

LUCILLA PIZZOLI, MATTHIAS HEINZ

# I vantaggi della ricerca su corpora per l'ampliamento e la verifica dei dati dell'OIM<sup>1</sup>

Nel contributo si intendono sottolineare le enormi potenzialità rappresentate dalla ricerca linguistica su corpora bilanciati rappresentativi dell'uso delle varietà linguistiche nelle quali sono stati censiti gli italianismi. Il controllo su corpora consente di rintracciare l'effettiva circolazione dei prestiti nell'uso e di misurarne il peso in modo più preciso rispetto a quanto documentato dalle fonti lessicografiche, nelle quali – soprattutto per il caso dei prestiti – possono intervenire fattori ideologici che sovrastimano o sottostimano l'impatto di una lingua straniera sulla lingua ricevente. Se il contributo di corpora sincronici risulta particolarmente prezioso nel caso dei neologismi, non sempre censiti in repertori lessicografici, anche l'indagine in prospettiva diacronica restituisce dati di grande interesse nell'indagine sugli scambi lessicali. La ricerca su corpora potrà dunque allargare la prospettiva di ricerca dei prestiti nelle tre direzioni di passato, presente e futuro.

*Parole chiave:* italianismi, corpora, lessicografia digitale, neologismi, arcaismi, archeologismi.

## 1. Introduzione

Non serve richiamare in questa sede come il fenomeno dei prestiti originati da parole italiane sia un tema di grande interesse, e che sia sempre più studiato da varie prospettive (anche in conseguenza della riscoperta del peso dell'italiano nel mondo, grazie alla constatazione dell'attrattività della lingua italiana come lingua di cultura, di studio, di lavoro)<sup>2</sup>. In quest'ambito si colloca il progetto OIM, che, come si

<sup>1</sup> Il presente testo è un prodotto collaborativo, la responsabilità redazionale per 1., 1.1, 1.3 e 2.1, 2.3 è di Lucilla Pizzoli e per 1.2 e 2.2 di Matthias Heinz.

<sup>2</sup> Si rimanda, per analoghe considerazioni, ai dati già richiamati in Pizzoli 2019; cfr. anche Mattarucco 2012 e Bombi & Orioles 2015, Bombi 2019; l'interesse per questo argomento è testimoniato dalla crescente mole di studi sulla diffusione dell'italiano nel mondo o in singole realtà geografiche e culturali. Basta vedere per esempio

potrà vedere, si può avvalere in modo particolarmente fruttuoso della ricerca su corpora.

### 1.1 Raccolte di prestiti nei repertori lessicografici

Partiamo da una considerazione generale: per lo più la ricerca di italianismi (e in generale di prestiti) si basa su estrazioni tratte dal lemma-rio di dizionari della lingua ospite, dal quale normalmente vengono selezionate le voci che hanno una marcatura specifica dal punto di vista etimologico, e quindi sono legate all'italiano come prestiti diretti, indiretti o calchi e quindi vengono isolati tramite questa trafila<sup>3</sup>. Tuttavia va precisato che la trafila etimologica può non essere ricostruita in modo omogeneo in tutte le lingue: l'accuratezza dei dati dipende dalla tradizione lessicografica di ogni lingua e soprattutto in molti dizionari potrebbe non essere aggiornato il comparto dei neologismi (perché prudentemente tenuti fuori dal lemmario in quanto frutto di contatto recente, legato a mode anche estemporanee); ancora, potrebbero esserci altri condizionamenti che determinano una

---

quanto pubblicato ogni anno nell'apposita sezione "L'italiano nel mondo" dalla rivista «Italiano LinguaDue» (<https://riviste.unimi.it/index.php/promoitals/index/>). Per i dati aggiornati sulla diffusione dell'italiano nel mondo si vedano i libri bianchi pubblicati a partire dal 2014 dal MAECI, a corredo delle edizioni degli *Stati generali della lingua italiana nel mondo*; 2014: *L'italiano nel mondo che cambia. Stati generali della lingua italiana nel mondo* (Firenze, Palazzo Vecchio, 21-22 ottobre 2014); 2016: *Italiano lingua viva. Stati generali della lingua italiana nel mondo* (Firenze, Palazzo Vecchio, 17-18 ottobre 2016); 2017: *L'italiano nel mondo che cambia. Stati generali della lingua italiana nel mondo* (Roma, Società Dante Alighieri, Palazzo Firenze, 18 ottobre 2017); 2018: *L'italiano nel mondo che cambia. Libro bianco degli Stati generali della lingua italiana nel mondo* (Roma, Villa Madama, 22 ottobre 2018); 2021: *L'italiano di domani. Stati Generali della Lingua e Creatività italiane nel Mondo* (Roma, Ministero degli Affari Esteri e della Cooperazione Internazionale – Sala Conferenze Internazionali, 29 novembre 2021).

<sup>3</sup> Tra le più recenti raccolte organiche, per esempio, si veda il ricco repertorio di italianismi in russo raccolto da Gherbezza 2019; si veda inoltre il DIFIT per inglese, tedesco e francese, e ancora Pinnavaia 2001 per l'inglese, Fábán-Szabó 2010 per l'ungherese, Gomez Gane 2012 per il catalano, Dashi 2013 per l'albanese, oltre a una cospicua mole di studi che indagano specifici ambiti semantici o epoche storiche, di cui è impossibile dare conto in questa sede e per la quale si rimanda alla raccolta bibliografica allestita proprio per il progetto OIM (consultabile online nel sito del progetto: <https://www.italianismi.org/fonti>).

attenzione ridotta – se non proprio l’esclusione – di alcune sezioni del lessico.

Nell’ottica di un osservatorio come l’OIM, concepito per misurare la profondità dell’impronta lessicale di una lingua sulle altre, conterà invece avere un quadro quanto più possibile esteso delle forme di contatto con la lingua ospite. Per questo motivo sarà utile avvalersi per la raccolta dei dati, oltre che delle banche dati lessicografiche, anche del tramite di altre fonti, non sempre prese in considerazione nella compilazione dei dizionari.

Per di più, all’interno del complesso paradigma della comunicazione nel mondo contemporaneo, hanno assunto un peso crescente le pratiche comunicative visibili nella sfera pubblica, in quella che viene classificata come “comunicazione sociale” e che include la fitta interazione tra gli individui per il tramite delle reti sociali e dell’ambiente esterno (Vespaziani 2018). In particolare, va notato che si riserva un’attenzione crescente, anche in linguistica, ai panorami linguistici urbani, proprio per monitorare la presenza di più lingue nel contesto multiculturale delle città<sup>4</sup>: i dati raccolti consentono dunque di misurare l’impronta semiotica di una o più lingue in un contesto specifico e dare conto del ruolo che rivestono le comunità linguistiche di antico o recente insediamento e degli equilibri che si vengono a creare tra diversi gruppi<sup>5</sup>. Altrettanto rilevante risulta, in questo quadro, il ruolo degli pseudoitalianismi, capaci di restituire l’immagine – sia pure semplificata e stereotipata – della lingua e della cultura italiana così come viene percepita nel mondo.<sup>6</sup>

---

<sup>4</sup> Gli studi sui *linguistic landscapes*, ben consolidati in area anglosassone (cfr. per es. Landry & Bourhis 1997 e Gorter 2006), sono stati applicati all’Italia e all’italiano grazie alle ricerche condotte nell’ambito dell’“Osservatorio linguistico nazionale dell’italiano diffuso fra stranieri e delle lingue immigrate in Italia” istituito presso l’Università per Stranieri di Siena nel 2000: cfr. Bagna & Barni 2006 e, da ultimo, per il panorama italiano negli ambienti urbani e scolastici, Bellinzona 2021.

<sup>5</sup> Indagini di questo tipo, avviate nell’Osservatorio senese, sono state documentate in studi concentrati su singole aree geografiche: cfr. per es. per il Giappone Vedovelli & Machetti 2006, per il Camerun Siebetcheu 2015.

<sup>6</sup> Per l’analisi di pseudoitalianismi in lingue e ambiti diversi si rimanda a Brincat 1991, Vedovelli 2005, Rieger 2008, Ille 2009, Vedovelli-Casini 2013, Casini 2015, 2017, 2018, Cassani 2015, Ferrini 2018 e ancora Siebetcheu 2015.

## 1.2 Un osservatorio sugli italianismi

Qui di seguito si darà qualche cenno sulla storia e lo sviluppo attuale del progetto di un osservatorio degli italianismi nelle lingue del mondo. L'idea nasce dalla pubblicazione nel 2008 in versione cartacea del *Dizionario degli italianismi in francese, inglese, tedesco* (DIFIT), concepito e redatto da H. Stammerjohann insieme a un gruppo internazionale di studiosi che ne curano le varie sezioni. Quasi nel contempo era in via di elaborazione un progetto editoriale di *Censimento* degli italianismi, curato da L. Serianni per l'editore UTET, che mirava alla documentazione di italianismi repertoriati in varie decine (oltre 70) di lingue del mondo intero. Gli autori del presente contributo erano già collaboratori dei rispettivi progetti. Del DIFIT si trae poi, per un complesso processo di retrodigitalizzazione, una prima versione in forma di banca dati interrogabile, messa a disposizione del pubblico dapprima all'interno del portale in rete VIVIT poi sul sito dell'*Osservatorio degli italianismi nel mondo* (OIM).<sup>7</sup>

Per ragioni di cambiamento di politica editoriale il progetto del *Censimento* non vede mai la luce in forma di opera a stampa (ne conseguono però alcuni trattamenti monografici basate sulle raccolte di italianismi come Gomez Gane 2012). Eppure, insieme con il nucleo dei dati del DIFIT una parte cospicua di tali materiali viene a costituire, sin dal 2014, la base della nuova piattaforma OIM, volta a fornire alla comunità scientifica uno strumento di lessicografia digitale atto a facilitare la ricerca sui prestiti passati dall'italiano in altre lingue. Definito come uno dei progetti strategici dell'Accademia della Crusca esso viene diretto da L. Serianni e M. Heinz e coordinato da L. Pizzoli, mentre la gestione della piattaforma informatica è affidata a M. Biffi e G. Salucci e l'elaborazione delle raccolte all'interno del gruppo di ricerca spetta a un numero crescente di unità operanti in varie sedi in Italia e all'estero.

Dalle tre lingue censite dal DIFIT si passa alle raccolte di italianismi in spagnolo, portoghese, catalano, polacco e ungherese mentre si stanno avviando i lavori di inserimento di ulteriori lingue di contatto quali il maltese, il neogreco, il macedone, il cinese mandarino e altre

---

<sup>7</sup> Il portale online VIVIT è stato progettato da F. Sabatini, N. Maraschio, D. De Martino, M. Biffi (<https://www.viv-it.org/schede/crediti>) mentre la banca dati DIFIT, opera di M. Biffi, G. Salucci, G. Seymer, M. Rago con la consulenza di H. Stammerjohann e M. Heinz (<https://difit.italianismi.org/crediti>), fa ormai parte dell'attuale sito dell'OIM ([www.italianismi.org](http://www.italianismi.org)).

in parallelo alla revisione e estensione dei dati sulle lingue francese, tedesco, inglese, prendendone in considerazione varietà anche al di là dei confini nazionali.<sup>8</sup> Allo stato attuale sono presenti nella piattaforma di inserimento dati più di 12.500 voci in lavorazione di cui circa 1.500 con dati completi. Complessivamente sono più di 8.900 le voci provenienti dalle raccolte di italianismi del DIFIT ma va crescendo il numero di quelle provenienti da altre lingue accanto al numero delle voci di partenza (“etimi”). Le schede relative a quelle voci sono affidate alla redazione italiana che ne verifica la semantica e le trafile e le inserisce nella piattaforma.

### 1.3 La classificazione dei prestiti nell’OIM

Si è ritenuto utile registrare nell’OIM tutte le forme di contatto, anche rilevate attraverso fonti estemporanee (scritture esposte, etichette di prodotti commerciali, ecc.) o segnalazioni di singoli rilevatori, purché se ne possa garantire una minima diffusione nell’uso e non si tratti di meri occasionalismi.

Per assegnare il giusto valore al canale di raccolta utilizzato, nella piattaforma dell’OIM vengono classificate con una diversa marcatura le fonti usate per la raccolta, a seconda del loro livello di affidabilità: ovviamente il livello di maggiore autorevolezza viene assegnato alle banche dati lessicografiche (dizionari dell’uso, storici, specialistici e dei termini stranieri), considerate fonti “autorevoli” in quanto fondate su criteri scientifici e studi che documentino i dati raccolti; vengono poi censiti gli italianismi raccolti tramite fonti “divulgative” (opere tradotte dall’italiano, guide turistiche, libri di viaggio, trattati o studi di taglio divulgativo che descrivono aspetti specifici della cultura italiana, articoli giornalistici, ecc.), cioè fonti documentate che non hanno le caratteristiche di scientificità dei lavori di ricerca ma presentano comunque un margine di affidabilità. L’etichetta di fonti “estemporanee” si applica invece a tutte le voci provenienti da segnalazioni occasionali raccolte per mezzo di motori di ricerca o tramite l’osservazione diretta sul campo operata dai ricercatori e tutte le forme di rilevazione non tradizionale (come per l’appunto scritture esposte, insegne, menu

---

<sup>8</sup> Per il francese e l’inglese, come anche per lo spagnolo, hanno un interesse particolare le varietà extraeuropee, nelle quali molti italianismi si sono prodotti come frutto del contatto dell’italiano degli emigrati con le lingue locali.

di ristoranti, pubblicità di prodotti, ecc.), utili per monitorare fenomeni recenti e a volte recentissimi<sup>9</sup>.

Tutte e tre le categorie individuate presentano dei limiti e dunque le raccolte che se ne ottengono non sono esenti da rischi, che varrà la pena di richiamare rapidamente.

I repertori lessicografici, ritenuti i più affidabili perché costruiti su ricerche condotte sulla base di criteri omogenei riscontrati in modo rigoroso, potrebbero essere influenzati dalla natura del corpus di partenza, non sempre rappresentativo della lingua nel suo complesso. Come sottolineava con qualche esempio Serianni (2008: 21-22), il dizionario potrebbe essere condizionato dalle caratteristiche specifiche delle singole raccolte lessicografiche, più sbilanciate sulla lingua scritta (e dunque poco rappresentative della diffusione di parole dell'uso vivo, anche recente): Serianni cita il caso del francese di Svizzera (ivi, p. 21, n.3), per il quale ci si basa sul *Dictionnaire suisse romand* (Thibault), fortemente incentrato sulla lingua scritta, specie amministrativa e politica. In altri casi il corpus di riferimento dei dizionari potrebbe essere falsato dall'esistenza precoce di repertori settoriali (è il caso di raccolte di termini musicali, come quelli pubblicati già nel Settecento per il francese; ancora Serianni 2008: 22); Ventura 2019 riporta l'esempio dell'inglese, arricchito da termini registrati in modo indebito da Florio e che hanno fatto entrare nei repertori inglesi voci tradotte da una presunta forma italiana mai usata<sup>10</sup>; e converso, molti repertori lessicografici difettano proprio di termini specialistici

---

<sup>9</sup> Sul tema della lingua italiana nel mercato globale si rimanda ai lavori avviati dal PRIN 2017 coordinato da Massimo Vedovelli su *Lingua italiana, mercato globale della lingue, impresa italiana nel mondo: nuove dinamiche linguistiche, socioculturali, istituzionali, economico-produttive* (Università di Siena Stranieri; Università degli Studi di Firenze; Università degli Studi Internazionali di Roma – UNINT; Università Telematica IUL). Per l'italianità linguistico-culturale nella comunicazione commerciale cfr. ora Mori 2021.

<sup>10</sup> È il caso delle forme inglesi *countergabbion*, *to countergabbion*, traduzioni delle corrispondenti voci italiane *contra-gabbione*, *contragabbionare* (non registrate in italiano dal GDLI) e introdotte nell'OED con un'unica attestazione nella seconda edizione del *World of words* di John Florio (1611). Tali parole sono «da ascrivere piuttosto al fascino dell'italiano da cui il Florio si lascia talvolta trascinare. Il suo dizionario, prezioso perché notoriamente molto attento al lessico tecnico (e che nasceva proprio dall'intento di rendere leggibili per il pubblico inglese anche le opere italiane contemporanee non letterarie), va perciò attentamente vagliato sulla base di altre fonti, contemporanee e posteriori» (Ventura 2019: 192).

perché non hanno introdotto, nei corpora di riferimento, testi di carattere tecnico: è il caso per esempio delle voci di ambito militare (cfr. ancora Ventura 2019: 170).

Inoltre, la presenza o l'assenza di termini stranieri potrebbe essere influenzata da fattori esterni, in particolare di tipo ideologico: il repertorio potrebbe essere più o meno prescrittivo (più o meno orientato verso una qualche forma di difesa della purezza della lingua) e condizionato dalla considerazione dell'immagine che storicamente la lingua da cui provengono i termini ha guadagnato in un certo ambito (in altri termini, il valore simbolico di certe parole potrebbe portare a sovrastimarne l'impatto sulla lingua ospite, determinandone l'accoglimento nei relativi dizionari). Oppure, potrebbero essere sottostimati dati riguardanti termini di recente ingresso che pure potrebbero descrivere le tendenze in atto nelle relazioni tra l'Italia e un altro paese. Un interessante settore di indagine potrebbe essere per esempio quello della manifattura, nella quale come è noto l'Italia detiene un primato a livello europeo: le proposte di singole aziende potrebbero aver favorito l'esportazione di termini legati a prodotti di successo. Tali termini potrebbero non essere stati ancora registrati nei dizionari ed essere recuperabili solo tramite la consultazione di repertori terminologici di ambito settoriale o attraverso i cataloghi che descrivono prodotti brevettati in Italia e poi esportati all'estero.<sup>11</sup>

Infine, va considerato il fenomeno che Della Valle & Patota 2013 definiscono degli archeologismi, cioè parole che, a differenza degli arcaismi veri e propri, rappresentano dei "residui passivi", cioè termini mai circolati in una lingua (o circolati solo marginalmente) e che tuttavia vengono trasferiti in modo inerziale da un'edizione all'altra dei dizionari (anche di quelli sincronici, che dovrebbero invece fotografare la lingua realmente "in uso") senza che se ne controlli attentamente la reale consistenza. Questo fenomeno è probabilmente più accentuato nella lessicografia italiana, più restia, per il timore reverenziale nei confronti della tradizione letteraria, ad abbandonare le forme del passato. Tuttavia potrebbe verificarsi un fenomeno analogo anche nei dizionari delle lingue straniere, nei quali potrebbero continuare ad essere registrati come vitali prestiti ormai appartenenti a un momento della storia di quella lingua in cui il contatto con l'italiano era più stretto.

---

<sup>11</sup> Si rimanda per queste considerazioni a quanto già presentato in Pizzoli 2019, dove si propone, come esempio, il caso di *pergotenda*.

Naturalmente anche le fonti cosiddette “divulgative” potrebbero essere condizionate dal testo fonte, specie nel caso di traduzioni, o di guide turistiche o testi riguardanti l’Italia, e proporre parole italiane o calchi di espressioni non realmente diffuse nella lingua ospite, ma rese necessarie dall’esigenza di descrivere elementi culturospecifici. Analogamente, anche le fonti “occasionalì” potrebbero raccogliere neologismi transeunti o veri e propri *hapax legomena* e devono dunque essere gestite con molta cautela.

Per avviare a queste possibili deviazioni, potrà essere molto opportuno il riscontro su corpora bilanciati, che potrà restituire la reale consistenza della diffusione delle singole parole, anche monitorandone, dove possibile, il percorso nel breve e lungo periodo, un’esigenza sottolineata in più occasioni da tutti coloro che si sono misurati con la ricerca di prestiti italiani: oltre al già citato caso del lessico militare, si può ricordare quanto si è osservato per la presenza di italianismi nell’arte (Motolese 2012), nel commercio (Wilhelm 2013), nella musica (Bonomi & Coletti 2015).

I vantaggi del riscontro su corpora possono essere evidenziati soffermandosi in particolare su due – se non tre – possibili direzioni: nel passato (eliminare gli arcaismi) e nel presente (individuare i neologismi); ma forse anche verso il futuro (prevedere la stabilizzazione dei neologismi).

Per quanto riguarda il passato, controllare le parole su corpora che presentano testi distribuiti in diacronia ci consente di collocare al giusto posto ogni parola, fissando meglio le datazioni della prima attestazione. Individuare una retrodatazione è un lavoro che viene fatto di continuo e che ovviamente il ricorso a corpora di grande entità che comprendono testi antichi consente sempre di perfezionare. Ormai quasi un paio di decenni fa Luigi Matt proponeva, in modo piuttosto innovativo per l’epoca, la consultazione di corpora (in particolare del catalogo di SBN) per la retrodatazione di termini specialistici che avevano attestazioni tardive solo perché i testi che li registravano non erano stati considerati come campioni rappresentativi della lingua e dunque non comparivano nei repertori tradizionali ma si trovavano nei titoli di opere dei rispettivi settori disciplinari<sup>12</sup>.

---

<sup>12</sup> Cfr. Matt 2004, che è riuscito a retrodatare oltre 1000 termini tecnici: «La relativa facilità con cui tale risultato è stato raggiunto permette di prevedere che un’applicazione a tappeto di questo metodo potrà fruttare un miglioramento sostanziale dei

Questo stesso controllo permette di monitorare l'arco di vita dei termini classificando eventualmente come arcaici o desueti termini non più attestati o scarsamente circolanti (e magari rimasti nei dizionari, anche sincronici, in modo inerziale). Documentare, anche per epoche passate, il rango di frequenza delle parole può inoltre servire a monitorare la vitalità dei termini e ricostruire la loro area di diffusione.

Per quanto riguarda il presente, il controllo su corpora sincronici consente di verificare l'esistenza e la consistenza della distribuzione nell'uso di parole che non sono state ancora censite nei dizionari di una lingua di interesse, magari lavorando in un'ottica comparata. Per esempio, partendo dal presupposto che alcuni termini potrebbero aver avuto una stessa trafila di diffusione in lingue geograficamente e culturalmente vicine, si può ipotizzare che la loro assenza in un repertorio dipenda dal fatto che la registrazione non è ancora stata completata ma che la parola sia comunque circolante. La conferma di questa ipotesi può essere garantita solo tramite il controllo su un corpus sincronico ampio e affidabile.

Dunque, a partire dalla lista di italianismi registrati in una lingua che può contare su una descrizione lessicografica particolarmente accurata, si potrebbe estendere il controllo su corpora rappresentativi di lingue che hanno avuto un analogo contatto con l'italiano per verificarne la presenza in queste lingue, specie ora che la globalizzazione rende così facili fenomeni di diffusione rapida di parole di moda.

Per quanto riguarda invece l'ultima possibile direzione, quella futura, il controllo su corpora di testi contemporanei consentirebbe di fare delle stime sulla stabilizzazione dei neologismi: l'analisi dell'andamento dei prestiti e del picco di frequenza, infatti, consentirebbe di ottenere elementi utili per valutarne la prospettiva di durata e l'opportunità di accoglierli nel lemmario di un dizionario.

Ancora una considerazione: il riscontro sistematico su corpora rappresentativi di testi autentici è in grado di restituire il dato del peso effettivo dell'italiano nella lingua ospite in modo molto più affidabile rispetto alla semplice registrazione lessicografica. Va ricordato infatti che i numeri di per sé non sono indicativi del contatto, né tantomeno della sua durata: molti italianismi potrebbero essere entrati in tem-

---

dati disponibili sui lessici settoriali» (p. 185). L'applicazione agli archivi elettronici è stata sperimentata anche da Telve 2002 per retrodatare voci onomatopiche e interiettive in testi letterari. Cfr. ora anche Cortelazzo 2022.

più remoti ed essere ora desueti (se non del tutto arcaici) nelle lingue ospiti, oppure risultare vitali solo in aree marginali della lingua<sup>13</sup>. Il controllo su corpora bilanciati e affidabili permette di assegnare all'italianismo un valore relativo rispetto alle diverse aree della lingua e misurare in modo preciso la diffusione del termine nella lingua ospite, eventualmente facendo un confronto più preciso anche della differenza tra una lingua all'altra (se i corpora sono confrontabili quantitativamente e qualitativamente).

Per l'OIM stiamo studiando un sistema di controllo che non si limiti solo al numero di prestiti registrati, ma consideri anche altre variabili: assegnando valori numerici fissati in base alla presenza della parola in un certo corpus e differenziando in base a parametri stabili si potrà misurare l'impatto dell'italiano sulle altre lingue<sup>14</sup>. In particolare, partendo dall'assunto che non tutti i prestiti valgono allo stesso modo, si potrebbe misurare il peso di variabili sia di tipo culturale (distinguendo tra prestiti legati alla vita privata e prestiti tecnici e culturali), sia di tipo più strettamente linguistico (analizzando cioè la presenza di termini concreti vs astratti, connotativi vs denotativi, permanenti vs effimeri).

Al numero di attestazioni (con la relativa percentuale nel campione di testi considerati) si potranno aggiungere dunque parametri di tipo culturale (rango di frequenza, ambito semantico, persistenza in diacronia, ecc.) e parametri di tipo linguistico (categorie grammaticali interessate dal contatto, livello di adattamento, produttività, ecc.). Il livello di penetrazione sarà massimo se la parola è entrata nel lessico di base, in categorie grammaticali meno effimere del nome (verbi, avverbi, ma anche affissi), se si è adattata e ha prodotto derivati (i cosiddetti "prestiti di secondo grado"); viceversa, l'impatto sarà meno rilevato se il prestito si limita ai nomi, si ferma alla periferia del repertorio di una

---

<sup>13</sup> Come è noto, Vidos 1965, ad esempio, elenca una serie di parole italiane passate ad altre lingue europee, soprattutto di ambito marinaresco, commerciale, musicale basandosi su studi di inizio Novecento. Prima di registrare le voci come prestiti sincronici, andrà controllato se le voci sono tuttora citate da dizionari moderni; solo il controllo su corpora, però, confermerà se quelle voci sono oggi vitali, anche soltanto in ambito tecnico; o se invece sono desuete (in tal caso ne dedurremo o che già le fonti di Vidos citavano le due forme per via libresca e inerziale o che queste erano vitali ancora ai primi del Novecento, ma sono oggi uscite d'uso).

<sup>14</sup> Rimandiamo alla proposta già abbozzata in Pizzoli 2017.

comunità (in un linguaggio specialistico), non presenta fenomeni di adattamento morfosintattico, ecc.<sup>15</sup>

## 2. *Il controllo sui corpora*

### 2.1 Esempi per lo spagnolo

Facendo un esperimento per lo spagnolo abbiamo provato a verificare la presenza di italianismi su due diversi corpora: quello offerto da Sketch engine (Spanish web 2018) e da Hemero<sup>16</sup>. Si può rilevare che la percentuale di occorrenze rispetto all'insieme delle parole raccolte nel corpus varia quando i numeri sono bassi. Per esempio, l'italianismo *arquitrabe*, un termine specialistico dell'arte attestato in spagnolo dal 1600 (cfr. DECH), risulta poco diffuso nella lingua comune ed è più soggetto ad oscillazioni nei valori percentuali nel confronto tra i due corpora (Sketch engine: 1853 (solo sing.) con una percentuale di 0.09 per million tokens; Hemero: 3 (2 sing; 1 pl. *arquitrabes*), con una percentuale di 0.00045 per million tokens). Uno scarto simile si ritrova anche in una parola come *cháchara* (dall'it. *chiacchiera*, attestato in spagnolo con i significati di "abbondanza di parole inutili", 1729; "conversazione frivola", 1551 e "oggetti di scarso valore", 1950: cfr. DECH; DLE 2014; NTLLE 2001): le occorrenze in Sketch engine sono 721 (sing.); + 277 (plur.), per una percentuale di 0.04 per million tokens + 0.01 per million tokens; in Hemero sono registrati 78 casi (73 sing; 5 pl. *chácharas*) per una percentuale di 0.12 per million tokens. Essendo un termine poco circolante, ha percentuali evidentemente casuali nei due corpora di controllo.

<sup>15</sup> Su problemi di questo tipo cfr. ancora quanto precisato da Serianni, che ha dimostrato come la qualità dell'apporto di una lingua vada valutata tenendo conto del comparto del lessico che viene coinvolto: limitando l'indagine al francese, una lingua pur profondamente legata all'italiano, si può misurare in poco più del 5% l'apporto totale degli italianismi al lessico francese fondamentale (Serianni 2008: 33).

<sup>16</sup> Sketch engine (<https://auth.sketchengine.eu>) è uno strumento che consente ricerche di tipo linguistico su corpora molto ampi di lingue diverse: uno dei corpus disponibile per lo spagnolo (Spanish web 2018) contiene testi provenienti da *Spanish Wikipedia*, *European Spanish web*, *American Spanish* raccolti tra febbraio e aprile 2018, per un totale di 17,553,075,259 parole; Hemero, purtroppo attualmente non più disponibile online, è un corpus di testi tratti da giornali spagnoli e ispanoamericani pubblicati tra il 1997 e il 2009, contenente oltre 660 milioni di parole.

Al contrario la parola *fiasco*, un termine di larghissima diffusione in molte lingue europee (sia con il significato di “bottiglia”, sia con quello di “fallimento, insuccesso in uno spettacolo, specie teatrale”, in spagnolo dal 1884, cfr. NTLLE), avendo numerose attestazioni mostra una percentuale più stabile: in Sketch engine: 28.694 attestazioni, per una percentuale di 1.41 per million tokens, in Hemero 900 attestazioni, per una percentuale di 1.36 per million tokens.

È inoltre particolarmente utile il riscontro su corpora per avere la conferma della scarsa o nulla circolazione di termini che già i dizionari definiscono desueti: nel DECH l’italianismo *fazoletto* viene definito «antic. y raro», e documentato dal 1611 in Covarrubias, con la precisazione che «Sólo este autor documenta el vocablo en castellano»<sup>17</sup>.

Questo termine, prevedibilmente assente in entrambi i corpora di controllo, si può definire un vero e proprio archeologismo (già nel 1726 il *Diccionario de autoridades* della Real Academia Española precisa che «Tiene poco o nignun uso»): si tratta con ogni evidenza di una parola entrata per via lessicografica e poi rimasta senza una reale ragione nel repertorio.

## 2.2 Il caso di *quarantena*

Una parola il cui uso è aumentato sensibilmente dal 2020 in poi, in ovvio rapporto con la pandemia globale Covid19, è *quarantena*, il che vale per forme simili in varie lingue. Una vera e propria impennata delle occorrenze di *quarantena* in italiano si conferma ricorrendo ad una risorsa per la ricerche sul lessico giornalistico come *l’Archivio la Repubblica*. Così alle 1,967 occorrenze nel periodo compreso tra 1984 e inizio 2020 ne corrispondono ben 14,628 tra inizio 2020 e fine 2021.<sup>18</sup> Secondo studi recenti, una breve nota commento di Tomasin del 2020 cui fa seguito l’articolo detagliato di Parenti/Tomasin 2021,

<sup>17</sup> L’informazione, per la quale ringraziamo Gloria Clavería Nadal, si ricava dal DECH, che ipotizza qualche spiegazione ulteriore: «de él pasó a Minsheu y a Aut., que lo declara sin uso. Covarr. explica que lo habian traído de Italia los cortesanos o los soldados. Para el origen de la voz italiana, controvertido, vid. Skok, ARom. V, 252-8; Gamillscheg, R.G. II, s. v.; Rohlf, ASNSL CLXVIII, 257. Forma abbreviada y usada realmente en España, pero sólo en germanía, es *fazo*, registrado por Juan Hidalgo (1609)».

<sup>18</sup> Simile l’aumento per forme analoghe come ted. *Quarantäne*, facilmente verificabile nei dizionari basati su corpora che offrono la funzionalità di visualizzare statistiche di frequenza (come DWDS, risorsa lessicografica per il vocabolario del tedesco).

sarebbe da escludere la pretesa origine veneziana del significato relativo a regolamenti sanitari<sup>19</sup>, pur essendo stato coniato in ambito settentrionale nel sec. XVI e diffuso nel secolo successivo:

- (1) “Nell’accezione sanitaria che ci interessa, il termine *quarantena* si è diffuso – verosimilmente a partire dall’italiano, nel corso del secolo XVII – in tutte le principali lingue europee, comprese quelle (come l’inglese, *quarantine*, o il tedesco, *Quarantäne*) nelle quali il legame con la parola che significa quaranta è ovviamente venuto meno.” (Tomasin 2020)

La prima attestazione di questa accezione, 1630 secondo DELI, andrebbe retrodatata di vari decenni come dimostrano Parenti e Tomasin riassumendo l’intricata trafila etimologica del termine sanitario. Essa parte da forme del numerale collettivo *quarantena* (basate su forme latine contemporanee) già in uso per designare un intervallo di quaranta giorni (con evidenti rapporti al periodo canonico di purificazione e digiuno nella cultura giudaico-cristiana, cfr. 2021: 27, 32) per acquisire solo più tardi, con la pestilenza degli anni ’70 del Cinquecento, il significato di ‘(periodo di) isolamento sanitario’, nato con ogni probabilità in ambito settentrionale:

- (2) “possiamo affermare che la voce può ben essersi formata nell’Italia settentrionale, perché anche in quest’area, come in Francia, erano in uso voci affini che possono essere chiamate in causa come premessa: anzitutto, come base di partenza, il numerale collettivo *quarantena* ‘*quarantina*’” (2021: 33)

L’etimologia di *quarantena* nell’accezione sanitaria è alquanto discussa anche nella lessicografia delle singole lingue in cui esiste il termine. Per l’inglese *quarantine* OED propone una trafila etimologica che riferisce la forma ad una variante it. *quarantina* (accanto alla forma desueta *quarentina*), con rimando esplicito ad un *quarentena* di origine regionale con la precisazione “Venice” (Venezia):

- (3) “In sense 4 probably < Italian *quarantina*, †*quarentina* (1630 in this sense, originally in the regional (Venice) form *quarentena*; a1311 denoting a set of forty (with reference to units of

<sup>19</sup> Il termine sinonimo veneziano è invece, come dimostrano i due studiosi con fior di esempi, *contumacia*, in uso corrente nei lazzeretti verosimilmente dalla seconda metà del sec. XVI (cfr. Parenti/Tomasin 2021: 24s.).

time), 14th cent. denoting a period of forty days, originally specifically one set aside for penance; see below).” (OED, s.v.)

Il dizionario etimologico di Pfeifer (online, s.v.) menziona l’origine italiana delle varianti tedesche *Quarantena*, *Quarantia* (al plurale *Quarentennas*, *Quaranten*) che risalirebbero al primo Seicento mentre si sarebbe poi affermato, nel corso del Settecento, la forma francesizzante *Quarantaine* (oggi *Quarantäne*), soppiantando le forme di origine italiana. Ovviamente una tale ricostruzione etimologica si avvale di testimonianze scritte che la fanno sembrare l’unica plausibile. In questa prospettiva ted. *Quarantäne* sarebbe da considerare un italianismo indiretto, inoltrato, come avviene spesso nella storia del lessico tedesco, dal francese come lingua di mediazione, e così sarebbe da trattare anche dai dizionari dei prestiti.

In una risorsa lessicografica come l’OIM, che mira ad arricchire la documentazione degli italianismi lessicali, subentrano però funzionalità supplementari per dare conto del fatto che i fenomeni di contatto linguistico si manifestano, oltre le dimensioni del lessico e della semantica che interessano i prodotti stessi del contatto (prestiti), a più livelli sistematici della lingua (fonetica/morfosintassi/pragmatica). Infatti, al fine di offrire un quadro il più possibile completo del suo percorso diacronico, di ciascuna voce viene indicato, come parte integrante del corredo di informazioni microstrutturali, la trascrizione fonetica (con eventuali varianti foniche) e un file audio della pronuncia nella lingua di arrivo. In tedesco si notano due varianti fonetiche di *Quarantäne*, circoscritte nella loro diffusione diatopica. Da una parte esiste la variante principale [karan'te:nə] basata su una pronuncia francesizzante della sillaba iniziale di parola che si rifa a fr. [karã'ten], diffusa in gran parte della Germania e percepita da molti parlanti come variante standard, mentre nell’area tedescofona meridionale, soprattutto nel dominio delle varietà bavaresi (sia della Germania che dell’Austria), è molto presente anche l’altra variante che invece dell’occlusiva velare sorda [k] inizia per il nesso della velare con la fricativa labiodentale sonora, [kvaran'te:nə]. Nel maggiore dizionario di riferimento del tedesco moderno, DUDEN (s.v.), le varianti di pronuncia figurano di fianco, la seconda etichettata come “seltener” (‘più rara’) senza men-

zione di una distribuzione diatopica.<sup>20</sup> Una differenziazione areale che vale anche per le pronunce colte, con frequenza decisamente alta anche della seconda variante, specie nei parlanti oltre i 30 anni di età, è invece suggerita dalle inchieste recenti di Meier-Vieracker (2020) per le varietà della Germania e di Soukup (2021) per l'area austriaca. Essendo quest'ultima area linguistica anche in altri casi zona privilegiata del contatto con l'Italoromania, a differenza del resto della Germania<sup>21</sup>, è altamente verosimile che qui la differenza di pronuncia non sia un effetto secondario (dovuto per es. alla grafia) bensì conseguenza delle diverse forme di partenza. Alla luce delle considerazioni sopraccitate di Parenti/Tomasin (2021) sono quindi da aggiornare le trafile etimologiche di ted. *Quarantäne*, ricostruite finora solo parzialmente dalla lessicografia tedesca:

- (4) ted. standard [kaʀan'tɛ:nə] < fr. *quarantaine* (forse con precedenza dell'it. *quarantena*)
- (5) ed. austriaco: [kvaʀan'tɛ:nə] < it. *quarantena* (forse con influsso a posteriori del fr. *quarantaine*)

I singoli fattori diatopici e sociolinguistici di una tale distinzione abbisognano senz'altro di futuri accertamenti attraverso i (rari) corpora di parlato con dettagliata annotazione sociofonetica. Prendendo in considerazione anche la fonetica degli adattamenti formali dei lemmi di partenza la banca dati OIM, che attinge a fonti lessicografiche, divulgative e corpora, si prefigge di documentare nelle sue schede i tratti sistematici degli italianismi nelle lingue di arrivo così da gettare le basi per ulteriori ricerche.

### 2.3 Il caso di *gabbione*

Per mostrare invece come può esser sfruttato un corpus di testi recenti per introdurre termini non ancora registrati dai dizionari, possiamo portare il caso di *gabbione*. In italiano la parola ha assunto il significa-

<sup>20</sup> DUDEN ne indica una terza, ancor più vicina a quella francese e ugualmente marcata 'più rara', con la vocale nasale [kaʀā'tɛnə], della quale però non si ha traccia in nessuno degli studi attuali. Sono inoltre pensabili realizzazioni variabili della vocale tonica (semiaperta o semichiusa).

<sup>21</sup> Cfr. ad es. la varietà di verdura chiamata *Melanzani* (< it. *melanzane*) nelle varietà dialettali di stampo bavarese-austriaco e *Aubergine* (< fr. *aubergine*) nelle altre varietà tedesche (inclusa quella standard; cfr. DWDS, s.vv.).

to specialistico nella terminologia militare di “elemento di fortificazione, costituito da un grosso cesto di vimini riempito di terra e sassi, usato per la costruzione di trincee e parapetti” e in quella idraulica di “elemento di difesa idraulica collocato a riparo di argini, ponti, scarpate fluviali e sim., costituito da un contenitore di rete metallica riempito di ciottoli e pietre” (GRADIT). Nel linguaggio recente il termine viene usato anche per designare la “gabbia in cui prendono posto gli imputati durante alcuni processi in tribunale” (ancora GRADIT; quest’ultima accezione non è registrata in Zingarelli 2022, nonostante sia l’uso più frequente nel linguaggio giornalistico: consultando l’archivio di Repubblica dal 1984 al 2021 si ottengono 154 occorrenze, prevalentemente concentrate in questo uso). Come prestito, però, la parola si è specializzata nella direzione della struttura portante riempita di sassi e ghiaia, usata con una certa frequenza specie per l’arredo dei giardini. In tedesco l’italianismo *gabbione* non è registrato fino a tempi recenti: non se ne trova traccia nel DWDS (che raccoglie testi fino al 1999); nel DIFIT era censito nella forma *Gabion* come prestito indiretto dal francese, dove aveva diversi significati (1. Fortificazione; 2. (region), Grande cesto con manici per il trasporto del letame, della terra. 3. sport / equit. Casotto per i cacciatori di selvaggina acquatica). In tempi recenti compare, come italianismo, in DUDEN, nella forma *Gabione*. La conferma che la parola ha acquisito una nuova vitalità, oltre che una nuova forma grafica, arriva proprio dal controllo su Sketch engine, che per il corpus tedesco (German web 2018, costituito da 5,346,041,196 parole) raccoglie oltre duemila occorrenze del prestito (343 sing. *Gabione* e 1758 pl. *Gabionen*), in gran parte concentrate in testi legati al giardinaggio.

Il successo di questa struttura e la sua graduale espansione anche al di fuori della Germania potrebbero far supporre un analogo radicamento anche in lingue di paesi vicini: *Gabion* ha oltre 7000 occorrenze nel corpus inglese (English Web 2020: 5004 sing., 2013 pl.); troviamo inoltre 1262 *gabion* in francese (French Web 2017), 3077 *gavión* in spagnolo (Spanish Web 2018), 677 *gabião* in portoghese (Portuguese Web 2011); 190 *gabion* in ungherese (Hungarian Web 2012). In queste ultime due lingue il numero più basso potrebbe essere collegato al fatto che i corpora raccolgono testi dei primi anni Dieci del XXI secolo, quando forse questo elemento di arredo non era ancora così diffuso.

Nell'OIM, per l'ungherese, la parola viene segnalata come prestito indiretto, mediato dal francese e dal tedesco, e se ne registra, accanto al significato abituale di fortificazione militare (indicato come neologismo nel 2010), anche con il significato più recente di elemento costruttivo usato specie nei recinti di giardini: la segnalazione arriva proprio per il tramite di fonti "estemporanee" (descrizioni storiche e – ultimamente – numerose occorrenze su Internet (p.es. <https://www.pannongabion.hu/gyik>) e di siti di venditori di materiali da costruzione)<sup>22</sup>: si potrebbe pensare a una nuova parola internazionale, utile a descrivere un'evoluzione moderna, forse un po' meno artistica, del tipico giardino all'italiana che decorava le ville rinascimentali.

### *Riferimenti bibliografici*

- Bagna, Carla & Barni, Monica. 2006. Per una mappatura dei repertori linguistici urbani: nuovi strumenti e metodologie. In De Blasi, Nicola & Marcato, Carla (a cura di), *La città e le sue lingue. Repertori linguistici urbani*. Napoli: Liguori. 1–43.
- Bellinzona, Martina. 2021. *Linguistic landscape. Panorami urbani e scolastici nel XXI secolo*. Milano: FrancoAngeli.
- Bombi, Raffaella & Orioles Vincenzo (a cura di). 2015. *Italiani nel mondo. Una Expo permanente della lingua e della cucina italiana*. Udine: Forum.
- Bonomi, Ilaria & Coletti, Vittorio. 2015. *L'italiano della musica nel mondo*. Firenze: Accademia della Crusca.
- Brincat, Giuseppe. 1991. Spigolando nel "Sunday Times": parole italiane e pseudoitalianismi nell'inglese di oggi. In Coveri, Lorenzo (a cura di), *L'italiano allo specchio. Aspetti dell'italianismo recente. Saggi di Linguistica Italiana. Atti del Primo Convegno della Società Internazionale di Linguistica e Filologia Italiana (Siena 28-31 marzo 1989)*. Torino: Rosenberg & Sellier. 7–14.
- Casini, Simone. 2015. Italianismi e pseudoitalianismi nel mondo globale: il ruolo dell'enogastronomia. In Bombi, Raffaella & Orioles, Vincenzo (a cura di), *Italiani nel mondo. Una Expo permanente della lingua e della cucina italiana*. Udine: Forum. 89–102.

<sup>22</sup> L'informazione, raccolta da Zsuzsanna Fábán, è registrata nell'OIM (<https://www.italianismi.org/scheda-italiano/gabbione/1888>).

- Casini, Simone. 2017. Italianismi e pseudoitalianismi a Toronto: una ricerca tra gli studenti di italiano del St. George Campus della University of Toronto. *Italica*. 94. 153–176.
- Casini, Simone. 2018. Italianismi e pseudoitalianismi a Toronto: tra valori simbolici e prospettive di apprendimento. In Turchetta, Barbara & Vedovelli, Massimo (a cura di), *Lo spazio linguistico italiano globale: il caso dell'Ontario*. Pisa: Pacini. 225–254.
- Cassani, Vanina. 2015. Italianismi e pseudoitalianismi a Londra: l'italiano tra vie tradizionali e innovative di diffusione. *Rivista Studi Italiani di Linguistica Teorica e Applicata* XLIV (1). 361–385.
- Cortelazzo, Michele. 2022. Retrodatazioni al “DELI” da traduzioni letterarie ottocentesche. *Studi di lessicografia italiana*. XXXIX, 247–312.
- Covarrubias Orozco, Sebastián de. 1987. *Tesoro de la lengua castellana o española*, ed. de Martín de Riquer. Barcelona: Alta Fulla (rist. anast. dell'edizione del 1611).
- Dashi, Brunilda. 2013. *Gli italianismi nella lingua albanese*. Roma: Nuova cultura.
- DECH. 1980-1991. Coromines, Joan & Pascual, José Antonio. *Diccionario crítico etimológico castellano e hispánico*. Madrid: Gredos (edición en CD-ROM).
- Della Valle, Valeria & Patota, Giuseppe. 2013. Residui passivi. Storie di archeologismi. *Studi di lessicografia italiana*. XXX. 133–164.
- DIFIT. Harro Stammerjohann et al. (a cura di). 2008. *Dizionario di italianismi in francese, inglese, tedesco*. Firenze: Accademia della Crusca.
- DLE. 2014<sup>23</sup> (y actualización). Real Academia Española. *Diccionario de la lengua española*. Madrid: Espasa-Calpe (<https://dle.rae.es>).
- DRAE. 2003. *Diccionario de la lengua española*. 22ª ed. Madrid: Espasa-Calpe.
- DUDEN. 1999. *Das große Wörterbuch der deutschen Sprache*, herausgegeben von Scholze-Stubenrecht, Werner unter Mitarbeit von Alsleben, Brigitte. Mannheim: Dudenverlag (<https://www.duden.de/>).
- DWDS. *Digitales Wörterbuch der deutschen Sprache, Das Wortauskunftssystem zur deutschen Sprache in Geschichte und Gegenwart*, hrsg. v. d. Berlin-Brandenburgischen Akademie der Wissenschaften (<https://www.dwds.de>).
- Fábíán, Zsuzsanna & Szabó, Gyözö. 2010. *Dall'Italia all'Ungheria: parole di origine italiana nella lingua ungherese*. Udine: Forum Editrice.
- Ferrini, Caterina. 2018. Italianismi e pseudoitalianismi nelle Little Italy di Toronto: il linguistic landscape come termometro per misurare la “febbre

- da italiano". In Turchetta, Barbara & Vedovelli, Massimo (a cura di), *Lo spazio linguistico italiano globale: il caso dell'Ontario*. Pisa: Pacini. 255–292.
- GDLI. 1961-2002. *Grande dizionario della lingua italiana*, fondato da Salvatore Battaglia. Torino: Utet.
- Gomez Gane, Yorick. 2012. *Gli italianismi nel catalano. Dizionario storico-etimologico*. Roma: Aracne.
- Gorter, Durk (ed). 2006. *Linguistic Landscape: A New Approach to Multilingualism*. Clevedon-Buffalo-Toronto: Multilingual Matters.
- GRADIT. 1999-2003. *Grande dizionario italiano dell'uso*, diretto da Tullio De Mauro. Torino: Utet.
- Heinz, Matthias (a cura di). 2017. *Osservatorio degli italianismi nel mondo: punti di partenza e nuovi orizzonti. Atti dell'incontro (Firenze, 20 giugno 2014)*. Firenze: Accademia della Crusca.
- Ille, Karl. 2009. Italianismen und Pseudoitalianismen in der gastronomischen und kommerziellen Öffentlichkeit Wiens. In Ehmer, Josef & Ille, Karl (Hg.). *Italienische Anteile am multikulturellen Wien*. Innsbruck-Wien-Bozen: Studienverlag (Querschnitte, 27). 111–125.
- Landry, Rodrigue & Bourhis, Richard Y. 1997. Linguistic landscape and ethnolinguistic vitality: an empirical study. *Journal of Language and Social Psychology* 16(1). 23–49.
- Matt, Luigi. 2004. Retrodatazioni di tecnicismi da titoli di pubblicazioni. *Studi di lessicografia italiana* XXI. 183–246.
- Mattarucco, Giada (a cura di). 2012. *Italiano per il mondo. Banca, commerci, cultura, arti, tradizioni*. Firenze: Accademia della Crusca.
- Meier-Vieracker, Simon. 2020. Wie spricht man "Quarantäne" aus? Ergebnisse einer Umfrage. *Linguistische Werkstattberichte*. <https://lingdrafts.hypotheses.org/1539>
- Mori, Laura, 2021. Italian Heritage come categoria interpretativa dell'italianità linguistico-culturale per l'analisi della comunicazione commerciale nel mercato economico internazionale. *Cultura&Comunicazione* XI (19). 36–48.
- NLTLE. 2001. *Nuevo Tesoro Lexicográfico de la Lengua Española*, publicado por la Real Academia Española, 2 DVD. Madrid: Espasa Calpe.
- OED. 1989. *The Oxford English Dictionary*. by John A. Simpson & Edmund Weiner, 20 voll. Oxford: Clarendon press.
- OIM e DIFIT: [www.italianismi.org](http://www.italianismi.org)
- Parenti, Alessandro & Tomasin, Lorenzo. 2021. Su Quarantena, preteso venezianismo, e su Contumacia. *Lingua nostra*. LXXXII/1-2. 23–33.

- Pfeifer, Wolfgang (dir.). 1993<sup>2</sup>. *Etymologisches Wörterbuch des Deutschen*. Berlin: Akademie-Verlag (online: <https://www.dwds.de/d/wb-etymwb>).
- Pinnavaia, Laura. 2001. *The Italian Borrowings in the Oxford English Dictionary: A Lexicographical, Linguistic and Cultural Analysis*. Roma: Bulzoni.
- Pizzoli, Lucilla. 2017. Per un dizionario degli italianismi nel mondo: rilancio di un progetto. *Testi e linguaggi* 11. 171–182.
- Pizzoli, Lucilla. 2019. Italiano e italianismi nel mondo: osservazioni sulla ricerca di neologismi. In Bombi, Raffaella (a cura di), *Italiano nel mondo. Per una nuova visione*. Udine: Forum editrice. 151–158.
- Repubblica Archivio* (1/09/2021).  
<https://ricerca.repubblica.it/repubblica/archivio/repubblica>
- Rieger, Marie A. 2008. ‘Alles picco belli oder was?’ Form und Funktion pseudo-italienischer Produktnamen im deutschen Lebensmittelmarkt. *Onoma*. 43. 149–175.
- Serianni, Luca. 2008. 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-Paris: Fondazione Cassamarca-Unione Latina. 19–41.
- Siebetcheu, Raymond. 2015. La lingua italiana nei panorami linguistici urbani delle città camerunensi. In Pallante, Gianna & Kuitche Talé, Gilles (a cura di), *20 anni d’insegnamento dell’italiano L2 in Camerun: bilancio e prospettive*. Italiano LinguaDue. 2. 59–70.
- Soukup, Barbara. 2021. Geht Österreich in ‘Karantäne’ oder ‘Kwarantäne’? Ergebnisse einer Umfrage. *Wiener Linguistische Gazette* 90. 265–307.
- Telve, Stefano. 2002. Retrodatazioni di voci onomatopoeiche e interietive. Un esempio di applicazione lessicografica degli archivi elettronici. *Studi di lessicografia italiana* XIX. 229–277.
- Tomasin, Lorenzo. 2020. Una *quarantena* può durare anche “solo” quattordici giorni. *Italiano digitale* XII/1.  
(online: <https://id.academiadellacrusca.org>)
- Vedovelli, Massimo & Casini, Simone. 2013. Italianismi e pseudoitalianismi in Giappone: le radici profonde di una consonanza culturale nel mondo globale. In Gesuato, Maria Katia & Peruzzi, Paola (a cura di), *La lingua italiana in Giappone 2*. Tokyo: Istituto italiano di cultura. 34–106.
- Vedovelli, Massimo & Machetti, Sabrina. 2006. Italiano e lingue esotiche in contatto nella comunicazione sociale: il caso degli italianismi a Tokyo. In Banfi, Emanuele & Iannàcaro, Gabriele (a cura di), *Spazio linguistico italiano e le lingue esotiche: rapporti e reciproci influssi (Atti del XXXIX*

- Congresso internazionale di studi della Società di linguistica italiana – SLI: Milano, 22-24 settembre 2005*). Roma: Bulzoni. 181–206.
- Vedovelli, Massimo. 2005. L'italiano nel mondo da lingua straniera a lingua identitaria: il caso 'freddoccino'. *Studi Italiani di Linguistica Teorica e Applicata* (XXXIV/3). 585–609.
- Vedovelli, Massimo. 2014. L'italiano nel mercato globale delle lingue: prospettive, potenzialità, criticità. *Rapporto Italiani nel Mondo 2014*. 289–297.
- Vespaziani, Alberto. 2018. Comunicazione Sociale. *Il Libro dell'Anno del Diritto Treccani*. 282–284.
- Vidos, Benedek E. 1965. *Prestito, espansione e migrazione dei termini tecnici nelle lingue romanze e non romanze. Problemi, metodo e risultati*. Firenze: Olschki.
- Wilhelm, Eva-Maria. 2013. *Italianismen des Handels im Deutschen und Französischen*. Berlin/Boston: De Gruyter.
- Zingarelli 2022. *Vocabolario della lingua italiana*. Bologna: Zanichelli.