

Pietro G. Beltrami
(Opera del Vocabolario Italiano)
Presentazione

A sei anni dalla prematura scomparsa di Valentina Pollidori (18 luglio 2004), i colleghi dell'Opera del Vocabolario Italiano hanno compiuto il loro desiderio, che era anche un impegno, di ricordarla come si conviene, parlando di ricerca nel campo di studi che è stato suo, in una giornata che si è tenuta il 26 ottobre 2010, con ampia e, piace ricordarlo, calda partecipazione di pubblico, nella sede dell'Istituto, che è la stessa dell'Accademia della Crusca. La Presidente dell'Accademia, Nicoletta Maraschio, ha aperto i lavori con il suo saluto, e con un ricordo affettuoso di Valentina. Questo volume raccoglie ora tutti gli interventi della giornata, rispettando la decisione di alcuni dei relatori di pubblicare il loro testo così com'è stato pronunciato, di altri di pubblicarlo in una versione rielaborata.

L'ordine di pubblicazione è lo stesso del programma del convegno, nel quale gli interventi dei ricercatori dell'OVI, che hanno fatto il punto sullo sviluppo dei lavori lessicografici e informatici dell'Istituto, si sono intrecciati, in tre blocchi, con quelli degli studiosi ospiti, che hanno portato in discussione argomenti diversi, correlati per materia o per metodo, o hanno commentato aspetti del lavoro dell'OVI. Edgar Radtke ha affrontato una questione di grammatica dell'italiano antico, tema di particolare attualità dopo la pubblicazione della grande *Grammatica dell'italiano antico* diretta da Giampaolo Salvi e Lorenzo Renzi, alla quale ha dato un contributo importante, ampiamente riconosciuto, la banca dati testuale dell'italiano antico dell'OVI. Sui materiali di questa, e non solo, si fonda la poderosa ricerca di Roberta Cella sull'ordine dei clitici nel fiorentino antico, della quale, come si può immaginare, la relazione orale ha dato solo una sintetica presentazione. José Antonio Pascual ha portato in discussione alcuni dei problemi che si affrontano nell'elaborazione dell'attesissimo nuovo dizionario storico dello spagnolo, di cui attualmente si è aperta la prima fase dei lavori alla Real Academia Española. Alessandro Pancheri ha portato l'attenzione sul 'prima' del *Corpus OVI dell'Italiano antico*, ovvero sui testi prelevati per questo dagli «Studi di filologia italiana». Fra questi, va messa in bella evidenza l'edizione delle poesie di Guido Orlandi curata da Valentina, uscita in versione definitiva nel 1995, ma che già circolava negli studi a partire dalla sua prima versio-

Dizionari e ricerca filologica. Atti della Giornata di studi in memoria di Valentina Pollidori, Firenze, 26 ottobre 2010 (Supplemento III al *Bollettino dell'Opera del Vocabolario Italiano*), Alessandria, Edizioni dell'Orso, 2012.

ne in tesi di laurea, del 1984. Messo a frutto già nell'edizione di Cavalcanti curata da Domenico De Robertis nel 1986, questo lavoro eccellente e, quando fu fatto, innovativo ha poi dimostrato nel tempo una tenuta corrispondente alla sua qualità: ne ha già parlato efficacemente Lino Leonardi nel commosso ricordo di Valentina pubblicato negli «Studi di lessicografia italiana» del 2004. Al convegno Leonardi ha parlato del 'dopo' del *Corpus OVI*, cioè di prospettive progettuali sulle quali potrà svilupparsi la collaborazione, già sperimentata, tra la Fondazione Ezio Franceschini, l'Opera del Vocabolario Italiano, l'Università di Siena e altre Istituzioni di ricerca.

Al di là dell'affetto per una persona che aveva tutte le qualità per farsi voler bene, rimasto vivo in tutti coloro che l'hanno conosciuta, e testimoniato al convegno da colleghi e amici, i lavori della giornata hanno messo in piena evidenza l'importanza del contributo di Valentina Pollidori agli studi e all'impresa del vocabolario, in cui entrò stabilmente, dopo avervi collaborato dall'esterno, alla fine del 1988, quando vinse il concorso di Ricercatore del CNR (vinse poi quello di Primo ricercatore nel 2001). Mi riferisco, naturalmente, alla già più volte citata banca dati dell'italiano antico, realizzazione, in forma elettronica, di quell'archivio lessicale del vocabolario che fin dal primissimo progetto era pensato non solo per servire di base alla redazione, ma anche per costituire uno strumento autonomo di ricerca. Valentina ne ha preso in mano l'elaborazione in tempi veramente difficili, quando ciò che esisteva era sostanzialmente quanto ne era stato preparato dall'Accademia della Crusca prima di passare la mano al CNR, in larga parte ancora codificato non solo in testi singoli, ma in un formato antico da cui si stentava a riconvertirlo, con una grande quantità di concordanze per forme in tabulati lemmatizzati a penna, da trasferire – come lemmatizzando di nuovo! – sui files di nuovo formato via via che venivano prodotti all'Istituto di Linguistica Computazionale di Pisa, che fino ad allora aveva fornito il supporto informatico al vocabolario. Chi non ha passato quegli anni all'OVI non può rendersi conto di quale specie di miracolo paresse, alla fine del 1995, il primo corpus montato in DBT, sul quale poté avviarsi l'anno successivo, ancora timidamente, la redazione del *TLIO*; che miracolo non era, ma il frutto di un lavoro appassionato, ostinato e competente, in cui i problemi di trattamento informatico e i problemi filologici e linguistici erano affrontati congiuntamente senza sconti. Lo si vede, in realtà solo parzialmente, nelle schede della bibliografia che integra il corpus, nelle quali si è riversata l'attività quotidiana di Valentina, che su quel primo corpus ha continuato a lavorare, proseguendo la ricerca bibliografica e filologica, sovrintendendo alla digitalizzazio-

ne dei testi da aggiungere, collaborando con chi da una parte si occupava della lemmatizzazione e chi dall'altro mandava avanti la redazione del vocabolario. Lo si vede, questo lavoro (come ricorda Domenico Iorio-Fili in questo stesso volume), anche nel contributo dato da Valentina allo sviluppo di un programma informatico adeguato ai problemi dell'OVI, quel *GATTO* che è arrivato a maturazione nel 1998, consentendo finalmente di utilizzare appieno il corpus lemmatizzato, e che oggi è uno dei gioielli dell'Istituto. Valentina non è purtroppo giunta a vedere il corpus lemmatizzato operativo in rete in *GattoWeb*, nel 2005; è stata, però, precedentemente, protagonista della prima fase in cui la banca dati è entrata in rete, senza lemmatizzazione, nella versione *ItalNet*, nel 1998, diventando subito uno strumento di ricerca essenziale per gli studi filologici e storico-linguistici di italiano antico, che ne sono stati profondamente rinnovati. Valentina sentì il raggiungimento di questo traguardo come una svolta nella vita dell'OVI, che finalmente si presentava sulla pubblica scena degli studi con un prodotto di riferimento, e 'finalmente esisteva', come mi disse all'indomani della presentazione di *ItalNet* (20 maggio 1998), e ne conservo vivo il ricordo.

Non c'era dunque modo migliore di ricordare Valentina, da parte dei colleghi dell'OVI, che fare il punto sull'evoluzione dei lavori. Pär Larson ha così parlato della costituzione e dell'incremento del corpus testuale, con particolare riguardo agli aspetti filologici che lo rendono molto speciale rispetto alla maggior parte dei numerosi archivi testuali esistenti in rete. Elena Artale ha presentato aspetti e problemi della lemmatizzazione, che nel caso del corpus dell'OVI è un vero e proprio lavoro di interpretazione dei testi, con in più tutti i problemi di coerenza e di consultabilità che pone un corpus così complesso, plurilingue e costituito da ogni tipo possibile di testo e di edizione. Rinviano al contributo congiunto dei due ricercatori per i contenuti, non voglio tacere che è un motivo di vanto per l'OVI la regolarità dell'aggiornamento del corpus in rete, ogni quattro mesi, con incrementi e correzioni e con lo sviluppo del lavoro di lemmatizzazione. Paolo Squillacioti ha esposto le caratteristiche della redazione del *TLIO*, dando un'idea molto concreta delle procedure di lavoro, degli strumenti adottati e degli obiettivi parziali che si perseguono, senza nascondere i punti di difficoltà. Su una fase particolare della redazione si è soffermata Rossella Mosti, cioè sulla cosiddetta pre-redazione, nella quale si forma la nomenclatura del *TLIO* sulla base, principalmente, del lemmario del corpus lemmatizzato, ma non solo (poiché il *TLIO* include anche materiali estranei al corpus). Una ragione per cui questa, come tutte le altre di cui si parla, è un'operazione delicata, è che il lavoro si

esercita su una situazione in movimento, con un corpus in evoluzione e una lemmatizzazione *in fieri*, e non può essere diversamente, perché se si fosse voluto ottenere preliminarmente l'assestamento completo dei materiali si sarebbe rinviata per un tempo indefinibile, forse ad oggi nemmeno ancora iniziata, la redazione del *TLIO*; invece il vocabolario è attualmente consultabile per una percentuale non lontana dalla metà, in una forma che si potrà integrare, ma è già ricca delle informazioni essenziali. E infine, gli strumenti informatici, senza i quali l'elaborazione delle banche dati e del vocabolario non sarebbe possibile: Domenico Iorio-Fili, autore del già citato e ben noto *GATTO*, ha esposto i principi di funzionamento di un lemmatizzatore semi-automatico (il contributo spiega in che senso) adatto a operare in presenza non di un corpus di grandi dimensioni, ulteriormente espandibile a piacere, in una sola lingua, ma di una pluralità di corpora di ridotte dimensioni in lingue diverse, per i quali l'elaborazione di lemmatizzatori automatici rischia di essere persino più onerosa della lemmatizzazione occorrenza per occorrenza, attualmente operata. È una bella sfida, ma il software principale è già stato sperimentato con successo, e il prodotto sarà operativo nella versione 4 di *GATTO*, in preparazione. Andrea Boccellari, autore dell'attuale interfaccia di pubblicazione del *TLIO* in rete, ha presentato un altro importante software in fase di sviluppo e sperimentazione, cioè il futuro sistema di redazione del *TLIO*, che permetterà, fra l'altro, di abbattere i tempi della revisione gestendo automaticamente tutti gli aspetti formali delle voci, e al quale sarà affiancato un nuovo sistema di pubblicazione in rete, con un forte potenziamento delle funzioni di ricerca. Le due presentazioni hanno insomma puntato sul futuro, a partire da una situazione già ricca di solidi risultati. Ed è stato questo lo spirito di tutto l'impegno dell'OVI in questa giornata: di dedicare a Valentina un ricordo che fosse non solo rimpianto, ma anche una testimonianza di come sul suo lavoro si sia continuato e si continui a costruire.