

## IL PROGETTO OIM (OSSERVATORIO DEGLI ITALIANISMI NEL MONDO)\*

Lucilla Pizzoli<sup>1</sup>, Matthias Heinz<sup>2</sup>

**1.** L'OIM (Osservatorio degli Italianismi nel Mondo), banca dati mirata a raccogliere le parole italiane e di origine italiana entrate nell'uso di altre lingue, è stato accolto a partire dal 2014 tra i progetti strategici dell'Accademia della Crusca all'interno del portale Vivit. Fin dalla sua costituzione il progetto è stato affidato alla direzione di Luca Serianni e Matthias Heinz, che da anni hanno indirizzato i loro sforzi, insieme a un vasto gruppo di ricerca internazionale, per creare un punto di riferimento per la raccolta, verifica e condivisione di italianismi diffusi in diverse lingue del mondo<sup>3</sup>. L'improvvisa scomparsa di Luca Serianni rappresenta un'enorme perdita anche per questo progetto: tutti i ricercatori impegnati nell'OIM hanno partecipato al grande dolore della comunità scientifica ribadendo il loro impegno nel voler proseguire il lavoro. È dunque alla memoria di Luca Serianni che dedichiamo questa sezione, convinti che portare avanti l'Osservatorio sia il modo migliore per ricordare la figura di un grande maestro, che, tra tante imprese meritorie, ha dato un impulso decisivo anche al progetto OIM, indirizzandolo con competenza sicura e seguendone con attenzione gli sviluppi<sup>4</sup>.

Negli anni il progetto si è allargato attraverso confronti diretti tra studiosi provenienti da diverse aree del mondo ed è grazie a questi numerosi colloqui che si è riusciti a risolvere punti critici e trovare le soluzioni migliori per le difficoltà che si sono via via presentate<sup>5</sup>.

L'ultimo colloquio al quale Serianni ha partecipato è stato quello del 2021, tenutosi presso la sede dell'Accademia tra il 14 e il 15 ottobre in collegamento online con molte sedi, ma finalmente, dopo l'emergenza pandemica, in forma prevalentemente presenziale: in quell'occasione è stato possibile, con grande soddisfazione dell'intero gruppo, dare l'annuncio della prima pubblicazione *online* del sito OIM<sup>6</sup> e Serianni ha aperto i lavori esprimendo il suo apprezzamento per i risultati raggiunti.

L'articolo che presentiamo in questa sede intende dare conto dei risultati raggiunti finora grazie al sostegno generoso dei ricercatori impegnati nel progetto: il grande entusiasmo che accomuna gli studiosi attivi in OIM è l'elemento che spiega il fatto che si riesca a portare avanti un disegno tanto ambizioso con una gittata temporale

\*Sono da attribuirsi a Lucilla Pizzoli i paragrafi 1 e 3 e a Matthias Heinz il paragrafo 2.

<sup>1</sup> Università degli studi internazionali di Roma.

<sup>2</sup> Universität Salzburg.

<sup>3</sup> Diretto da Serianni e Heinz il progetto viene, sempre dal 2014, coordinato scientificamente da Lucilla Pizzoli mentre la gestione della piattaforma informatica è responsabilità di Marco Biffi e Giovanni Salucci.

<sup>4</sup> Per valutare l'apporto decisivo di Serianni al progetto basti rileggere il suo contributo al primo colloquio fiorentino dell'OIM (del giugno 2014: cfr. Serianni, 2017). Tra gli altri contributi dedicati al tema vale la pena ricordare anche Serianni (2008). Sul tema dell'italiano nel mondo, e del parallelo ruolo delle lingue altre in Italia, cfr. Serianni (2019).

<sup>5</sup> Si ricordano di seguito i colloqui OIM che si sono susseguiti dopo il 2014: Firenze (4-6 settembre 2017), Salisburgo (28-29 giugno 2018), Cosenza (5-7 novembre 2018), Firenze (4-5 aprile 2019), Atene (1 novembre 2019), Firenze (14-15 ottobre 2021), Firenze (14-16 settembre 2022).

<sup>6</sup> OIM: [www.italianismi.org](http://www.italianismi.org).

evidentemente molto lunga. Lo stesso Serianni, nel sottolineare il suo compiacimento per quanto ottenuto in questi anni, aveva rimarcato, rallegrandosene in modo particolare, lo spirito che tiene insieme il gruppo e aveva ringraziato calorosamente tutti coloro che avevano garantito gli avanzamenti<sup>7</sup>.

È dunque con particolare soddisfazione che diamo conto del lavoro OIM in questa sede, che ha ospitato proprio nel suo primo numero (1, 1, 2009) la presentazione del DIFIT<sup>8</sup> e che nel tempo ha seguito le ricerche nate in quest'ambito: il tema degli italianismi infatti è molto vicino alla sensibilità del comitato scientifico di "Italiano LinguaDue", che negli anni ha raccolto moltissimi contributi in questo settore di ricerca. A dimostrazione di questa particolare attenzione si deve infine, di recente, l'allargamento del comitato scientifico della rivista a Matthias Heinz e Lucilla Pizzoli, che si aggiungono al nutrito gruppo di esperti orientati in questa direzione.

Nei prossimi numeri, approfittando della presenza, all'interno della rivista, della sezione *Italoфонie* curata da Franco Pierno e che ospita anche questi primi contributi, si potrà descrivere più nel dettaglio lo stato di avanzamento delle singole raccolte.

## 2. COM'È NATO IL PROGETTO DI UN OSSERVATORIO DEGLI ITALIANISMI?

Progetti collaborativi come questo, caratterizzato da una certa complessità sia nell'intento scientifico che si prefigge che nella sua articolazione interna, richiedono una divisione del lavoro e un coinvolgimento di esperti nelle singole aree linguistiche di interesse, e tendono di più in più a ramificarsi. In quanto segue ripercorriamo brevemente la storia del progetto che ormai vede coinvolte almeno 15 unità di lavoro (per lo più composte da piccoli gruppi di studiosi e studiose delle aree in questione) operanti in una ventina di atenei italiani ed esteri tra Austria, Finlandia, Grecia, Macedonia, Malta, Polonia, Portogallo, Slovacchia, Spagna, Svezia, Ungheria, Canada e USA.

Com'è ben noto, la ricerca sugli italianismi si occupa dei lemmi di origine italiana presi in prestito da altre lingue a seguito di un contatto culturale e linguistico spontaneo, sporadico o duraturo, mettendo a fuoco i risultati lessicali di tale contatto. Le varietà che entrano in gioco possono essere più o meno limitrofe come il francese, il catalano, lo spagnolo, il tedesco e l'inglese oppure più lontane, in forma standard o dialettale, più o meno 'esotiche' da un punto di vista eurocentrico. Mentre i conseguenti prestiti sono registrati in un folto numero di studi e in una certa misura nei dizionari monolingui, pochi lavori hanno finora adottato una prospettiva più ampia, studiando le forme provenienti dal contatto con il patrimonio culturale dell'italiano in parallelo, comparandone le vicende in tante lingue. Gli unici esempi recenti di un sistematico approccio multilingue al contatto linguistico nel lessico sono il *Dictionary of European Anglicisms* (Görlach, 2001), il *Lehnwortportal* tedesco (progetto sui prestiti dal tedesco nelle sue storiche lingue di contatto in corso di realizzazione presso l'Institut für deutsche Sprache/IDS, Engelberg/Mayer),

<sup>7</sup> Va ricordato che il progetto OIM si avvale dell'appoggio istituzionale e del sostegno finanziario dell'Accademia della Crusca. Un impulso significativo all'ampliamento delle ricerche è stato garantito da un finanziamento ARCUS concesso dal Mibact nel 2017. Hanno contribuito all'avanzamento dei lavori l'Università di Salisburgo (2018), l'Università della Calabria (2018) e l'UNINT – Università degli studi internazionali (2020 e 2021). La parte principale del lavoro però è sostenuta dall'impegno personale dei singoli ricercatori, che dedicano molte delle loro energie a raccogliere dati e a renderli compatibili con i criteri usati nella piattaforma dell'OIM.

<sup>8</sup> <https://riviste.unimi.it/index.php/promoitals/issue/view/42>.

il progetto neerlandese *Nederlandse woorden wereldwijd* e *Uitleenwoordenbank* (Sijs, 2010, 2015) e, per l'area italomanziana, l'OIM che prende spunto da due imprese precedenti.

Per identificare retrospettivamente il punto di partenza bisogna risalire, da un lato, alla pubblicazione, nel 2008 presso le edizioni dell'Accademia della Crusca, del *Dizionario degli italianismi in francese, inglese, tedesco* (DIFIT), concepito e redatto da H. Stammerjohann insieme a un gruppo internazionale di studiosi responsabili per le tre sezioni linguistiche. Dall'altro va menzionato il progetto editoriale quasi parallelo di un *Censimento* di ampio respiro, volto a repertoriare italianismi in oltre 70 lingue del mondo intero, a cura di Luca Serianni per l'editore UTET. Gli autori del presente contributo sono stati già collaboratori dei rispettivi progetti.

Il DIFIT con la sua documentazione cartacea degli italianismi in tre lingue europee è poi stato retrodigitalizzato al fine di renderlo interrogabile in forma di banca dati: la prima versione è disponibile come parte integrante del portale VIVIT dal 2013/14, poi sul sito creato specificamente per l'*Osservatorio degli italianismi nel mondo*, per cui si viene a stabilire la sigla OIM<sup>9</sup>. Il progetto del *Censimento* invece non uscirà in forma di opera a stampa per ragioni di cambiamento di politica editoriale, benché ne risultino singole pubblicazioni di raccolte di italianismi come ad es. Gomez Gane (2012) per il catalano<sup>10</sup>. Tuttavia rimane dapprima inesplorata la ricchezza delle raccolte portate a termine, elaborate secondo criteri ben definiti, seppur carenti a quel momento di un supporto idoneo per renderli fruibili da parte degli specialisti o addirittura di un pubblico comune. L'idea di un osservatorio degli italianismi nelle lingue del mondo che possa offrire un tale modo di fruizione mediante il supporto informatico, nata già nell'ambito della digitalizzazione dei dati del DIFIT, prende poi avvio nel 2014<sup>11</sup>. I dati che stanno alla base della nuova piattaforma provengono dal DIFIT e dai risultati delle raccolte per il *Censimento* e confluiscono, appurati ed elaborati, nella banca dati. Visualizzati tramite l'interfaccia per gli utenti forniscono alla comunità scientifica uno strumento di lessicografia digitale per indagare le vicende dei prestiti passati dall'italiano in altre lingue.

All'interno del gruppo di ricerca le unità di lavoro si prendono cura della raccolta stessa dei materiali, della verifica in base agli strumenti lessicografici e *corpora* di riferimento per le rispettive lingue e dell'immissione dei dati così ottenuti, affiancati dal comitato scientifico e da un gruppo di informatici specializzati nell'umanistica digitale. La crescita del gruppo di lavoro ha prodotto anche germogli al di fuori dell'Europa come il gruppo OIM Nordamerica (con la sottounità OIM Canada con sede all'Università di Toronto), da anni attiva nell'indagine del contatto tra l'italiano e le varietà nordamericane dell'inglese, francese e spagnolo. Nella prospettiva di procedere *grasso modo* a cerchi concentrici si è passati, a partire dal 2016/17, dalle tre lingue censite dal DIFIT, idiomi di una certa centralità in Europa, alle raccolte di italianismi in spagnolo, portoghese, catalano, polacco e ungherese (le lingue iberiche entrambe nella loro varietà europea). L'apporto delle raccolte già iniziate nell'ambito del *Censimento* e la felice cooperazione tra collaboratori attivi già allora e altri specialisti competenti nella rispettiva area di contatto linguistico sono stati fattori decisivi per ottenere, nel giro di pochi anni, delle documentazioni complete o pressoché complete e adattate ai parametri della piattaforma informatica.

<sup>9</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)).

<sup>10</sup> Per la ricostruzione delle potenzialità e gli effettivi risultati di tale impresa cfr. Pizzoli (2017) e Serianni (2017).

<sup>11</sup> Cfr. Heinz (2017) per una prima rassegna degli spunti e traguardi di un tale osservatorio.

In una terza fase progettuale, dal 2019, sono iniziati e in parte già avanzati i lavori di raccolta e inserimento di italianismi in maltese, neogreco, macedone, slovacco, finlandese, svedese, cinese (mandarino) e altre lingue in parallelo alla revisione ed estensione dei dati sulle lingue francese, tedesco, inglese, prendendone in considerazione varietà anche al di là dei confini nazionali<sup>12</sup>. Lo scambio sulle sfide e i progressi di tale documentazione collettiva avviene grazie alla serie assai regolare di colloqui e seminari menzionata nella sezione iniziale (cfr. n. 5). Questi incontri, che si svolgono per lo più nella sede dell'Accademia e talvolta in altre sedi e/o in modalità virtuale, sono destinati in particolare ai membri del gruppo e a chi si accinge a formare una nuova unità in seno al progetto, e comunque si rivolgono spesso anche ad un pubblico più largo per dare conto dei progressi del progetto.

Attualmente sono registrate nella piattaforma oltre 12.600 voci di cui circa 1.500 categorizzate come 'complete'; quelle contrassegnate come 'in stato di lavorazione' per essere vagliate richiedono un ultimo passaggio di verifica. Nel complesso più di 8.900 voci provengono dalle raccolte di italianismi del DIFIT, pur essendo in continuo aumento il numero di quelle provenienti da altre lingue. Le schede relative ai nuovi lemmi vengono gestite dalla redazione italiana che in un *workflow* embricato ne verifica le trafile lessicali e gli sviluppi semantici al fine di inserirle nella piattaforma.

### **3. A CHE PUNTO SIAMO: LA PIATTAFORMA OIM E GLI SVILUPPI DELLA RICERCA SUGLI ITALIANISMI**

#### **3.1. *La piattaforma informatica del progetto OIM***

Negli ultimi contributi pubblicati dedicati al progetto OIM abbiamo potuto documentare alcuni dei risultati raggiunti e molti degli sviluppi previsti: la storia dell'OIM a partire dal 2014 è già stata ricostruita in questa sede da Matthias Heinz e molti dei problemi metodologici e delle potenzialità della piattaforma sono stati anticipati in altri contributi<sup>13</sup>.

Varrà la pena, dunque, presentare qui lo stato della ricerca a oggi, dando conto delle scelte operate per la raccolta dei dati e per la loro classificazione: pur avendo progettato la piattaforma usando criteri lessicografici adeguati ad una visualizzazione informatica di nuova concezione, è stato inevitabile scontrarsi con ulteriori difficoltà solo al momento dell'inserimento dei lemmi. La soluzione dei problemi (emersi soprattutto via via che venivano considerate lingue con strutture diverse) è arrivata, come si è detto, dal confronto serrato con il gruppo di lavoro e dalla verifica puntuale con gli informatici.

Va precisato infatti che la prima sperimentazione ha interessato un gruppo molto ristretto di lingue: solo aggiungendo dati in modo graduale si è potuto mettere a punto un sistema in grado di accogliere informazioni relative a lingue disomogenee da un punto di vista tipologico e soprattutto di descrizione lessicografica.

Forti dell'esperienza della raccolta avviata in anni ormai lontani per UTET (da Serianni definita a suo tempo fallimentare: "storia di un fallimento" è il sottotitolo scelto – con una certa dose di coraggiosa autocritica – per descrivere il lavoro originario, che aveva l'ambizione di presentare i prestiti registrati in oltre 70 lingue in un dizionario cartaceo:

<sup>12</sup> Le varietà extraeuropee sono di grande importanza nel caso del francese, francese, inglese, spagnolo e portoghese, lingue in cui si trovano molti italianismi che sono frutto del contatto tra le varietà degli italiani emigrati e lingue locali.

<sup>13</sup> Cfr. Biffi (2014), Heinz (2017), Pizzoli (2017, 2019), Heinz, Pizzoli, Serianni (2022), Pizzoli, Heinz (2022).

cfr. Serianni, 2017: 39), abbiamo capito che sarebbe stato impossibile tentare di rendere disponibile nello stesso momento una mole di dati così grande. Ma non è un caso che solo ora, in un'epoca di forte ripensamento dello strumento lessicografico, si sia potuto immaginare un metodo di lavoro completamente diverso.

È un dato di fatto assodato che la rivoluzione informatica abbia rappresentato un'enorme possibilità di sviluppo anche per la lessicografia, sia dal punto di vista della raccolta dei dati che da quello della loro fruizione: con i "sogni" del lessicografo digitale – ormai in gran parte divenuti realtà – si misurava già due decenni fa Gilles-Maurice de Schryver<sup>14</sup>; delle potenzialità rappresentate dalla LIZ per la lessicografia storica, per esempio, scriveva Serianni (2006) isolando tre possibili implicazioni<sup>15</sup>. E anche per la lessicografia sincronica non c'è bisogno di menzionare, in questo contesto, la ricaduta della marcatura di frequenza indicata da De Mauro nel GRADIT e basata per l'appunto su corpora bilanciati di testi; nella stessa linea, ormai, si muovono tutti i grandi progetti lessicografici, come dimostra tra tutti il caso del progetto VoDIM<sup>16</sup>. Riguardo alla rivoluzione della fruizione dei dizionari informatizzati, poi, qualunque utente a cavallo tra le due epoche è capace di misurare la differenza tra i due modi d'uso: se sfogliare un dizionario di carta può dare il pretesto per passeggiate fuori programma con incursioni in lemmi per i quali non era stata preventivata l'esplorazione e fare scattare così nuovi stimoli, l'interrogazione del lemmario tramite le applicazioni digitali (sempre a portata di mano grazie alle versioni studiate per dispositivi mobili) consente di attivare ricerche incrociate e trovare le informazioni in modo completamente nuovo, oltre che di disporre di integrazioni con altri media (audio, video, grafici animati o altre funzioni di rappresentazioni delle voci indagate in connessione con corpora di testi o altri progetti di ricerca: si vadano per esempio le funzioni attivate per il tedesco dal DWDS).

Vedremo più avanti come l'informatizzazione dei dati lessicografici relativi al prestito apra nuove frontiere nella ricerca e consenta di perfezionare la descrizione delle vie del contatto linguistico per l'OIM anche in prospettiva storica.

Partiamo intanto col descrivere le novità nella organizzazione dei dati nella piattaforma dell'osservatorio. Come è facile immaginare, per l'OIM questo enorme sviluppo della lessicografia digitale si è indirizzato nel passaggio da un progetto di tipo tradizionale, da concretizzarsi in un dizionario cartaceo (del tipo di quello che con il DIFIT era stato possibile realizzare felicemente, grazie al fatto che era stato selezionato un numero ridotto, più controllabile di lingue) a una struttura aperta, cioè a un vero e proprio osservatorio dinamico in cui raccogliere i dati in modo continuativo<sup>17</sup>.

<sup>14</sup> Cfr. de Schryver (2003) e gli studi ivi richiamati. Sulla rivoluzione della lessicografia nell'era dell'editoria digitale cfr. Lanfranchi (2014) e le osservazioni, a commento di queste prime riflessioni, condotte da Serianni (2016). Sull'editoria digitale e sul tramonto della cosiddetta "galassia Gutenberg" la bibliografia, non solo italiana, è sterminata (a partire dalla raccolta che ospita il contributo di Lanfranchi, 2014): qui ci si limita a rinviare ai lavori di Roncaglia (2000, 2010 e 2018) e di Marello (2018). Per l'uso dei *corpora* informatici per le retrodatazioni e le possibili insidie cfr. anche Telve (2002), Matt (2004), Maconi (2016), Zarra (2017), Cortelazzo (2022). Un prezioso archivio di datazioni lessicali è costituito dal 2017 presso l'Accademia della Crusca: cfr. Maconi (2020).

<sup>15</sup> Nello specifico, la risorsa informatica avrebbe consentito agli studiosi interessati di individuare 1. Retrodatazioni; 2. Attestazioni prodromiche (cioè «sintagmi non ancora cristallizzati, né formalmente né semanticamente, in una stabile unità polirematica o locuzione idiomatica»: l'interesse di queste forme sta nel fatto che «offrono la possibilità di verificare in che modo si è formato e poi consolidato un certo sintagma»); 3. Scoperta di altre accezioni: cfr. p. 45).

<sup>16</sup> Sul VoDIM (*Vocabolario dinamico dell'italiano moderno*) cfr. Biffi (2016), Marazzini, Maconi (2018), Barbera (2018), Biffi (2018). Sull'uso dei *corpora* per la ricerca linguistica cfr. Riccio (2016); sul caso del CoLIWeb cfr. Biffi, Ferrari (2020).

<sup>17</sup> Cfr. quanto già specificato in Heinz, Pizzoli, Serianni (2022: 119).

Ed è stato così che per un lavoro pensato con una finalità tanto a lungo termine si è deciso di procedere a piccoli passi. Usando un'immagine molto cara allo stesso Serianni, si è stabilito di allargare l'indagine secondo una disposizione degli obiettivi a cerchi concentrici: oltre alle lingue già censite nel DIFIT, dunque, si sarebbe partiti dalle grandi lingue di cultura molto vicine all'italiano (innanzitutto il castigliano, dapprima solo per la varietà spagnola, alla quale poi sarebbero state affiancate le varietà usate negli altri paesi ispanofoni, e poi il portoghese), proseguendo con lingue europee di rilievo culturale per le quali si disponeva già di raccolte avanzate (catalano, ungherese e polacco). In un cerchio successivo si sarebbero poi aggiunte le altre lingue rilevanti in relazione geografica e culturale all'italiano, valutate anche sulla base della disponibilità dei singoli ricercatori (al momento, come specificato da Matthias Heinz (§ 2), sono in lavorazione le raccolte relative agli italianismi in maltese, neogreco, macedone, slovacco, finlandese, svedese, cinese mandarino).

La progressione del lavoro per successive fasi di sviluppo, aspetto compatibile solo con una piattaforma informatizzata aggiornabile in tempo reale e suscettibile di modifiche in corso d'opera anche nell'impianto progettuale, ha reso possibile implementare funzioni diverse sulla base dell'inserimento delle nuove lingue, che spesso hanno caratteristiche che l'italiano non contempla (si è potuto dare seguito, per esempio, all'esigenza di registrare la trascrizione delle parole in alfabeti diversi da quello latino, o di indicare il genere neutro); altre funzioni, infine, sono state progettate tenendo conto dei meccanismi stessi del prestito. Nel presente contributo, concepito con l'obiettivo di descrivere le funzioni della piattaforma e le possibilità che questo strumento offre per la ricerca nell'ambito del contatto, si descriverà quanto realizzato finora (ponendo l'accento su questioni di metodo) e quanto messo in programma per le successive fasi di lavorazione dell'OIM.

Come si è anticipato, nel mettere insieme uno strumento che accoglie i dati selezionati da più gruppi di ricerca attraverso l'uso di più fonti lessicografiche, diverse per taglio, indirizzo ideologico, presenza di specifici campi, ecc., ci si è dovuti misurare con la enorme eterogeneità delle informazioni: fin dall'inizio, pertanto, è stato necessario rinunciare alla pretesa di esaustività e omogeneità nella descrizione dei prestiti, cercando un compromesso tra la necessaria rigidità della struttura informatica (che ha richiesto una classificazione rigorosa dei tipi di prestito: è stato fondamentale distinguere, per esempio, tra prestito diretto e indiretto) e la multiforme provvisorietà delle informazioni incerte. Il risultato di questo compromesso è stata la creazione di spazi di movimento, come per esempio la categoria dei "prestiti incerti", con cui etichettare tutti i lemmi per i quali non è ancora possibile, per varie ragioni, ricostruire con esattezza l'etimologia (per la difficoltà di distinguere tra prestito e formazione autoctona, oppure individuare un'eventuale altra lingua di mediazione, magari della stessa area romanza, se non del latino, o ancora per la mancanza di indizi sufficienti per stabilire la provenienza da una forma dello standard o di un dialetto dell'italiano). Si è reso necessario, infine, uno spazio di dialogo, che si è rivelato essenziale in molte circostanze, nel quale inserire testo libero per aggiungere precisazioni, attenuazioni, ipotesi di lavoro, richieste di controllo ad uso interno ed esterno<sup>18</sup>. Solo grazie a tutte queste accortezze si è potuto rendere disponibile sul portale <https://www.italianismi.org/> la prima scrematura di dati raccolti in un primo tempo con

<sup>18</sup> Va considerata tra l'altro la necessità che il sistema di inserimento fosse quanto più possibile amichevole e di facile utilizzo per gli studiosi che stanno contribuendo al lavoro: a tal fine la piattaforma è stata corredata di un normario e di diversi manuali di inserimento intesi a dare soluzioni a tutta la casistica di problemi prevedibili. Come accade per la stessa piattaforma, anche il materiale prodotto per agevolare il compito dei ricercatori viene costantemente aggiornato via via che si presentano problemi inediti.

maglie larghe, in attesa di raffinare ulteriormente la classificazione e attivare anche le chiavi di ricerca più sofisticate.

Accettare compromessi, in qualche maniera, è stata l'unica possibilità per cominciare a rendere pubblico il sistema e lasciare comprendere quanto spazio di ricerca offrirà il confronto tra lingue in cui si sono radicate parole italiane: l'attuale piattaforma, pur essendo stata concepita per rendere disponibili per la comunità scientifica i dati in forma completa, consente al momento di leggere la superficie, mentre lascia solo intravedere ciò che si sta ancora facendo (sono molte, inevitabilmente, le schede relative a parole in lavorazione sia per le lingue di cui si è appena iniziato il censimento sia per le stesse lingue già completate, ma per le quali mancano ancora passaggi di controllo fondamentali); è del tutto invisibile, invece, il cosiddetto lavoro di *backoffice*, che contempla una serie di funzioni innovative delle quali si vorrebbe dare qualche dettaglio in questa sede.

### 3.2. I criteri per la raccolta dei dati

Partirei dal descrivere i criteri generali che hanno orientato la prima organizzazione del materiale: innanzitutto la definizione stessa di prestito, che nel nostro caso, in un'ottica di osservatorio, ha contemplato tutte le forme di contatto tra lingue, capaci di rivelare l'impatto dell'italiano sulla lingua ospite, misurato anche in una forma sfumata di intensità. Per questo sono stati inclusi nel censimento, che si muove sempre a partire dalla voce italiana che ha dato origine alla nuova forma nella lingua ospite, sia i prestiti diretti sia quelli indiretti, ma anche gli pseudoprestiti, i deonomastici, fino ad arrivare agli italianismi di secondo grado<sup>19</sup>. Per garantire una certa uniformità al lemmario e dare all'utente la possibilità di ritrovare con maggiore facilità le forme di partenza, ci si è risolti a presentare i dati in un'ottica sincronica, registrando la voce nella forma in cui è diffusa nell'italiano contemporaneo, pur essendo ben consapevoli del fatto che in molti casi i prestiti sono stati originati da parole che avevano forme oggi desuete. Valga tra tutti il caso di *affresco*, una parola molto nota e ben rappresentativa dell'arte italiana: questa tecnica pittorica si è diffusa all'estero soprattutto nella forma oggi non più in uso *a fresco*, attestata dal 1535. E in effetti l'uso della voce unverbata *affresco* era biasimato ancora dall'Ugolini nel 1855 («*affresco*, sostantivamente usato non ha la nostra lingua, ma solo il modo avverbiale *a fresco*, *dipingere a fresco*: non dirai dunque affreschi, ma pitture a fresco: potrai però dir freschi per

<sup>19</sup>Di questa scelta si è già dato conto nei precedenti contributi dedicati all'OIM: in particolare, sulla nozione di italianismo di cui ci si è serviti per l'OIM e sui possibili problemi derivati da alcune delle scelte compiute per l'osservatorio cfr. Pizzoli (2017) e Heinz, Pizzoli, Serianni (2022); sull'utilità di registrare i prestiti indiretti mi permetto di rinviare ancora a Pizzoli (2017: 173-74). Proprio l'introduzione della categoria del prestito incerto, peraltro, ha permesso di registrare, sia pure prudentemente, una serie di lemmi per i quali l'attribuzione non è definitiva e la cui assenza (o viceversa presenza forzata) avrebbe potuto sovra- o sottostimare l'impatto dell'italiano sulle altre lingue. Da notare che anche per le parole italiane di partenza viene registrata l'etimologia remota, in modo da poter riconoscere i prestiti per i quali l'italiano è lingua tramite ma che originano da altre lingue straniere (tra i tanti esempi possibili si può citare il caso di *camuccà* "tessuto di seta damascata", un arabismo in italiano, oggi desueto, che in ungh. ha dato vita per il tramite dell'italiano a diverse forme, anche arcaiche: *kamuka*, † *kamuka*, † *camuco*, † *kamoka*. Per la nozione di prestito di secondo grado cfr. Serianni (2017: 42); sull'utilità di ampliare lo sguardo a tutte le forme del contatto, anche servendosi di fonti meno ortodosse rispetto a quelle lessicografiche, cfr. Pizzoli, Heinz (2022). Che la consultazione di materiale nuovo dia risultati interessanti è dimostrato per esempio dalle ricerche condotte per spagnolo e catalano (cfr. Clavería Nadal-Julà Luna, 2021), o per il macedone (cfr. Saržoska, Ivanovska-Naskova, 2021) che hanno consentito di ampliare la lista di italianismi attraverso il ricorso a raccolte di testi recenti.



affreschi»; cfr. DELI)<sup>20</sup>. Eppure un utente di oggi non avrebbe dubbi nel ricercare questa parola nella forma unverbata, l'unica registrata a lemma dai dizionari dell'uso. Naturalmente le forme arcaiche vengono puntualmente indicate nella scheda perché non sia sacrificata la possibilità di ricavare informazioni anche in diacronia: nell'esaminare gli italianismi, infatti, non può essere trascurata la componente essenziale, in chiave culturale, dell'irradiazione nei secoli passati delle parole italiane anche quando si sia modificata la forma o il significato (come si vedrà più avanti per il caso di *fracassare*).

Seguendo questa logica, le voci poste a lemma nell'OIM sono classificate sulla base della loro attuale diffusione: per rendere quanto più possibile confrontabili le valutazioni, ci si è affidati al GRADIT, il dizionario sincronico con il lemmario più consistente: si è stabilito di considerare neutri i termini registrati senza nessuna marcatura in GRADIT, mentre di segnalare come neologismo le voci non registrate dal GRADIT (né dal GDLI, per evitare di etichettare in modo incongruo parole desuete; si considerano "arcaiche/desuete" le parole o le accezioni che nel GRADIT sono marcate con LE (letterario) e/o OB (obsoleto: per es. *anguistara* 'caraffa' o *abbronzare* nel senso di 'dare un colore simile a quello del bronzo, brunire') mentre si considerano "non più in uso" (e vengono per questo evidenziate con la *crux*) le voci non registrate dal GRADIT ma presenti nel GDLI (per es. *andrivello* 'cavo che passa per una taglia fermata in cima a un albero a pennone'). Ulteriori sfumature dal punto di vista diacronico, come l'etichetta BU (basso uso) proposta dal GRADIT, o qualunque altra possibile osservazione, sono aggiunte nelle note a campo aperto (è il caso di *parrucca*, nell'accezione spregiativa di 'parruccone')<sup>21</sup>.

Analogamente, le forme arcaiche di parole oggi diffuse in una variante diversa sono comunque indicate (specie perché spesso sono proprio quelle che hanno dato origine ai prestiti in epoche antiche), e sempre accompagnate dalla precisazione della loro uscita d'uso<sup>22</sup>.

<sup>20</sup> Da notare inoltre che anche dal sost. *fresco* nei significati legati alla temperatura (da cui anche il significato specializzato nell'ambito della moda, in relazione al tessuto "fresco di lana") si sono originati dei prestiti, in particolare in polacco, tedesco, ungherese. Va registrata anche un'accezione in ambito gastronomico nell'ingl. *fresco* 'bevanda rinfrescante' (dal 1880: cfr. DIFIT).

<sup>21</sup> Questo tipo di schedatura, per conseguenza, fa sì che in prospettiva le informazioni inserite nelle caselle con campo libero non siano recuperabili dalla maschera di ricerca come filtro specifico: tuttavia, come in ogni dizionario elettronico, si prevede comunque la ricerca a tutto campo e dunque alcune informazioni sono sempre disponibili. Da notare che anche nel caso delle notazioni di tipo diacronico l'indicazione sulla diffusione nell'italiano contemporaneo viene agganciata alla scheda significato, e quindi ripetuta per ogni singola accezione della voce sia in italiano sia nelle lingue ospiti: l'italiano *arlecchino* (sost.), per esempio, diffusissimo in molte lingue – al momento catalano, francese, portoghese, spagnolo, tedesco, inglese, polacco, ungherese, maltese – anche in diverse forme, ha generato prestiti riconducibili al significato principale ('Persona travestita da Arlecchino'), ma anche a usi figurati ('Persona che ama gli scherzi, buffone' e 'Cose composte da elementi eterogenei, discordanti, che non possono essere accostati né fusi insieme') e specialistici della zoologia (nella forma *moretta arlecchino*: 'Anatra selvatica (*Histrionicus histrionicus*), unica specie del genere Istrionico', rintracciabile nell'inglese – mediato dal francese – *harlequin*, accanto all'uso esteso, innovativo, a 'una razza di cani dal pelo maculato' e all'"opale nobile o orientale"); il significato originato dall'uso veneziano di gelato o sorbetto 'costituito da gusti diversi variamente colorati', attestato nel catalano e nello spagnolo, non è considerabile vitale nell'italiano di oggi e viene marcato con l'etichetta BU tratta da GRADIT. Viene marcata come innovativa l'accezione del francese 'Piatto formato da avanzi diversi', pur evidentemente derivata dalle precedenti, così come l'inglese (raro) *to harlequin* 'Far sparire come per magia, come Arlecchino in una pantomima', 'Recitare la parte di Arlecchino', 'Colorare, decorare con colori a contrasto', che prevede addirittura un cambio di categoria morfologica e può dunque essere considerato un prestito di secondo grado.

<sup>22</sup> Valga per questa tipologia il caso di *quarantena*, che in alcune lingue, come il polacco, produce prestiti a partire dalla forma antica *quarantana* o dal lat. *quarantanna* (cfr. la ricostruzione di Heinz in Pizzoli, Heinz, 2022).



### 3.3. La scheda di inserimento dei dati

Varrà la pena di descrivere più nel dettaglio quali sono le informazioni raccolte nelle schede, che, in un futuro auspicabilmente vicino, saranno tutte recuperabili tramite l'interrogazione della maschera e selezionabili con specifici filtri<sup>23</sup>.

Proprio come in un albero genealogico (organizzato per famiglie linguistiche apparentate dal punto di vista tipologico), il sistema prevede che possa essere ricostruita la relazione di dipendenza di una forma dalla parola italiana di partenza, con il rispettivo grado di acclimatemento morfologico: dall'italiano *calzone*, per esempio, si è sviluppata nel francese la forma adattata *caleçon* (con le varianti *caleçons*, *calleson*, *caleson*, *calçon* e i rispettivi significati: 'mutande da donna', 'mutande da uomo', 'costume da bagno'), ma anche gli italianismi di secondo grado, di uso familiare e gergale, *cal(e)bard* o *calebar* e *cal(e)cif*.

In tedesco invece il prestito è arrivato tramite il francese, e dunque in quel caso la parola *caleçons* sarà registrata come prestito indiretto, con l'indicazione della lingua di mediazione.

Da notare che dall'italiano si è diffusa anche la forma *calzone* in riferimento alla popolare specialità della cucina meridionale ('Disco di pasta di pane ripiegato in metà e farcito con mozzarella, pomodoro, prosciutto e sim., da mangiarsi cotto in forno o fritto'), che troviamo in molte altre lingue come conseguenza del successo recente dei prodotti da forno italiani (sul successo internazionale di *ciabatta*, per esempio, cfr. Coveri, 2016). Nelle lingue ospiti, dunque, possono convivere prestiti con diversi significati (e a volte anche diverse forme), che sono registrate nella piattaforma in base alla trafilata di arrivo: per il tedesco, dunque, si troverà sia la forma adattata *caleçons*, arrivata dal francese e relativa al vestiario, sia la forma non adattata *calzone*, usata per l'ambito gastronomico. La stessa convivenza di forme, probabilmente non percepita dal parlante comune, si osserva in inglese: l'OED registra sia la forma oggi arcaica *calzoons* (o *calsoun*) 'calzoni, capo di abbigliamento in uso presso gli orientali', dal 1615, e *calzone* (cibo), dal 1933 (come prestito originariamente dell'inglese statunitense). L'ungherese, che evidentemente non era stato raggiunto in epoca più antica dalla forma legata all'abbigliamento, ha raccolto invece il solo prestito gastronomico, nella forma non adattata *calzone*, come accade comunemente nei prestiti più recenti<sup>24</sup>.

Come si è detto, alcune categorie sono state appositamente predisposte per l'osservatorio: oltre al caso già spiegato del prestito indiretto, sono state aggiunte le categorie dello "pseudoitalianismo" e della "formazione ibrida". Con il primo termine si categorizzano i casi, ben descritti nella letteratura sul prestito, di formazioni imitative "su modello italiano"<sup>25</sup>. L'etichetta di "formazione ibrida", invece, viene assegnata ai casi in

<sup>23</sup> Come anticipato, al momento la voce si presenta in forma semplificata senza possibilità di ricerche incrociate: sono ricavabili, oltre ai dati relativi alla voce italiana di partenza (forma o varianti della forma grafica, informazioni grammaticali, significato), la presenza nelle lingue straniere di una o più forme di prestito corredate dai principali tratti di interesse lessicografico (data della prima attestazione, categoria grammaticale, genere, numero, significato).

<sup>24</sup> Tutti gli esempi citati in questo contributo sono tratti dalla banca dati OIM: ringrazio le redazioni dei gruppi OIM che li hanno messi a disposizione. I dati per francese e inglese, registrati già in DIFIT, sono tratti rispettivamente da TLFi e OED.

<sup>25</sup> Sono i casi, già ricordati da Stammerjohann (2010), delle «formazioni che sembrano italiane ma non lo sono, per es. la locuzione *tutti frutti*, che si trova in molte lingue nel senso di 'frutta mista', poi 'gelato misto', quindi usato per qualsiasi *mixtum compositum* [...]. Similmente, il romeno *contropapa* è composto da due parole italiane, mentre il composto italiano è *antipapa*». Gli pseudoitalianismi sono parole create sulla base di assonanze con forme italiane (si copia dunque il modello, la struttura morfologica) ma che non hanno nessuna circolazione in Italia, anche se potrebbero essere classificate come italiane nel dizionario della lingua ospite (la base di partenza, dunque, sarà una parola "non esistente in italiano"). Gli pseudoitalianismi vanno

cui il prestito è composto con elementi italiani aggiunti a quelli locali: è il caso, per esempio, del nomignolo *mangiakekeka* (diffuso nell'inglese canadese e statunitense anche nella variante *munjacake*), usato dagli italo-canadesi per designare un canadese di origine inglese, oppure di *locateli*, usato popolarmente nelle varietà di spagnolo di Perù, Cile, Argentina, Uruguay, Bolivia ed Ecuador per riferirsi in modo scherzoso a persone bizzarre, pazze (in questo caso il suffisso tipicamente italiano *-ello* viene aggiunto a una base originaria *loco* 'pazzo')<sup>26</sup>. La formazione ibrida si distingue, sia pure a volte con una certa difficoltà, dall'italianismo di secondo grado, considerando la produttività dei formanti nella lingua ospite. Se la composizione avviene in modo occasionale ibridando elementi italiani e locali si parla di formazioni ibride; il secondo grado si presuppone invece quando il prestito diventa produttivo e dà vita a nuove forme derivate che si creano secondo i meccanismi di formazione indigeni (è il caso, per esempio, dell'agg. maltese *cucesk* 'cosa da ignoranti', derivato da *cuc* 'somaro, ignorante', prestito originato dal siciliano *ciucciu*, o del s. maltese, di ambito marinairesco, *arbulatura* 'il cordame degli alberi di un veliero' (derivato da *arblu* 'albero di barca o di bandiera, palo', a sua volta proveniente dal siciliano *arbulu*).

### 3.4. Sulla semantica dei prestiti

Ogni scheda relativa alla forma è poi completata da una scheda (o più schede) che specificano il significato: a costo di fornire dati sovrabbondanti, si è stabilito di ripetere per ogni prestito il significato registrato dalle fonti, perché solo in questo modo possono essere intercettate le sfumature d'uso che ogni parola ha sviluppato nella lingua nella quale si è acclimata.

Possiamo considerare l'esempio della parola *altana*, che in italiano è attestata dalla fine del XV sec. con il valore di 'loggia aperta che si eleva sul tetto di un edificio' (GRADIT) e che si è diffusa nell'Europa orientale con significati totalmente sovrapponibili (ted. *Altan* 'loggia aperta che si eleva sul tetto di un edificio'; ungherese *altan*, oggi desueto, 'loggia aperta sul tetto di un edificio'), mentre nel polacco *altana* ha assunto il significato più specifico di 'gloriette, piccolo padiglione in giardino'.

Sfumature analoghe riguardano anche forme non più in uso: la parola italiana *aranciata* è stata adattata in francese come *orangeade*, termine che in passato è stato usato per indicare la marmellata di arance; altre volte il passaggio che ha prodotto lo slittamento semantico è meno intuitivo, e produce cortocircuiti più curiosi (dalla forma dialettale arcaica *articiocco*, si è prodotta, oltre al fr. *artichaut* e all'ingl. *artichoke*, con molte varianti storiche, anche la forma tedesca *Artischocke* che a partire dall'inizio del XX sec. viene usata, oltre che per indicare l'ortaggio, per definire una 'persona strana, curiosa, il cui modo d'agire è inconcepibile'; dall'italiano *lampion*, oltre al riferimento all'illuminazione, il francese *lampion* ha assunto, dal XIX sec., il significato di 'tricorno' nell'ambito dell'abbigliamento e della moda, mentre il portoghese *lampião*, dal XX sec., viene usato popolarmente per identificare un 'tifoso della squadra di calcio del Benfica'). È interessante notare che a volte gli slittamenti di significato possono riguardare aree linguistiche più estese: la parola

distinti dai prestiti adattati (che hanno un riscontro di tipo fonomorfológico con le caratteristiche della lingua ospite) e dagli slittamenti semantici, attraverso i quali vengono assegnate a formazioni originariamente italiane nuove valenze. A proposito del fenomeno speculare dei "false anglicisms" in italiano e delle questioni terminologiche cfr. Furiassi (2015: 253-277).

<sup>26</sup> Il termine viene usato, solo in Argentina, anche per indicare un piccolo panino ripieno, solitamente con prosciutto e formaggio (cfr. DA, s.v. *locateli*: <https://www.asale.org/damer/locateli>).

italiana *aggiornamento*, per esempio, ha una diffusione nello spagnolo d'Argentina (*aggiornamiento*) e nel maltese (*aggiornament*) con un valore simile all'italiano ('rinnovamento e revisione – d'una cognizione, d'un'opinione, d'un costume, d'un'opera –, in rapporto col progresso dei tempi'), mentre in inglese britannico, francese, tedesco, polacco è entrata come forma non adattata (*aggiornamento*) solo nel significato ristretto all' 'adattamento della chiesa cattolica, per estensione anche di altre chiese e religioni, al mondo moderno' (in polacco con un riferimento più circostanziato al programma di rinnovamento religioso avviato da Papa Giovanni XXIII).

Accertare che un prestito abbia subito un cambiamento semantico e che possa essere ritenuto innovativo rispetto all'uso che di quel termine si fa in Italia è un'operazione molto delicata, perché ci si deve basare necessariamente su un riscontro puntuale con la documentazione storica: può capitare, infatti, che una parola abbia conservato nella lingua ospite un significato che invece in italiano è ormai arcaico. È il caso di *fracassare*, che i dizionari dell'uso contemporaneo spiegano solo in riferimento all'atto di rompere (GRADIT: 1. fare a pezzi qcs., rompere sconquassando; 2. fig., ridurre qcn. in malo modo, conciarlo male; 3. Intr. OB crollare rovinosamente, sconquassarsi, frantumarsi; Zingarelli (2023): fare a pezzi violentemente e con rumore; intr. † crollare in frantumi; Treccani *online*: Rompere, fare in pezzi, sconquassare battendo violentemente; con riferimento a persona, ridurre in cattivo stato, conciare in malo modo). Solo ricorrendo al GDLI si possono recuperare le accezioni figurate oggi desuete ('confutare (un'idea, un argomento)'; 'rovinare economicamente') che si sono però radicate nel portoghese *fracassar* 'fallire' e nello spagnolo *fracasar* 'fallire (detto di un progetto)' e 'avere un risultato negativo in un affare (detto di una persona)<sup>27</sup>.

La stessa cautela va usata per le accezioni al momento non registrate dai dizionari italiani ma che potrebbero avere più avanti una diffusione anche in Italia, magari per effetto del mercato globale (si tratterebbe di quelli che si definiscono "cavalli di ritorno"): basti considerare l'esempio della parola *arancino*, che in italiano viene usata esclusivamente in riferimento alla tradizionale preparazione gastronomica ('crocchetta di riso fritta di forma sferica, ripiena di formaggio, rigaglie di pollo e ragù di carne, tipica della cucina meridionale') e che però fuori d'Italia viene usata anche per commercializzare le 'scorzette di arancia al cioccolato', come è stato documentato per il tedesco, specie nella varietà austriaca, nelle forme *arancini* e *aranzini* fin dal 1839 (cfr. DIFIT)<sup>28</sup>.

Il significato introdotto dall'industria dolciaria è senz'altro innovativo perché non registrato nella lessicografia italiana, ma non possiamo escludere che in un prossimo futuro l'uso si diffonda anche in altre zone. Anche se l'antichità di attestazione nel tedesco austriaco lascerebbe pensare a un'innovazione locale, in assenza di documentazione più larga non è da escludere che la nuova accezione sia nata altrove o sia comunque promossa dall'Italia (come dimostra la presenza, per esempio, dello stesso prodotto nel catalogo dell'azienda ciampinese *Menicucci*, in quella amalfitana *Sal de riso*, nella forlivese *Gardini* e nella alassina *Balzola*)<sup>29</sup>.

<sup>27</sup> Da segnalare che in port. si registra anche il significato di 'fare un gran rumore' e in spagn. quello di 'rompersi contro uno scoglio (detto di un'imbarcazione)'.  
<sup>28</sup> Ritroviamo quest'uso anche nell'azienda vitivinicola ticinese *Matasci*, che vende questa specialità per il mercato locale e forse non solo: cfr. <https://matasci-vini.ch/products/arancini>.

<sup>29</sup> Cfr., rispettivamente, <https://www.cioccolatomenicucci.it/prodotto/arancini-ricoperti-di-cioccolato-extra-fondente/>; <https://www.salderisoshop.com/it/gli-sfizioli/106-arancini-al-cioccolato.html>; <http://www.gardinicioccolato.it/prodotto/astuccio-filetti-arancia-canditi-ricoperti-cioccolato-extra-fondente-62-cacao/>; <https://www.bacidialassio.com/it/la-pasticceria/25-arancini.html>.

### 3.5. Prospettive di lavoro future

Anche solo questo esempio mostra chiaramente come abbia senso che l'osservatorio sia concepito come una struttura aperta, in continuo aggiornamento in linea con l'evoluzione della lingua contemporanea e le possibili nuove acquisizioni in un settore lessicale dinamico e non facilmente circoscrivibile, nelle forme e nei significati, come questo; da una parte, man mano che si rinvencono prestiti derivati da parole italiane non ancora registrate, possono essere aggiunte schede nuove (sia dal lato italiano, sia da quello della lingua ospite); d'altra parte, si deve poter tornare su schede già completate per molteplici ragioni: oltre all'esigenza di correggere qualche errore, si può presentare la necessità di arricchire la voce italiana di partenza con le nuove accezioni registrate nelle lingue inserite in fasi successive<sup>30</sup>; si possono perfezionare le schede con gli apporti di ricerche aggiornate; si possono avviare controlli (a campione o sistematici) grazie al raffronto tra coppie di lingue: per esempio, se un italianismo è valutato come prestito indiretto in una lingua, sarà bene che lo stesso prestito sia anche registrato come diretto in quella che viene indicata come lingua di mediazione. È il caso di *arresto*, che è documentato in polacco, nella forma *areszt*, come prestito arrivato tramite il tedesco; ma evidentemente le fonti usate per censire gli italianismi in tedesco non offrono questo dato, che andrà verificato tramite altre risorse ed eventualmente aggiunto; lo stesso si verifica per il portoghese *maçapão*, dall'italiano *marzapane*, per il quale si indica lo spagnolo come lingua di mediazione, anche se nell'elenco degli italianismi in spagnolo la voce risulta ancora assente<sup>31</sup>.

Ma i controlli sulla direzione del prestito valgono anche nel caso di parole che in italiano sono considerate forestierismi e che invece dalle fonti straniere risultano classificate come termini provenienti dall'italiano: è il caso, per esempio, della voce regionale *bazzoſſia* ('minestrone o gener. vivanda abbondante e grossolana', 1585, ma poi anche 'insieme di cose disordinate', av. 1850), data come ispanismo in GRADIT, ma che già GDLI ricollegava «all'it. antico *battisòſſia* e *battisòſſiola* 'palpitazione di cuore, spavento, confusione', attraverso le forme dial. settentrionali», precisando che anche lo spagnolo *bazofia* va considerato come derivato dall'italiano: la conferma di questa ipotesi verrebbe anche dalla presenza di forme analoghe, ritenute italianismi, in catalano (*gasòſſia*, con il significato di 'piatto formato dagli avanzi di diverse pietanze mescolati') e in portoghese (*bazóſſia*, che vale 'bollito preparato con gli avanzi' ma ha assunto anche il significato innovativo di 'boria').

<sup>30</sup> Si consideri che nella scheda italiana vengono registrate solo le accezioni che hanno dato vita a prestiti attestati nelle altre lingue (altrimenti il lemmario italiano dovrebbe essere l'equivalente di quello di un dizionario storico); con un effetto apparentemente paradossale, dunque, può capitare che nella scheda italiana manchino significati che nell'italiano contemporaneo possono risultare prevalenti: è il caso per esempio di *camerata*, attestato nel fr. *camerade*, nel port. *camarata*, nel ted. *Kamerad*, con significati tutti riconducibili al significato militare di 'compagno d'armi, commilitone' e 'militari, o i collegiali, che dormono in una stessa camerata'; nella scheda italiana manca il significato oggi più marcato, usato 'come appellativo reciproco degli iscritti al partito fascista' (GRADIT), che al momento non risulta attestato in nessuna lingua straniera. Nel caso in cui l'aggiunta di una nuova scheda rendesse necessaria una specificazione ulteriore si potrà provvedere successivamente: il sistema segnala con un apposito avviso i nuovi italianismi agganciati a una scheda italiana già completa, in modo che si possa verificare se il significato del prestito è compatibile con quanto registrato fino a quel momento nella scheda italiana o se non servano integrazioni.

<sup>31</sup> Sulle relazioni tra lingue limitrofe, per esempio, si rimanda ai casi, già citati in Heinz, Pizzoli, Serianni (2022: 120 n.2), dello sp. *artesano*, passato per il tramite del cat. *artesa*, o dello sp. *bergantín*, dall'italiano *brigantino* tramite il catalano *bergantí* (con un possibile influsso francese), su cui cfr. anche Gomez Gane, Julià Luna (in stampa).

Altri campi che potranno utilmente essere rivisti, alla luce dei dati presenti nelle schede già validate, sono quelli relativi alla datazione. Il sistema prevede che – come in qualunque repertorio lessicografico – vengano indicate le date, sia per le parole italiane di partenza sia per i prestiti (dove possibile anche specificando le date delle diverse accezioni), sulla base di quanto rinvenuto sulle fonti: l’indicazione può essere puntuale nel caso di attestazione certa, ma anche più generica (con un arco temporale anche piuttosto ampio e con diverse forme di attenuazione: circa XVII sec., prima metà del XV sec., fine del XIX sec., 1455-62, ecc.). Le datazioni approssimate sono visibili nella scheda come testo libero, ma, per consentire il recupero di termini appartenenti allo stesso periodo, è stato studiato un sistema di equivalenze numeriche che consentirà un’estrazione massiva dei dati nonostante questi siano presentati sotto forme diverse.

È ben noto come la datazione di un termine sia sempre suscettibile di revisioni, e che anche le schede OIM potranno essere aggiornate tenendo conto dei dati acquisiti tramite la pratica che Serianni aveva definito il «facile gioco delle retrodatazioni» (2006: 43). Tuttavia nell’osservatorio l’informazione sulla data di prima attestazione assume una valenza piuttosto rilevante, perché consente di fare verifiche incrociate e ricostruire con maggiore precisione la trafila del prestito, determinando, per esempio, se un prestito è arrivato in una determinata lingua in forma diretta o indiretta. Un caso molto tipico di incertezza sulla direzione del contatto è quello che riguarda le possibili interferenze tra lingue della stessa tipologia: solo grazie all’indicazione della data di prima attestazione si potrà capire in quale lingua il prestito ha attecchito per poi essere trasmesso a lingue vicine<sup>32</sup>.

Altri importanti riflessi dell’analisi della data della prima attestazione hanno a che vedere con la questione dell’appartenenza della parola di origine al repertorio dell’italiano o non piuttosto a un dialetto. In molti casi, infatti, il contatto si è attivato a partire da una voce dialettale che solo in un momento successivo si è poi diffusa in tutto il territorio italiano passando allo *status* di dialettismo: per esempio la parola *pesto*, una delle specialità regionali più conosciuta all’estero e accolta in diverse lingue nel mondo, è entrata nei dizionari italiani a partire dal 1918 ma viene registrato in inglese dall’OED, evidentemente per tramite dialettale, già prima del 1848 (e poi registrato in testi novecenteschi, sempre in riferimento all’Italia).

Metterà conto, in ogni caso, accogliere criticamente le indicazioni fornite dai dizionari, che, specie nel caso dei prestiti (non sempre tempestivamente registrati per via scritta), potrebbero dare adito a ricostruzioni non sempre affidabili. Basti l’esempio, suggerito da Paolo Cherchi, di *cosa nostra*, che l’OED riconduce a una prima attestazione in un giornale inglese (l’*“Economist”*) nel 1963 e solo come seconda quella di un testimone americano (Saul Bellow, nel celebre romanzo *Herzog*, 1964): in questo modo sembrerebbe di poter dire che la locuzione sia arrivata negli Stati Uniti solo dopo essersi acclimata in Gran Bretagna. Ma l’esistenza di un’organizzazione criminale radicata prevalentemente negli Stati Uniti e di così grande notorietà oltreoceano lascerebbe sospettare che il termine circolasse già da tempo nei media americani, ben prima di essere registrato da una fonte inglese<sup>33</sup>.

<sup>32</sup> Un caso piuttosto interessante per l’analisi del contatto per vicinanza territoriale è quello, ricostruito da Gomez Gane, Julià Luna (in stampa), del termine musicale *battuta* ‘unità di tempo’ (GRADIT), passato dall’italiano a catalano, portoghese e castigliano, con cambiamenti di significato anche piuttosto rilevanti che solo la comparazione delle date di prima attestazione consente di mettere nel giusto ordine.

<sup>33</sup> Peraltro la spiegazione di *cosa nostra* offerta dall’OED è proprio “The Mafia in the U.S.A.”. Una testimonianza della larghissima circolazione della locuzione nell’area americana viene dall’uso che se ne fa

Tutti gli esempi citati e le questioni che sollevano rendono bene il senso della complessità di questo lavoro: un'opera che, come si è detto, aspira a costituire un punto di riferimento per gli studi sul contatto linguistico a partire dalla lingua italiana ma che non può prescindere da un continuo lavoro di accomodamento dei criteri e delle scelte generali<sup>34</sup>.

La consapevolezza delle difficoltà che si incontrano nel mettere insieme informazioni provenienti da sistemi tanto eterogenei consentirà anche un uso più accorto dei dati raccolti e della necessità di trattarli in forma provvisoria. È proprio sul senso di questa precarietà che vorremmo si costruisse però, con un provocatorio paradosso, un impianto solido e di lunga durata.

## RIFERIMENTI BIBLIOGRAFICI

- Actes du XXIXe Congrès international de linguistique et de philologie romanes, (Université de Copenhague, 1-6 de julio de 2019), Société de Linguistique Romane / Éditions de linguistique et de philologie, Vol. I, pp. 791-803.
- Barbera M. (2018), “Approssimazioni al VoDIM”, in *Italiano digitale*, VII, 4, pp. 138-158.
- Biffi M. (2016), “Progettare il corpus per il vocabolario postunitario”, in Marazzini, Maconi, 2016, pp. 259-280.
- Biffi M. (2018), “Strumenti informatico-linguistici per la realizzazione di un dizionario dell'italiano post-unitario”, in Fioredistella Iezzi D., Celardo L., Misuraca M. (a cura di), JADT '18. *Proceedings of the 14th international conference on statistical analysis of textual data*, Universitalia, Roma, vol. I, pp. 99-107.
- Biffi M. et al. (2014), *Osservatorio degli Italianismi nel mondo*, Accademia della Crusca, Firenze: [www.Italianismi.org](http://www.Italianismi.org).
- Biffi M.-Ferrari A. (2020), “Progettare e realizzare un ‘corpus’ dell'italiano nella rete: il caso del ‘CoLIWeb’”, in *Studi di lessicografia italiana*, XXXVII, pp. 357-374.
- Caffarelli E. (2016), “I marchionimi italiani e la loro diffusione internazionale”, in D'Achille P., Patota G. (a cura di), *Italiano della creatività: marchi e costumi, moda e design*, Accademia della Crusca, goware, Firenze, pp. 29-48.
- Clavería Nadal G., Julià Luna C. (2021), “Los italianismos en el español y en el catalán contemporáneos”, in Schøsler L., Juhani Härmä (eds.), *Actes du XXIXe Congrès international de linguistique et de philologie romanes* (Copenhague, 1-6 juillet 2019), Eliphi, Strasbourg, pp. 791-803.
- Cortelazzo M. (2022), “Retroduzioni al «DELI» da traduzioni letterarie ottocentesche”, in *Studi di lessicografia italiana*, XXXIX, pp. 247-312.

oggi come parola-simbolo nelle insegne di esercizi commerciali, per esempio a Buenos Aires, per individuare e personalizzare il locale in lingua italiana (cfr. Caffarelli, 2016: 45).

<sup>34</sup> Si potrebbero citare, ancora, le difficoltà di accordare altri campi, come quello della assegnazione dei termini ad ambiti semantici specifici (si è scelto di prevedere alcune categorie molto ampie, come “diritto/politica/burocrazia/amministrazione” e di riservare una casella a parte per ambiti più circoscritti, come “chiesa/religione”; altri ambiti specialistici, come per es. quello della geografia, sono ricondotti a categorie più generali, come le “scienze naturali”); lo stesso discorso andrebbe fatto per l'indicazione delle marcature d'uso sull'asse della diafasia, posto che non tutti i dizionari si servono delle stesse categorie (oltre a singoli problemi di traduzione, vanno fatte attente valutazioni sulla diversa considerazione di etichette come dialettale, comune, popolare, gergale e simili).

- Coveri L. (2016), “Neoitalianismi gastronomici dalle regioni al mondo: il caso di ciabatta ‘tipo di pane’”, in Morgana S., Demartino D., Stanchina G. (a cura di), *L’italiano del cibo* (Atti del convegno, Milano, 30 settembre - 2 ottobre 2015), Accademia della Crusca, Firenze, pp. 101-112.
- DA = *Diccionario de americanismos* (2010), Asociación de Academias de la Lengua Española. Editorial Santillana, Lima.
- de Schryver G.M. (2003), “Lexicographers’ dreams in the electronic-dictionary age”, in *International Journal of Lexicography*, XVI, 2, pp. 143-199.
- DELI = *Dizionario Etimologico della Lingua Italiana* (1999), di Cortelazzo M. e Zolli P., seconda ed. a cura di Cortelazzo M. e Cortelazzo M. A., Zanichelli, Bologna.
- DIFIT (2008) = Stammerjohann H. et al. (a cura di), *Dizionario di italianismi in francese, inglese, tedesco*, Accademia della Crusca, Firenze (<https://difit.italianismi.org>; ora consultabile anche nel portale dell’OIM: [www.italianismi.org](http://www.italianismi.org)).
- 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>.
- Engelberg S., Meyer, P. (progetto in corso), *Lehnwortportal*, Institut für deutsche Sprache/IDS, Mannheim: <http://lwp.ids-mannheim.de/>.
- Furiassi C. (2015), “False borrowings and false Anglicisms: Issues in terminology”, in Furiassi C., Gottlieb H. (eds.), *Pseudo-English: false Anglicisms in Italian*, De Gruyter, Berlin-Boston, pp. 253-277.
- GDLI (1961-2002) = *Grande dizionario della lingua italiana*, fondato da Battaglia S., UTET, Torino.
- Gomez Gane Y, Julià Luna C. (in stampa), “Gli italianismi nel catalano: bilancio e prospettive”.
- Gomez Gane Y. (2012), *Gli italianismi nel catalano. Dizionario storico-etimologico*, Aracne, Roma.
- Görlach M. (2001), *A dictionary of European Anglicisms. A usage dictionary of Anglicisms in sixteen European languages*, Oxford University Press, Oxford.
- GRADIT (1999-2008) = *Grande dizionario italiano dell’uso*, diretto da De Mauro T., UTET, Torino.
- Heinz M. (a cura di) (2017), *Osservatorio degli italianismi nel mondo. Punti di partenza e nuovi orizzonti* (Atti dell’incontro OIM, Firenze, Villa Medicea di Castello, 20 giugno 2014), Accademia della Crusca, Firenze.
- Heinz M., Pizzoli L., Serianni L. (2022), “Lo studio dell’italiano a contatto con altre lingue: l’Osservatorio degli Italianismi nel Mondo (OIM)”, in Heinz M., Minniti Gónias D., Serianni L. (a cura di), *Gli italianismi nel neogreco e i suoi dialetti e in altre lingue*. Atti del Convegno “Italoellenica. Incontri sulla lingua e la traduzione” (1 novembre 2019), Università Nazionale e Kapodistrias di Atene, Atene, pp. 118-129.
- Lanfranchi E. (2014), “Il rinnovamento del vocabolario. Dalla crisi della carta alle potenzialità nuove dell’era digitale”, in Marazzini, 2014, pp. 191-230.
- Maconi L. (2016), “Retrodatazioni lessicali con Google Libri: opportunità e inganni della Rete”, in Marazzini, Maconi, 2016, pp. 73-94.
- Maconi L. (a cura di) (2020), *Laboratorio di ArchiDATA 2020. Retrodatazioni lessicali: storia di cose e di parole*, Accademia della Crusca, Firenze, pp. 209-238: <https://accademiadellacrusca.it/it/contenuti/quaderni-di-archidata/8054>.
- Marazzini C. (a cura di) (2014), *L’editoria italiana nell’era digitale. Tradizione e attualità*, Accademia della Crusca-Goware, Firenze.



- Marazzini C., Maconi L. (2018), “Il Vocabolario dinamico dell’italiano moderno rispetto ai linguaggi settoriali. Proposta di voce lessicografica per il redigendo VoDIM”, in *Italiano digitale*, VII, 4, pp. 100-119.
- Marazzini C., Maconi L. (a cura di) (2016), *L’italiano elettronico: vocabolari, corpora, archivi testuali e sonori* (Atti del convegno, Firenze, 6-8 novembre 2014), Accademia della Crusca, Firenze.
- Marello C. (2018), *Lexicografía digital y enseñanza de lenguas. Lessicografía digital e insegnamento delle lingue*, Memoria para optar al grado de doctor, Directora Natalia Borreguero Zuloaga, Universidad Complutense de Madrid, Facultad de Filología, Madrid.
- Matt L. (2004), “Retroduzioni di tecnicismi da titoli di pubblicazioni”, in *Studi di lessicografia italiana*, XXI, pp. 183-246.
- OED (1989<sup>2</sup>) = *Oxford English Dictionary*, edited by John Simpson and Edmund Weiner, Clarendon Press, Oxford: <https://www.oed.com/>.
- Pizzoli L. (2017), “Per un dizionario degli italianismi nel mondo: rilancio di un progetto”, in *Testi e linguaggi*, 11, 2017, pp. 171-182.
- Pizzoli L. (2019), “Italiano e italianismi nel mondo: osservazioni sulla ricerca di neologismi”, in Bombi R. (a cura di), *Italiano nel mondo. Per una nuova visione*, Forum editrice, Udine, pp. 151-58.
- Pizzoli L., Heinz M. (2022), “L’uso dei corpora elettronici per l’OIM (Osservatorio degli italianismi nel mondo)”, in Cresti E., Moneglia M. (a cura di), *Corpora e Studi Linguistici. Atti del LIV Congresso della Società di Linguistica Italiana* (online, 8-10 settembre 2021), Milano, Officinaventuno, pp. 397-417.
- Riccio A. (2016), *Gli strumenti per la ricerca linguistica. Corpora, dizionari e database*, Carocci, Roma.
- Roncaglia G. (2010), *La quarta rivoluzione. Sei lezioni sul futuro del libro*, Laterza, Roma-Bari.
- Roncaglia G. (2018), *L’età della frammentazione. Cultura del libro e scuola digitale*, Laterza, Roma-Bari, seconda ed. accresciuta, 2020).
- Roncaglia G., Ciotti F. (2000), *Il mondo digitale*, Laterza, Roma-Bari.
- Saržoska A., Ivanovska-Naskova R. (2021), “I neoitalianismi in macedone tra dizionari e testi”, in Saržoska A. (a cura di), *L’italianistica nel terzo millennio: le nuove sfide nelle ricerche linguistiche, letterarie e culturali*, Atti del convegno internazionale, Università di SS. Cirillo e Metodjo, Skopje, pp. 207-217.
- Serianni L. (2006), “Gli archivi elettronici e la lessicografia storica”, in Schweickard W. (a cura di), *Nuovi media e lessicografia storica*. Atti del colloquio in occasione del settantesimo compleanno di Max Pfister, Niemeyer, Tübingen, pp. 41-58.
- Serianni L. (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), Fondazione Cassamarca-Unione Latina, Treviso-Paris, pp. 19-41.
- Serianni L. (2016), “Ha un futuro il dizionario dell’uso?”, in Marazzini, Maconi, 2016, pp. 33-45.
- Serianni L. (2017), “L’italiano nel mondo. Intenti e propositi di un progetto editoriale sugli italianismi”, in Heinz, 2017, pp. 39-54.
- Serianni L. (2019), “L’italiano, gli italiani, i migranti”, in *Carte di viaggio*, 12, pp. 127-136.
- Sijs, N. van der (2010), *Nederlandse woorden wereldwijd*, SDU Uitgevers, Den Haag.
- Sijs N. van der (2015), *Uitleenwoordenbank*, Meertens Instituut, Amsterdam: [www.meertens.knaw.nl/uitleenwoordenbank/](http://www.meertens.knaw.nl/uitleenwoordenbank/).
- Simone R. (2016), “Il dizionario del futuro”, in Marazzini, Maconi, 2016, pp. 17-32.
- Stammerjohann H. (2010), “Italianismi”, in *Enciclopedia dell’italiano*, dir. da R. Simone et al., Istituto della Enciclopedia Treccani, Roma: [http://www.treccani.it/enciclopedia/italianismi\\_\(Enciclopedia-dell’Italiano\)/](http://www.treccani.it/enciclopedia/italianismi_(Enciclopedia-dell’Italiano)/).

Telve S. (2002), “Retrodatazioni di voci onomatopeiche e interietive”, in *Studi di lessicografia italiana*, XIX, pp. 229-277.

TLFi = *Trésor de la langue française informatisé*: <http://www.atilf.fr/tlfi>.

Zarra G. (2017), “Retrodatazioni di islamismi”, in *Studi linguistici italiani*, XLIII, 2, pp. 257-266.