

## Il Tesoro della Lingua Italiana delle Origini: Caratteristiche, problemi, futuro

Il *Tesoro della Lingua Italiana delle Origini (TLIO)* è il ben noto dizionario storico dell'italiano antico elaborato dall'Istituto Opera del Vocabolario Italiano del Consiglio Nazionale delle Ricerche, che lo pubblica in corso d'opera in rete, all'indirizzo [tlio.ovi.cnr.it](http://tlio.ovi.cnr.it) (raggiungibile anche dall'indirizzo dell'Istituto, [www.vocabolario.org](http://www.vocabolario.org)), oppure [www.oivi.cnr.it](http://www.oivi.cnr.it)). Il XXVII<sup>e</sup> Congrès international de linguistique et de philologie romanes mi dà un'importante occasione di fare il punto sui lavori di quest'opera davanti alla comunità scientifica internazionale, alla fine del mio lungo periodo di direzione dell'Istituto e del vocabolario, iniziato nel lontano 1992.

Tutte le risorse prodotte dall'OVI sono accessibili dalla prima pagina del sito dell'Istituto; qui si trovano un tasto per l'accesso al *TLIO*, un altro per l'accesso alle banche dati, e un terzo per scaricare dal sito il software *GATTO*, con il quale sono costruite le banche dati dell'Istituto. Questo software è offerto gratuitamente, per uso di ricerca senza scopo di lucro, nella versione più aggiornata, che comprende anche le ultime eventuali correzioni di errori di funzionamento riscontrati con l'uso (le funzioni più comuni sono da lungo tempo ben collaudate); è possibile scaricare invece il solo aggiornamento, e sono disponibili un manuale didattico e un piccolo corpus esemplificativo che serve per provare tutto ciò che è descritto nel manuale ed eseguire gli esercizi proposti.

Con il tasto «consulta il vocabolario» si accede alla pagina iniziale del *TLIO* (noncludo immagini alla comunicazione a stampa, perché tutto ciò che descrivo può essere facilmente riprodotto da ciascuno sul proprio computer). La ricerca più semplice si esegue scrivendo nella casella di ricerca l'entrata di una voce, o anche solo i suoi primi caratteri: nella colonna di sinistra della finestra compaiono le entrate di tutte le voci presenti nel *TLIO* che cominciano con tali caratteri; la voce viene visualizzata cliccando sull'entrata: per es. *fuoco*.

La voce si presenta con l'entrata e la categoria grammaticale in alto a sinistra («fuoco s.m.»), una fascia di tasti (che tutti insieme, con il contenuto cui danno accesso, costituiscono, nel gergo dell'OVI, l'intestazione), la prima definizione seguita dai contesti citati e, di seguito, le altre definizioni con le loro citazioni. L'ultimo tasto dell'intestazione («Tutto/Stampa») permette di scaricare la voce completa (incluso il contenuto dell'intestazione), stampandola su stampante o su file.

Ogni esempio è introdotto da un'abbreviazione bibliografica che può essere aperta con un clic, visualizzando la scheda relativa. Per es. la prima citazione in *fuoco* è da

«S. Francesco, Laudes, c. 1224 (assis.)»: la scheda che si apre contiene per esteso il nome dell'autore («Francesco d'Assisi (santo)»), il titolo del testo («Laudes creaturarum (o Cantico di Frate Sole)»), l'edizione da cui il testo è citato («Poeti del Duecento, a cura di Gianfranco Contini, Milano-Napoli, Ricciardi, 1960»), l'area linguistica “specificata” (una città) cui il testo è assegnato («assis.», cioè assisiate) e l'area linguistica “generica” (una regione) di cui questa fa parte («umbr.», umbro), la datazione («c[irca] 1224»), il genere («lir[ica]»), la forma («versi»), e il tipo (testo «originale»).

Nella stessa scheda bibliografica appena descritta, va segnalato il link «Doc. filol.» (documentazione filologica), aggiunto di recente. Per circa 700 testi dei quasi 2000 del corpus di riferimento del *TLIO*, questo permette di aprire, in una finestra successiva, uno o più files nei quali sono descritti gli interventi filologici apportati ai testi citati, come correzioni introdotte in base al controllo dei manoscritti (eseguito tipicamente nel caso di testi documentari tramandati da un solo manoscritto), sostituzioni di lezioni del testo con lezioni dell'apparato, e anche semplici correzioni di refusi: si vedano, per es., nella voce *fuoco*, le schede dei *Miracole de Roma*, citazione [2] sotto la prima definizione, che mostra la copia del testo a stampa con le correzioni manoscritte; del *Microzibaldone pisano*, cit. [8], che mostra un errata corrige, come anche le schede del *Glossario latino-aretino*, cit. [12], e del *Serapiom* padovano, cit. [13]; la scheda del *Libro de la destructione de Troya* napoletano, cit. [18], che mostra la cosiddetta ‘scheda verde’ che si elaborava, per la preparazione dei testi per la digitalizzazione, nella fase antica dei lavori del vocabolario; ecc. In effetti, è rimasto finora poco evidente per chi lo consulta che il *Corpus TLIO* non è un insieme di testi digitalizzati meccanicamente, ma è, almeno per una parte dei testi, il risultato di un attento lavoro filologico; ed era rimasto sensibile il problema che le differenze fra i testi del corpus e le edizioni di riferimento non fossero giustificate: il lettore in molti casi poteva anche prenderle per degli errori di digitalizzazione. Aver finalmente non solo riordinato, ma anche nuovamente verificato la documentazione filologica, e averla collegata al vocabolario (seguirà velocemente il collegamento anche alla banca dati) è un bel risultato, maturato negli ultimi due anni, che accresce notevolmente il valore dello strumento lessicografico.

Il primo tasto dell'intestazione, «lista forme», apre l'elenco di tutte le forme grafiche del corpus che almeno in una occorrenza sono riferibili all'entrata (in altre occorrenze possono essere omografi):

*feu, ffoco, ffuoco, fo, fò, foc', foccu, foch', fochi, focho, fochy, focchi, foco, fòco, focora, focu, fog, foghi, fogho, fogi, fogo, fogu, fogui, foki, fooco, fos, fouco, fous, fucho, fuco, fucu, fughi, fugo, fuoc', fuochi, fuocho, fuochu, fuochy, fuoco, fuocora, fuog, fuoghi, fuogho, fuogo, fuoko, fuoro*

Questa semplice lista alfabetica ha la funzione di stabilire il collegamento con la banca dati: cliccando l'icona raffigurante un gattino giallo a destra dell'entrata, nella colonna di sinistra, si avvia infatti una ricerca nel corpus di tutte le forme elencate. Il corpus nel quale è lanciata la ricerca è una versione particolare del *Corpus OVI*, detta *Corpus storico*: poiché la banca dati viene aggiornata con correzioni, con integrazioni

di nuovi testi e anche con sostituzioni di edizioni superate, le citazioni del *TLIO* potrebbero non trovare riscontro nella versione corrente; il *Corpus storico* conserva a questo fine tutti gli stati testuali precedenti, in modo che la ricerca vada a buon fine. Poiché questo è il suo unico fine, non può essere oggetto di ricerche condotte non a partire dal *TLIO*, ma direttamente nella banca dati, che naturalmente si consulta nella forma più aggiornata; perciò per chi consulta la banca dati rimane invisibile.

Il secondo tasto dell'intestazione apre la nota etimologica, che di regola consiste in un semplice rinvio al *LEI* (o *LEI-Germ*) o, in mancanza, ad un altro dizionario etimologico, preferibilmente il *DELI* o il nuovo *Etimologico* di Nocentini (Nocentini 2010) o il vecchio *DEI*, talvolta ad altri dizionari o saggi; nel redigere il *TLIO*, infatti, con eccezioni per qualche rara voce, non si fa ricerca etimologica, visto che è attivo il poderoso cantiere del *LEI*, del quale il *TLIO* è anzi una fonte per l'italiano antico<sup>1</sup>. Il terzo e quarto tasto visualizzano le prime attestazioni della voce: il terzo la prima attestazione in assoluto, il quarto le prime attestazioni distribuite per provenienza linguistica, divise in quattro sezioni: testi toscani (o anche toscanizzati, come i testi dei poeti della Scuola siciliana nei canzonieri toscani delle origini), con i quali vanno anche i corsi, testi settentrionali, testi mediani e meridionali, testi siciliani.

Il tasto successivo («Note ling[uiastiche]») dà accesso ad una serie di annotazioni linguistiche che nella maggior parte saranno inserite in una seconda fase della redazione; attualmente vi si annotano eventuali particolarità, come per es. i dubbi riguardanti l'accento o il timbro della vocale tonica, il fatto che un aggettivo o un sostantivo sia invariabile, il fatto che un verbo sia attestato solo in forma sostantivata, o solo pronominale; in futuro troverà posto qui una descrizione delle costruzioni dei verbi, che nella serie delle definizioni è sacrificata ad una descrizione esclusivamente, o almeno prevalentemente semantica. Sistematicamente, invece, nella forma attuale della redazione, è inserito in questo punto l'elenco alfabetico, di tutte le cosiddette 'polirematiche' (con termine mutuato dal *GDIU*) individuate ed esemplificate nella voce, con il rinvio alla definizione sotto cui sono trattate, numerose nella voce *fuoco* s.m. (l'interfaccia del *TLIO* in rete permette di passare da ogni elemento della lista alla definizione cliccando sul numero):

*a fuoco* 1.3; *a fuoco e armi* 4; *a fuoco e a fiamma* 4; *a fuoco e ferro* 4; *a fuoco lento* 3.2; *aggiungere fuoco a fuoco* 4; *a lento fuoco* 3.2; *al fuoco!* 3.5, 4; *al fuoco! al fuoco!* 3.5, 4; *andare a fuoco* 4; *battere fuoco* 3; *battere il fuoco* 3; *battesimo di fuoco* 5.1.1.1; *cacciare fuoco* 4; *cerchio del fuoco* 2.1; *cielo di fuoco* 2.1; *cerchio del fuoco* 2.1; *città di fuoco* 1.8; *condannare al fuoco* 3.5; *condannare nel fuoco* 3.5; *con fuoco e con ferro* 1.3.6.1; *con fuoco e ferro* 4; *covare il fuoco* 3; *crescere legna a fuoco* 3.1; *da fuoco* 3.2, 3.2.1; *dannare al fuoco* 3.5; *dare fuoco a* 3.2; *dare fuoco a cencio* 1; *dare il fuoco a* 1.3.6.1; *di fuoco* 1.1, 1.3.4, 1.5.1, 3.2, 4.1; *di fuoco e di coltello* 4; *donare fuoco* 3.5.1; *emisfero del fuoco* 2.1; *essere a fuoco* 1; *essere cenere senza fuoco* 1; *essere esaminato col fuoco* 1.3.2.1; *essere esaminato di fuoco* 1.3.2.1; *essere fuoco e fiamma* 1; *fare fuoco* 1; *fare fuoco di* 1; *fuoco di paglia* 3.1; *fuoco di sant'Antonio* 6; *fuoco! fuoco!* 4; *fuoco greresco* 4.1.1; *fuoco greco* 4.1.1; *fuoco sacro* 6; *fuoco selvatico* 6; *fuoco temperato* 4.1.1; *gettare fuoco* 4; *giudicare al fuoco* 3.5;

<sup>1</sup> Sul rapporto del *TLIO* con il *LEI* cf. Beltrami (2012).

*luogo del fuoco 3; mal fuoco l'arda 3.5; mandare al fuoco 3, 3.5; mettere a fuoco e a ferro 4; mettere a fuoco e a fiamma 4; mettere fuoco 1.3.5, 4; mettere il fuoco 4; mettere in fuoco 5; mettere legna al fuoco di 3.1; mettere legna nel fuoco di 3.1; mettere a fuoco 4; mettere in fuoco 4; mettere nel fuoco 3; morte di fuoco 3.5; pena del fuoco 3.5; per fuoco e ferro 4; per fuoco e per ferro 4; per fuoco e per coltello 4; porre fuoco 4; prendere fuoco 1; prova del fuoco 1.3.2.1; regione del fuoco 2.1; rosso fuoco 1.4; saggio del fuoco 1.3.2.1; sentenziare al fuoco 3.5; sfera del fuoco 2.1; sopra il fuoco lento 3.2; trarre dalla brace e gettare nel fuoco 3; trarre fuoco dall'acqua 1.2; trarre fuoco chiaro da fredda neve 1.2; trarre l'acqua dal fuoco 1.2; uscire dalla brace e rientrar nel fuoco 3*

Si tratta, nella terminologia del *TLIO*, dell'insieme delle 'locuzioni' (sintagmi che equivalgono a una parola per significato e funzione sintattica), delle 'espressioni fraseologiche' (espressioni il cui significato è diverso dalla somma degli elementi, con una qualche forma di scarto, o anche espressioni più o meno fisse e almeno potenzialmente ripetibili) e delle 'collocazioni' (sintagmi per qualsiasi ragione degni di essere evidenziati). L'etichettatura di ogni polirematica è inserita nella definizione (*locuz. nom.* = 'locuzione nominale', *locuz. verb.* = 'locuzione verbale' ecc.; *fras.* = 'espressione fraseologica'; nessuna etichetta per le collocazioni). L'elenco ha due funzioni: la prima, di presentare l'insieme delle polirematiche, che nella voce non sono poi raggruppate insieme, ma distribuite sotto le definizioni pertinenti; la seconda, di consentire al sistema di generare automaticamente una tabella delle polirematiche, accessibile dalla pagina 'tutto sul *TLIO*', che permette la ricerca a partire da ogni polirematica, e presenta tutte le voci in cui ognuna è contenuta. Il tasto «Note» contiene una diversa serie di annotazioni; in particolare, nella voce *fuoco*, vi si trova la registrazione di un toponimo (*Poggio del Fuoco*, citato da Matteo Villani). Il tasto «Lista definizioni» raggruppa le definizioni di primo e secondo livello, per consentire un esame sintetico del contenuto della voce, soprattutto nel caso di voci lunghe e molto articolate. Il tasto «Redattore» mostra la firma del primo estensore della voce, che prima di arrivare alla pubblicazione è passata, come d'uso, attraverso una serie di revisioni, che possono averla modificata anche considerevolmente; i nomi dei revisori sono però normalmente sottintesi, salvo nel caso in cui una voce sia stata ripresa in mano a distanza di tempo con modifiche rilevanti. Importante è la data dichiarata accanto al nome del redattore, che permette di determinare lo stato del corpus al momento della redazione (la voce, per es., non terrà conto di testi inseriti nel corpus dopo la redazione). Nel caso di revisioni che comportino un nuovo esame del corpus, questa data viene modificata di conseguenza. Se invece la revisione comporta solo interventi 'interni', senza ricorso al corpus, ciò viene registrato alla fine della voce, con l'annotazione [u.r. (= ultima revisione) *data*], che segnala ogni tipo di innovazioni, anche minime (come le correzioni di refusi), ed è ugualmente aggiornata quando viene aggiornata la data inserita nel campo 'redattore'.

Del *TLIO* sono state redatte a tutt'oggi circa 28.000 voci, delle quali, all'inizio di ottobre 2013, 26.500 sono edite in rete; all'incirca metà dell'opera. Con l'eccezione di un paio di anni di particolare difficoltà, l'accrescimento è stato costantemente di circa 2000 nuove voci per anno, che non sono poche, se si tiene conto del fatto che

ogni voce comporta uno studio di prima mano sui materiali della banca dati, prima di tutto, e poi anche su tutti gli strumenti di studio e di consultazione correnti (studio di prima mano significa dedurre le definizioni dall'interpretazione dei testi e non dai lessici, né dall'introspezione del redattore, e affrontare spesso problemi interpretativi complessi su testi difficili in una grande varietà di lingue dell'Italia medievale molto diverse tra loro). Un problema è dunque sempre stato quello di garantire all'impresa energie sufficienti, con una buona percentuale di collaboratori stabili, intorno ai quali organizzare il contributo di collaboratori occasionali; un altro, quello di garantire l'omogeneità di un lavoro svolto da un gruppo che necessariamente si modifica nei tempi lunghi dell'impresa. Questo, in particolare, comporta la necessità che tutte le voci passino almeno una volta attraverso la revisione di una o due persone che siano sempre le stesse, con un inevitabile collo di bottiglia che pone dei limiti obiettivi alla velocità di sviluppo: che però, in presenza di risorse adeguate, potrebbe ancora crescere fino almeno al doppio. Le risorse, appunto, sono un problema pressante per un'impresa di lungo periodo, difficilmente assimilabile alla tipologia dei progetti finanziati (si deve però riconoscere che il *TLIO* sta meglio, ed ha a tutt'oggi prospettive più favorevoli, di altre imprese lessicografiche analoghe, in Europa, che sono entrate in una fase di gravi ristrettezze economiche).

La banca dati dell'italiano antico, che è disponibile in rete per le ricerche degli studiosi dal 1998 nella versione creata dal consorzio *Italnet* e, abbandonata questa<sup>2</sup>, dal 2005 nella versione in *GattoWeb*, nell'ultimo aggiornamento (agosto 2013) contiene complessivamente 2314 testi, o si dovrebbe meglio dire unità bibliografiche o senz'altro files, perché ogni unità è un file che può contenere un poema, un trattato, una raccolta di poesie, un sonetto, un breve documento contabile. La quantità si esprime dunque meglio in termini di occorrenze totali, 23.160.300 in questa versione. Il criterio di base è di includere nella banca dati tutti i testi prodotti in qualunque varietà dell'italo-romanzo (esclusi sardo e friulano) entro il secolo XIV all'incirca e disponibili in edizioni affidabili. Il criterio cronologico posto all'inizio dei lavori era più stretto e più preciso, con limite al 1375 (morte di Boccaccio); l'idea attuale è di documentare con impegno all'esaustività tutto ciò che è databile entro il 1375, e di includere invece in modo non sistematico la documentazione posteriore, che, anche quando non fosse immediatamente pertinente per il *TLIO*, tornerà utile per gli sviluppi futuri del vocabolario storico. L'esaustività, ovviamente, è una tendenza; a partire da un corpus come è stato messo insieme in questi anni, che rappresenta in ogni caso tutto ciò che è significativo per una storia della lingua e del lessico ben documentata, significa continuare la ricerca e tener dietro agli sviluppi della filologia: edizioni di nuovi testi, nuove edizioni di testi già inclusi. In effetti il criterio di utilizzare edizioni affidabili, in astratto ovvio e del tutto cogente per un lavoro scientifico,

<sup>2</sup> Il corpus di *Italnet*, implementato in *PhiloLogic* (con consultazione solo per forme), è ancora accessibile nel sito dell'ARTFL di Chicago, ma dal 2005 non è stato più aggiornato né corretto. I corpora di cui l'ОВI si assume la responsabilità scientifica sono solo quelli accessibili attualmente dal suo sito, tutti in *GattoWeb*.

è quello che tiene meno, se l'idea è di esplorare il lessico della lingua e non quello di un corpus definito dalla casualità del progresso filologico (che casuale non è, naturalmente); il fatto che sul corpus si debba costruire il vocabolario impone di utilizzare anche testi editi male, senza i quali il vocabolario presenterebbe buchi vistosi (ne ho scritto in Beltrami 2010). Qualificante, per un'impresa come quella del *TLIO*, è saper gestire con attenzione e consapevolezza filologica la qualità variabile delle fonti, precisando ogni volta che sia pertinente l'affidabilità dei dati che vengono presentati, con particolare riguardo alle voci rare, per le quali l'affidabilità della fonte fa veramente differenza (cf. anche Beltrami 2011).

L'insieme dei testi (nel senso detto), raccolto e consultabile unitariamente nel *Corpus OVI dell'Italiano antico*, o semplicemente *Corpus OVI* ([gattoweb.ovi.cnr.it](http://gattoweb.ovi.cnr.it)), è diviso in due corpora, il *Corpus TLIO* e il *Corpus TLIO aggiuntivo*. Il primo, lemmatizzato, è il corpus di riferimento per la redazione del *TLIO* (nel senso che la voce del *TLIO* si redige prima di tutto su questo, allargando poi la ricerca ad altre fonti per integrarla, se opportuno); il secondo non è lemmatizzato, e contiene testi in un certo senso in attesa, da inserire in futuro nel *Corpus TLIO*, ma solo dopo averli lemmatizzati almeno quanto basta per non introdurre in esso forme non associate ad alcun lemma. La lemmatizzazione del *Corpus TLIO*, infatti, non è top-down (uno strumento fatto di lemmi, forme e regole di associazione da applicare alla lemmatizzazione delle singole occorrenze), ma bottom-up, dall'interpretazione dell'occorrenza singola all'associazione di forma e lemma al dizionario di macchina che propone suggerimenti per le associazioni successive. Per spiegare con un esempio, si consideri di lemmatizzare l'occorrenza *disnar* nel *Vulgare de elymosinis* di Bonvesin da la Riva, v. 746:

Lo bon segnor, vezando lo pover peregrin,  
Al so *disnar* l'invidha col so volt alegrin...<sup>3</sup>

Il lemmatizzatore (umano) legge il testo e lo interpreta, e registra 'manualmente', operando nell'ambiente di lemmatizzazione del software *GATTO*, che questa precisa occorrenza di *disnar* è un'occorrenza del sostantivo *desinare* (lemma) nella forma *disnar*; questa occorrenza viene perciò collegata al record «*disnar* | *desinare* | s.m.» del dizionario di macchina; se il record non esiste, questa operazione lo crea. In questo modo tutti i lemmi del dizionario di macchina sono creati a partire dall'interpretazione dei testi. Le ragioni di questo modo di operare, che fa sì che, fatalmente, solo una parte, anzi una minoranza delle occorrenze venga lemmatizzata<sup>4</sup>, dipendono dalla natura del corpus, fatto di numerosi sottoinsiemi di testi (sottocorpora) di lingue diverse quanto basta perché uno strumento automatico costruito per una di queste non funzioni per l'altra, e, tranne forse il toscano del secondo e terzo terzo del Trecento, troppo piccoli per giustificare l'elaborazione di tale strumento. Va però detto che all'*OVI* la ricerca su questo punto è andata avanti, e che la versione di *GATTO*

<sup>3</sup> «Il buon signore, vedendo il povero pellegrino, lo invita al suo pranzo con viso allegro».

<sup>4</sup> Nella versione del *Corpus TLIO* di agosto 2013 le occorrenze lemmatizzate sono 3.653.328 su 21.862.472.

attualmente in elaborazione (*GATTO 4*) includerà un lemmatizzatore ‘semiautomatico’ pensato per il trattamento di corpora di limitate dimensioni (per chiarire che cosa sia rimando a Iorio Fili 2010 e 2012). Tornando dunque alla lemmatizzazione del *Corpus TLIO*, a questo punto è possibile cercare il lemma «desinare s.m.» in due modi diversi: o per ottenere le sole occorrenze lemmatizzate, con risultati sicuri, ma parziali, o per ottenere tutte le occorrenze di tutte le forme associate a questo lemma, con risultati completi, ma che includono gli omografi: nel caso di *disnar* anche l’infinito del verbo, per esempio nella *Parafresi pavese del «Neminem laedi»*, cap. 13: «El fa bisogno altro cha `sti son de chioche chi vol *disnar* con Cristo in questo mondo e cenar conso in lo so’ alto regno»<sup>5</sup>. Per *fuoco*, dalla cui voce siamo partiti, la ricerca del lemma «fuoco s.m.» nelle sole occorrenze lemmatizzate dà nel *Corpus TLIO* 1496 risultati, la ricerca comprensiva delle occorrenze non lemmatizzate ne dà ben 21.982. Compromettente, con altre, l’omografia della forma *fo*, tre occorrenze nei *Sermoni subalpini* e solo in questi, con la voce del passato remoto di «essere», lo stesso che *fu*, che aggiunge 10.819 occorrenze. Il redattore del *TLIO* deve naturalmente esaminare tutto il materiale; da esempi come questo si vede che il compito non è sempre agevole (per questo, è auspicabile che in futuro il numero delle occorrenze lemmatizzate sia molto innalzato, con l’aiuto, se gli esperimenti daranno buon esito, anche del lemmatizzatore semiautomatico).

L’esplorazione attenta di tutto il materiale è insostituibile, ed è il primo e vero punto di novità del vocabolario; la redazione potrà tuttavia essere agevolata, in un prossimo futuro, da un nuovo strumento informatico, un editor attualmente in fase di prima sperimentazione. Il suo scopo è di automatizzare la maggior parte delle operazioni non interpretative, se non tutte, abbattendo in particolare tutti quegli errori formali, anche piccoli, che allungano notevolmente i tempi di revisione. Elementi come le abbreviazioni bibliografiche, le marche d’uso, le etichette delle polirematiche ecc. saranno caricati da liste gestite dal sistema, e collocati correttamente al loro posto; la struttura della voce (livelli delle definizioni, ordinamento degli esempi) sarà gestita automaticamente. Per l’evoluzione sostanziale del *TLIO*, nel futuro, due punti sono particolarmente pertinenti. Il primo è un’integrazione al campo etimologico della voce: il nuovo sistema consentirà di introdurre separatamente prefissi e suffissi, permettendo indicizzazione e ricerca anche in base a questi elementi. Il secondo è un’integrazione alle definizioni: si potrà introdurre, definizione per definizione, un’etichetta onomasiologica (che farà riferimento al sistema di Hallig/Wartburg 1963); dell’importanza di questa si è discusso a lungo nelle riunioni del progetto *COST Medioevo europeo*, rivolto alla nuova sfida dell’interoperabilità sulla rete di strumenti lessicografici informatizzati come il *TLIO*. Per quanto riguarda l’esecuzione effettiva, l’una e l’altra integrazione potranno essere effettuate in forma sistematica sull’insieme delle voci già redatte, mentre le voci redatte in seguito conterranno anche le nuove informazioni.

<sup>5</sup> «Ci vuol altro che questi suoni di campane se uno vuole pranzare con Cristo in questo mondo e cenare con lui nel suo alto regno».



Il *TLIO* è nato, già con le prime voci uscite alla fine del 1997, come un dizionario concepito per trarre profitto dalle possibilità di adattamento e di evoluzione offerte dalla rete anche alle voci già redatte, un'opera in grado di ristrutturarsi costantemente nel suo insieme per adeguarsi al progresso degli studi e far fronte alle nuove sfide, che allora, in un passato prossimo ma che la velocità dei nostri tempi rende già remoto, non erano nemmeno pensabili. Ciò che mi pare più importante è che alle sfide del futuro abbia continuato a far fronte con la tranquilla solidità dello studio filologico e lessicografico dei dati, che non cambia natura con il passare del tempo.

Università de Pisa et  
CNR-Opera del Vocabolario Italiano (Firenze)

Pietro G. BELTRAMI

## Referenze bibliografiche

- Beltrami, Pietro G., 2010. «Lessicografia e filologia in un dizionario storico dell'italiano antico», in: Ciociola, Claudio (ed.), *Storia della lingua e filologia. Atti del VII Convegno ASLI (Pisa-Firenze, 18-20 dicembre 2008)*, Firenze, Cesati, 235-248.
- Beltrami, Pietro G., 2011. «Il mito dell'edizione per lessicografi e il *Tesoro della Lingua Italiana delle Origini*», in: Overbeck, Anja/Schweickard, Wolfgang/Völker, Harald (ed.), *Lexicon, Varietät, Philologie. Romanistische Studien Günter Holtus sum 65. Geburtstag*, Berlin, de Gruyter, 341-349.
- Beltrami, Pietro G., 2012. «*Il TLIO e il LEI, compagni di viaggio*», in: Lubello, Sergio/Schweickard, Wolfgang (ed.), *Le nuove frontiere del LEI*, Wiesbaden, Reichert, 193-200.
- Hallig, Rudolf/Wartburg, Walther von, 1963. *Begriffssystem als Grundlage für die Lexikographie. Versuch eines Ordnungsschemas / Système raisonné des concepts pour servir de base à la lexicographie. Essai d'un schéma de classement*, Berlin, Akademie-Verlag.
- Iorio Fili, Domenico, 2010. «Un nuovo strumento di lemmatizzazione automatica per corpora testuali di ridotte dimensioni. Applicazione all'italiano antico», *Bollettino dell'Opera del Vocabolario Italiano*, XV, 367-391.
- Iorio Fili, Domenico, 2012. «Il lemmatizzatore semiautomatico di GATTO4», in: [Beltrami, Pietro G. (ed.)], *Dizionari e ricerca filologica. Atti della Giornata di Studi in memoria di Valentina Pollidori*, Alessandria, Edizioni dell'Orso (Supplemento III al *Bollettino dell'Opera del Vocabolario Italiano*), 2012, 41-56.
- MEDIOEVO EUROPEO. <[www.medioevoeuropeo.org](http://www.medioevoeuropeo.org)>
- Nocentini, Alberto, 2010. *L'etimologico. Vocabolario della lingua italiana con CD-Rom e online*, Firenze, Le Monnier.