLF*¹⁹; le basi metodologiche sono chiaramente di stampo prettamente franco-francese, con un modello lessicografico palesemente individuabile nel *Trésor de la langue française*.

Al tempo stesso, l'interesse del TLFQ, in sintonia con l'evoluzione delle politiche linguistiche e, in parte, con alcune delle tendenze della linguistica nordamericana, consisteva (e consiste) nel ricostituire in chiave lessicologica l'identità linguistica del Québec. Le stesse scelte lessicografiche hanno dunque una spinta sostanzialmente culturale:

D'autres équipes ou organismes abordent la question de la langue d'un point de vue politique ou d'un point de vue correctif. Le choix qu'a fait le TLFQ est de se concentrer sur le vocabulaire et de traiter les mots comme étant des objets culturels. Les historiens considèrent que les objets physiques, ceux qu'on place dans des musées, peuvent renseigner sur les pratiques anciennes et, indirectement, sur l'évolution de la société et de ses valeurs. À plus forte raison en est-il des mots dont l'étude peut aider à comprendre comment s'est constituée notre conscience collective et à dégager les caractéristiques fondamentales de notre identité (dalla *Présentation* del TLFQ).

Una ricerca sugli italianismi nel *FdQ* non potrà dunque privarsi di questo strumento, ma, forse, senza larghe aspettative sull'eventuale contributo che esso potrà fornire: il paesaggio lessicale delineato dal TLFQ, proprio in forza delle questioni identitarie e culturali che lo animano, difficilmente tende a un'inclusione e a un'ammissione di possibili (e, certo, esistenti) forestierismi; questo non solo a causa dell'atteggiamento dei redattori, ma anche delle fonti compulsate, spesso già marcate da un senso di appartenenza geografico e storico ben preciso.

¹⁹ (www.tlfq.ulaval.ca [14/07/2015]).

A ogni modo, la consultazione del *Fichier* non è totalmente inutile. Ecco un esempio di italianismo (*terrazzo*), reperito grazie alla paziente inserzione nella stringa di ricerca di italianismi presi dal DIFIT:

The screenshot shows the search results for the word 'terrazzo' in the Fichier lexical database. The results are organized into a table with columns for the word, its definition, and the source. The word 'terrazzo' is listed with several entries, including a definition in French and a reference to the Dictionnaire de la langue française.

Entrées simples	Titre	Année
1. terrazzo, n. En français des matériaux à couler le ciment on obtient une infinité de combinaisons de couleurs. Comme le terrazzo se travaille de la même façon que le béton, il est nécessaire de le couler dans des formes attribuées ou laissées à découper [...]. [...] Dictionnaire de la langue française, vol. 11, p. 79	quadr., marquet	1940
2. terrazzo, n. On pavimente dans le vestibule et surtout les cinq marches du terrazzo qui menaient au rez-de-chaussée. MONTMAYEUR, Yves (1914) : 58. Dictionnaire de la langue française, p. 603	idébitaire	1989
3. terrazzo, n. J'ai été assis à une table verte qui demandait à un petit vestibule déformé par de vieilles chaises où agonisait une femme en pain, couverte de tabliers indiennes, puis l'air peu réjoui les deux longues volées de marches de terrazzo qui menaient au deuxième étage [...]. MONTMAYEUR, Yves (1914) : 58. Dictionnaire de la langue française, p. 150-151.	idébitaire	1989
4. terrazzo, n. «Campes de bois», sans que cela paraisse sur les traits, une quantité impressionnante de fibres en fil, lorsque le rubanier, malgré l'essai, à un peu servir une voie plieuse, lorsque le serra enroulé des pins entiers de fibres sans que la tête s'élève fuyant le plancher de terrazzo de la traverse, [...] se sent en bonnet, sans fil» MONTMAYEUR, Yves Dictionnaire de la langue française, p. 74-75	idébitaire	1990
5. terrazzo, n. Ces tesselles de pavés de marbre ou de terre, de conservation entre voisins échangés à son-elle parce que le volume de la Masse doit mal confiné, d'annonces d'habitat faites au milieu par des voix qui se connaissent pas encore tout à fait leur recevoir ailleurs, de couleur éblouissante qui claquent de façon péroratoire [...]. TRENHALL, Michel (1942) : 51. TRENHALL, Michel, p. 79-80	idébitaire	1994

Tab. 1. Schermata del Fichier lexical con parola inserita

Per quello che riguarda i repertori di natura non specificamente lessicografica, per il *CE* si può disporre di alcune banche dati concepite e realizzate in un contesto accademico, con obiettivi scientifici generalmente di tipo pragmatico e con categorizzazioni di tipo etnico. Il dipartimento di linguistica dell'Università di Toronto ha prodotto corpora consistenti in trascrizioni di conversazioni in inglese, registrate in seno a diversi gruppi etnici (preliminarmente selezionati) della comunità urbana di Toronto²⁰:

Corpora in the Classroom (University of Toronto, Linguistics Department)

Corpus of English Conversational Speech (Toronto). Dir. S. Tagliamonte (University of Toronto)

Risultati che provengono ancora da inchieste etnolinguistiche condotte a Toronto sono invece raccolti in un corpus

²⁰ <https://corpora.chass.utoronto.ca/login.php> [14/07/2015].

organizzato da due professori di linguistica dell'Università di York, James Michael e Micol Hoffman, significativamente denominato *Ethnolects in the City*²¹.

Una banca dati molto più interessante e di libero accesso è quella dello *Strathy Unit*, il *Corpus of Canadian English*, realizzato presso il dipartimento di linguistica della Queen's University (Ontario, Kingston)²². Si tratta di un corpus di circa 50 milioni di parole, provenienti sia da documenti scritti (romanzi, periodici, quotidiani, trattatistica, testi amministrativi) sia da documenti di natura orale, consistenti perlopiù in trascrizioni di trasmissioni della CBC, appartenenti a un periodo di circa ottant'anni (dagli anni venti del Novecento agli anni duemila). L'interesse di questo corpus è molteplice: non solo materiale orale e sincronico, ma anche una vasta scelta di testi scritti. Ecco come si presenta la schermata prima di una ricerca lessicale:



Tab. 2. Schermata del *Corpus of Canadian English* prima della ricerca lessicale

²¹ Tutti questi corpora sono online, ma consultabili solo con un'autorizzazione da parte dei curatori. Purtroppo, le mie richieste di consultazione non hanno finora ricevuto risposta. Al tempo stesso, ai fini di un'inchiesta sulla presenza di italianismi, si presume che i corpora in questione abbiano un interesse limitato dal punto di vista lessicale. Lo schema delle domande risulta ripetitivo e presumibilmente esiguo lo spazio riservato alla produzione lessicale delle risposte degli interrogati.

²² corpus.byu.edu/can [14/07/2015].

L'inserimento di un comune italianismo come *alfresco* fornisce i seguenti risultati:

The screenshot displays the 'CORPUS OF CANADIAN ENGLISH (STRATHY)' search results for the term 'alfresco'. The interface includes a search bar with the term 'alfresco' and a 'SEARCH' button. Below the search bar, there are several filters and options, including 'SEARCH BY', 'FILTERS', and 'SORTING'. The main area shows a table of results with columns for 'RANK', 'WORD', and 'CONTEXT'. The table lists the word 'alfresco' and provides a sample context: '...the house was built after World War II. Perfect for alfresco summer dining under the stars. Not to worry about the large amount of patio furniture.'

Tab. 3. Schermata del *Corpus of Canadian English* con parola inserita

Lo *Strathy* permette dunque una ricerca d'italianismi in contesto canadese efficace per la consistenza del materiale messo a disposizione, ma anche per la comprensione diacronica e diamesica dei risultati ottenuti.

Quanto al *FdQ*, uno strumento proveniente dal mondo accademico quebecchese e disponibile online, è il *Corpus du Français Parlé au Québec* (CFPQ)²³. Il CFPQ è frutto del lavoro di linguisti dell'Università di Sherbrooke, all'est di Montreal, vicino al confine con il Maine, una significativa posizione geo-accademica: si tratta di un territorio storicamente e geograficamente più compromesso con l'elemento britannico o anglo-americano, rispetto alla cittadina di Québec (Laval). La stessa università è un'istituzione giovane (fondata nel 1954, mentre Laval ha radici storiche che risalgono al Seicento), certamente sensibile alle tendenze scientifiche delle vicine istituzioni anglofone, soprattutto statunitensi. Il CFPQ, del resto, sembra concepito con schemi, esigenze e metodi tipici della sociolinguistica americana, con fonti, come dice del resto il titolo, solo orali, ben strutturate secondo criteri di età e di *gender*. L'oggetto della ricerca, il *FdQ*, è innanzitutto

²³ <http://recherche.flsh.usherbrooke.ca/cfpq> [14/07/2015].

descritto grazie alla trascrizione (con attenzione non solo alla dimensione verbale, ma anche paraverbale e gestuale) ed è interrogabile in chiave non solo lessicale, ma anche semantica e pragmatica. Sicuramente, un lavoro immediatamente effettuabile è quello di una ricerca degli eventuali italianismi presenti in questo corpus, tramite la funzione *Liste des mots*, organizzata per ordine decrescente di frequenza. Mantenendo lo stesso italianismo utilizzato come esempio per la ricerca nel *Fichier*, ecco il risultato ottenuto sondando il CFPQ:

UNIVERSITÉ DE SHERBROOKE Centre d'analyse et de traitement informatique du français québécois Corpus de français parlé au Québec

Faculté des lettres et sciences humaines CFPQ

Accueil Présentation Conventions Vue d'ensemble Renseignements Recherche vendredi 06 juin 2014

Statistiques

Aucun critère spécifique

Tous les sous-corpus

Expression recherchée terraz*

1 occurrence de **TERRAZ*** dans 1 ligne de transcription (1 sous-corpus)

SOUS-CORPUS 4 (1)

1 - R : écoute ben j'ai travaillé déjà là-dessus planchers de **terrazzo** qu'il y a dans la grande allée dans l'église j'ai travaillé là-dessus moi ils étaient après bâtir l'église en aux alentours de cinquante ça dix-neuf cent cinquante quelque chose de même [sous-corpus 4, segment 4, page 51, ligne 1] - 100%

CHERCHER

STATISTIQUES

471,575 mots

21,016 mots différents

VIDER ACTUALISER

SEXE

Homme (40 participants)

Femme (58 participants)

AGE

SCOLARITÉ

Tous droits réservés © Université de Sherbrooke 2500, boul. de l'Université, Sherbrooke (Québec) CANADA J1K 2R1
Mise à jour le 25 avril 2014 - Application développée avec cadiciel Yii (version 1.1.12)

Tab. 4. Schermata del CFPQ con parola inserita

4. METODI

4.1. Tipologia degli italianismi

Innanzitutto, occorrerebbe stabilire una precisa tipologia degli italianismi e, questo, in modo preliminarmente indipendente dalle classificazioni adottate dall'OIM, a cui spetterà poi il compito di selezionare e, eventualmente, di ridefinire le differenti categorie.

Adottando la terminologia proposta da Luca Serianni (2008), gli italianismi da includere saranno quelli diretti, provenienti dall'italiano, per via popolare o letteraria, forse

il gruppo meno fertile, nonostante i numeri elevati della presenza italiana in Canada.

Poi, gli italianismi indiretti, trasmessi da una lingua intermedia. In questa categoria tenderei a inserire gli italianismi provenienti anche dall'inglese britannico e dal cosiddetto «français de référence» (*FrdRef*), diverso dal «français standard» (*FrSt*) che pure potrebbe essere considerato come una lingua intermedia²⁴. Come vedremo, tuttavia, il rifornimento di italianismi da parte dell'inglese britannico nei confronti del *CE* e del *FrdRef* nei confronti del *FdQ* non è così scontato.

Si aggiungono poi gli pseudo-italianismi endogeni (*PsItEnd*), prodotti perlopiù dal contesto immigratorio, anche se non bisogna dimenticare, risalendo la scala diafasica (e diastratica), il ruolo svolto da alcune riviste specializzate (per esempio, nel campo dell'architettura, della moda, etc.). In tal caso, un aiuto consistente potrebbe provenire dalla ricerca condotta attraverso lo *Strathy*. La categoria degli *PsItEnd* è suddivisibile, seguendo una recente classificazione fornita da FURIASSI 2012, in alcune sottocategorie, distinte fondamentalmente per modo di formazione:

- AC: autonomous compounds (es. *tutti-frutti*)
- AD: autonomous derivatives (es. *pepperoni*)
- CE: compound ellipses (es. *latte* < *caffèlatte*)
- C: clippings (es. *dildo* < *diletto*)
- SS: semantic shifts (es. *alfresco*)
- T: toponyms (es. *bologna*)

Gli *PsItEnd* non sono da confondere con quelli che Massimo Vedovelli ha definito «Esotismi di comunicazione sociale» (*ECS*): questa categoria metalessicale, non sembra godere della qualità di stabilizzazione, cara ai lessicografi, perché, in genere,

²⁴ Il *FrdRef* è il francese descritto dai dizionari, distinto dal *FrSt*, cfr. POIRIER 2000, pp. 146-47.

sinonimo di termini effimeri, in balia della comunicazione sociale e dell'interazione collettiva, esprime invece il risultato lessicale di dinamiche sociali (come quelli con suffisso *-ccino* visti poco sopra). Come lo stesso Vedovelli argomenta, gli *ECS* non dovrebbero avere un valore lessicografico in funzione dell'attestazione colta o del fatto che siano stabilmente radicati perché insostituibili termini di *realia*, ma anche in funzione dell'evento individuale e particolare della *parole*, nel senso saussuriano del termine, ossia dell'uso, della forma lessicale dovuta alle esigenze espressive derivanti dalle pressioni interazionali (cfr. VEDOVELLI 2005, pp. 587-90).

In tal caso, una lessicografia moderna (o contemporanea) che si affida (come del resto l'OIM) alla presentazione online dei dati lessicali si presta perfettamente alla registrazione degli *ECS*, per la rapidità dei tempi d'inserimento, di aggiornamento semantico (e, eventualmente, anche di decadenza dei termini stessi). Un esempio è fornito dal termine *friulano*, sopra menzionato: derivato da un nome proprio commerciale, si è affermato come nome comune; è del tutto possibile che il nome possa estinguersi seguendo i destini commerciali del prodotto oppure sopravvivere in forma autonoma.

4.2. La «*démarche de type différentiel*»

Il successivo passo metodologico consisterebbe nell'adottare una «*démarche de type différentiel*» (POIRIER 2000, p. 145) grazie alla quale un confronto tra i diversi repertori (lessicografici, banche dati, etc.) aventi come termine di confronto e base lemmatica il DIFIT chiarisca culturalmente e cronologicamente quali italianismi appartengano effettivamente al contesto canadese.

Questa applicazione contrastiva potrebbe innanzitutto contribuire a una riclassificazione diacronica delle aree semantiche di alcuni italianismi. Si veda l'esempio di *alfresco* nel

CE: rispetto ai dati offerti dal DIFIT²⁵ i risultati ottenuti dallo *Strathy* (confermati dall'ultima edizione del COD²⁶), permettono una definizione in cui il campo semantico pertinente alle tecniche artistiche non è più documentato dalle fonti del *CE*.

Oltre allo svecchiamento semantico, un altro risultato utile è la constatazione di italianismi endogeni, non derivati da logiche commerciali (che spesso caratterizzano la comparsa degli *ECS* e, in parte, dei vari *PsIt*), ma da una normale consuetudine. Sono stati già citati gli esempi di *calabrese* e *friulano*; si può ancora offrire quello di *caponata* (COD 2004 e COD 2005, s.v.; confermato da *Strathy*):

caponata /,kapp'natə/ *noun* a dish of eggplant, olives, and or tomato sauce, seasoned with herbs, served typically as an app [Italian *capponata*]

Anche per il *FdQ*, il *Fichier* del TLFQ e il CFPQ confrontati non solo col DIFIT, ma anche con i maggiori repertori del *FrdRef* possono rivelare la presenza di italianismi; un esempio, già registrato nel *CE* (probabilmente ereditato dal *British English*), è quello del termine (già citato) *terrazzo*, forse assorbito dal *CE*.

5. CONFRONTO CON I PARLANTI

La ricerca effettuata sui repertori deve essere necessariamente completata e convalidata dalle “indagini sul terreno”, ossia dal classico confronto con i parlanti, ma anche da altro materiale ascrivibile al patrimonio dei locutori e reperibile principalmente in Internet. Per quello che riguarda le inchieste orali,

²⁵ DIFIT, s.v. *Al fresco*: per la sezione inglese, la prima accezione è quella già conosciuta, ‘All’aperto’; la seconda è quella designante una tecnica artistica, ‘A fresco’.

²⁶ Che registra (s.v. *alfresco*) solo il seguente significato: «in the open air».

occorre costituire gruppi di parlanti interrogabili seguendo criteri eminentemente sociolinguistici; in tal caso, l'attività dei colleghi linguisti di Toronto, che hanno già delineato solidi criteri in questa direzione e hanno descritto con precisione tendenze etnolinguistiche, potrebbe risultare utile. Quanto al materiale attingibile da Internet, una limitata esplorazione è stata finora condotta monitorando alcuni forum con una vocazione precipuamente lessicografica, come, per esempio, quello del sito Wordreference.com; si tratta di un coacervo di dichiarazioni certo da filtrare, ma che merita attenzione in quanto risorsa metalinguistica (o, meglio, epilinguistica).

Un esempio è fornito da un italianismo del *CE*, il già citato *mangiakekka*. Il controllo lessicografico, completato da un'indagine condotta presso i locutori (oralmente, in modo limitato e non ancora con criteri ben definiti, e attraverso la navigazione in Internet) ha permesso di attestare un'effettiva caduta dell'uso dell'italianismo in questione (secondo lo stesso DIFIT risalirebbe agli anni novanta), una conseguente storicizzazione e, addirittura, una continuazione semantica attraverso una variante formale, *caker*, che di *mangiakekka* potrebbe essere un derivato²⁷.

6. CONCLUSIONI

Lavoro sui corpora disponibili, indagini sul terreno, esplorazioni in Internet: sono queste dunque le prospettive di ricer-

²⁷ L'indagine condotta in data 16 maggio 2014 sul sito Wordreference.com (sezione «Forums»), è confortata dalle informazioni trovate in un blog «Caker Cooking» (caker.cooking.blogspot.ca [14/07/2015]): «What a “caker”? In short, an Anglo-saxon who cooks with Chez Whiz, Cream of Mushroom soup and Cool Whip. The word “caker” is a short form of “mangicake”, a term coined by Italians» (quest'ultima affermazione, tuttavia, non sembra immune da tentazioni paraetimologiche); ma anche da un locutore canadese che non conosceva la parola *mangiakekka*, ma che, informato del significato, ha lui stesso proposto il termine *caker*.

ca preliminari per l'equipe che si occuperà degli italianismi canadesi. Non occorre però prescindere dalla collaborazione con chi lavorerà sugli italianismi nell'inglese degli Stati Uniti; infatti, per contiguità geografica e linguistica, diversi italianismi sono comuni a entrambi i paesi e, spesso, non è facile individuare l'esatta provenienza geografica.

Per concludere, mi permetto ancora un'osservazione, riguardante la natura dell'approccio scientifico.

La prospettiva lessicologica (e lessicografica) inaugurata dal DIFIT resta ovviamente prioritaria. Tuttavia, le prime esperienze di ricerca hanno messo in rilievo come i progetti, sopra citati, dei colleghi linguisti di Toronto, aprano anche prospettive di tipo morfosintattico. In effetti, alcune delle loro recenti ricerche insistono sull'influenza che una L2 esercita sulla L1 nel caso del parametro *pro-drop*. Si tratta di studi che paragonano l'uso del pronome soggetto in parlanti anglofoni nativi, cinesi di prima e seconda generazione e italiani di prima e seconda generazione, tutti gruppi che vivono in stretto contatto interazionale. In alcuni settori ben precisi (da un punto di vista generazionale e *gender*) del gruppo dei parlanti anglofoni nativi si denota appunto un *pro-drop*, una caduta del pronome soggetto, come probabile frutto del contatto con parlanti degli altri gruppi etnici (cfr., per esempio, HOFFMAN-WALKER 2010 e MARR 2011).

L'approccio morfosintattico certo si allontana da quello, eminentemente lessicale, dell'OIM, ma come Žarko Muljačić si augurava circa venticinque anni fa (in articolo divenuto ormai un punto di riferimento critico), accanto allo studio globale e comparativo della lingua italiana come fonte di rinnovamento lessicale esogeno di altre lingue occorrerebbe, almeno in prospettiva, tener conto anche del suo impatto fonologico e morfosintattico (MULJAČIĆ 1996, soprattutto p. 519).

FRANCO PIERNO
University of Toronto

Bibliografia

- CHAPMAN-KIPFER 1998 = Robert L. Chapman - Barbara A. Kipfer (a cura di), *American Slang. The Abridged Edition of the Dictionary of America Slang*, New York, Harper Collins.
- CLIVIO 1975 = Gianrenzo Clivio, *The assimilation of English loanwords in Italo-Canadian*, in P.A. Reich (a cura di), *The Second LACUS Focus: Proceedings*, South Carolina, Hombeam Press, pp. 584-89.
- CLIVIO 1985 = G. Clivio, *Su alcune caratteristiche dell'italiese di Toronto*, in «Il Veltro», XXIX, pp. 483-93.
- CLIVIO 1986 = G. Clivio, *Competing loanwords and loanshifts in Toronto's italiese*, in C. Bettoni (a cura di), *Altro Polo. Italian Abroad. Studies on language contact in English-speaking countries*, Sydney, Frederick May Foundation for Italian Studies - University of Sidney, pp. 129-46.
- COD 1998 = Katherine Barber (a cura di), *Canadian Oxford Dictionary*, Toronto, Oxford University Press.
- COD 2004 = K. Barber (a cura di), *Canadian Oxford Dictionary*, Toronto, Oxford University Press (seconda edizione).
- COD 2005 = K. Barber (a cura di), *Canadian Oxford Dictionary*, Toronto, Oxford University Press (edizione online: www.oxfordreference.com).
- DELLA VALLE 1993 = Valeria Della Valle, *La lessicografia*, in L. Serrianni - P. Trifone (a cura di), *Storia della lingua italiana*, I. *I luoghi della codificazione*, Torino, Einaudi, pp. 29-91.
- DCHP 1967 = W.J. Avis (a cura di), *Dictionary of Canadianisms on Historical Principles*, Toronto, Gage.
- DIFIT = Harro Stammerjohann *et al.*, *Dizionario di italianismi in francese, inglese, tedesco*, Firenze, Accademia della Crusca, 2008.
- FURIASSI 2012 = Cristiano Furiassi, *False Italianisms in British and American English: A Meta-Lexicographic Analysis*, in R. Vatvedt Fjeld - J.M. Torjusen (a cura di), *Proceedings of the 15th EURALEX International Congress (7-11 August 2012)*, Oslo, Department of Linguistics and Scandinavian Studies, University of Oslo, pp. 771-77.
- GÖRLACH 2001 = Manfred Görlach, *Dictionary of European Anglicisms*, Oxford, Oxford University Press.
- HOFFMAN-WALKER 2010 = Michael F. Hoffman - James A. Walker,

- Ethnolects and the city: Ethnic orientation and linguistic variation in Toronto English*, in «Language Variation and Change», 22, pp. 36-67.
- LANZILOTTA 2014 = Luca Lanzilotta, *Il caso Starbucks: l'italiano come lingua di commercio e cultura negli Stati Uniti*, in «Italice», XCI, pp. 71-88.
- LIGHTER 1994-1997 = Jonathan E. Lighter (a cura di), *Historical Dictionary of American Slang*, 2 voll., New York, Random House.
- MARAZZINI 2014 = Claudio Marazzini, *Come la linguistica scoprì il parlato e finì per trascurare lo scritto*, in «Quaderni grigionitaliani», LXXXIII, pp. 24-30.
- MARR 2011 = Isobel Marr, *Imposition and Identity in Null Subject: Contact effects among speakers of Chinese, Italian and Anglo background in Toronto*, MA Forum paper, University of Toronto (si tratta di una ricerca non edita).
- MINGARELLI 1957 = Giosafat Mingarelli, *Gli Italiani di Montreal. Note e profili*, Montreal, Centro Italiano.
- MITFORD 1951 = Mathews Mitford (a cura di), *A Dictionary of Americanisms on Historical Principles*, Chicago, Chicago University Press.
- MULJAČIĆ 1996 = Žarko Muljačić, *Innovazioni lessicali irradiatesi dall'Italia nel mondo. Per uno studio comparato degli italianismi nelle bocche e nelle menti europee, maghrebine e vicino-orientali*, in I. Baldelli - B.M. Da Rif (a cura di), *Lingua e letteratura italiana nel mondo oggi*, Atti del XIII congresso AISLLI (Perugia, 30 maggio - 3 giugno 1988), Firenze, Olschki, vol. 2, pp. 519-31.
- PIETROPAOLO 1974 = Domenico Pietropaolo, *Aspects of English interference on the Italian language in Toronto*, in «The Canadian Modern Language Review», XXX, pp. 234-41.
- POGGI 2008 = Irene Poggi, *La lingua italiana a Montréal*, in «Altre Italie», 36-37, pp. 324-36.
- SERIANNI 2008 = Luca Serianni, *Gli italianismi nelle altre lingue romanze: prime riflessioni*, in *Italianismi e percorsi dell'italiano nelle lingue latine*, Atti del Convegno (Treviso, 28 settembre 2007), Treviso-Parigi, Fondazione Cassamarca-Unione Latina, pp. 19-41.
- POIRIER 2000 = Claude Poirier, *Le français de référence et la lexico-*

graphie différentielle au Québec, in M. Francard (a cura di), *Le français de référence. Constructions et appropriations d'un concept*, Actes du colloque de Louvain-la-Neuve (3-5 novembre 1999), Louvain-la-Neuve, «Cahiers de l'Institut de linguistique de Louvain», XXVI, 1-4, pp. 139-53.

VEDOVELLI 2005 = Massimo Vedovelli, *L'italiano nel mondo da lingua straniera a lingua identitaria: il caso freddoccino*, in «Studi italiani di linguistica applicata», XXXIV, pp. 585-609.

ZUCCHI 1998 = John E. Zucchi, *Italians in Toronto. Development of a National Identity (1875-1935)*, Kingston-Montreal, McGill-Queen's University Press.



INDICE

<i>Premessa</i> di Matthias Heinz	Pag. 5
<i>Saluto</i> di Nicoletta Maraschio	» 7
<i>Presentazione</i> di Harro Stammerjohann	» 11
Matthias Heinz, <i>Dal DIFIT all'OIM. Sfide lessicografiche e prospettive di implementazione</i>	» 17
Luca Serianni, <i>L'italiano nel mondo. Intenti e propositi di un progetto editoriale sugli italianismi</i>	» 37
Peter Meyer, <i>The limits of lexicographical abstraction. Some strengths and problems of the data architecture in the Lehnwortportal Deutsch</i>	» 53
Giuseppe Brincat, <i>Gli italianismi nella lingua maltese. Non solo prestiti</i>	» 75
Elżbieta Jamrozik, <i>Gli italianismi in polacco. Storia e realtà attuale</i>	» 87
Franco Pierno, <i>Gli italianismi nell'inglese di Toronto e nel francese di Montréal. Stato delle ricerche e progetti sugli italianismi in CanadaLingua inglese all'Università?</i>	» 109





