

Diccionarios históricos y fuentes documentales **Entrevista a Roberta Cella**

Rosalía Cotelo García
Instituto Historia de la Lengua de Cilengua

Roberta Cella nació en Savona. Estudió en la Università di Pisa donde se licenció en 1994 y se doctoró en 1999. Desde julio de este año hasta octubre de 2006 estuvo contratada como investigadora en el proyecto «Opera del Vocabolario Italiano» (Istituto del Consiglio Nazionale delle Ricerche de Florencia), a cargo de la lematización de la base de datos del *Tesoro della lingua italiana delle Origini* (TLIO), y de la redacción de más de 200 voces para ese repertorio. Desde noviembre de 2006 es profesora asociada de Lingüística italiana en la Università di Pisa, primero en la Facoltà di Lingue e letterature straniere, y actualmente en el Dipartimento di Filologia, letteratura e linguistica. Ha sido profesora invitada para impartir Filología Romanza en la Facoltà di Lettere e Filosofia de la Università degli Studi di Sassari (2005-2006), y también en la Cattedra Emile Lorand de Linguistica italiana de la Faculteit Letteren en Wijsbegeerte, Vrije Universiteit Brussel (2009-2010).

Sus líneas de investigación giran en torno a la lengua italiana antigua, desde un enfoque fonomorfológico y sintáctico, así como desde una perspectiva histórico-cultural, y también dirigida hacia la edición de textos de los siglos XIII y XIV tanto literarios como documentales; con un interés especial por el contacto entre lenguas, las fuentes documentales, la formación del tipo lingüístico nacional italiano, la lexicografía y la lengua literaria. Ha publicado las siguientes obras: *I gallicismi nei testi dell'italiano antico (dalle origini alla fine del sec. XIV)* (Florencia, Accademia della Crusca, 2003), *La documentazione Gallerani-Fini nell'Archivio di Stato di Gent (1304-1309)* (Florencia, SISMEL-Edizioni del Galluzzo, 2009), *La prosa*

narrativa. Dalle origini al Settecento (Bologna, il Mulino, 2013) y la *Storia dell'italiano* (Bologna, il Mulino, 2015); también se ha ocupado de Dante, Brunetto Latini, Giovanni Villani y Antonio Pucci, la lengua vulgar del género epistolar del siglo XIV, y los préstamos en los textos toscanos mercantiles escritos en la Italia transalpina, de los clíticos y la morfología verbal en el florentino del siglo XIV, de los sonetos cruzados entre Benedetto Varchi e Lionardo Salviati, de la gramática escolástica. Actualmente está trabajando en una monografía sobre la lengua vulgar de Francesco Petrarca.

ROSALÍA COTELO GARCÍA: *Desde su experiencia en el TLIO, ¿cuáles son, en su opinión, las bases fundamentales, teóricas y metodológicas, para un buen diccionario histórico?*

ROBERTA CELLA: Scherzando, potrei rispondere —anche se credo che qualcuno lo abbia già detto prima di me— che il miglior dizionario storico è quello che esiste: meglio qualche difetto d'impostazione o qualche errore interpretativo che la mancanza di uno strumento essenziale per tutte le discipline che si fondano sul testo e sulla critica delle fonti. Infatti, anche in epoca di grande disponibilità di informazione primaria solo uno strumento che classifichi il patrimonio lessicale e ne interpreti l'evoluzione e la stratificazione —un dizionario storico improntato a criteri scientifici moderni, appunto— può innescare il circolo virtuoso della filologia: i dizionari, anche quelli imperfetti, permettono di fare edizioni critiche migliori, e le edizioni critiche migliori correggono gli errori dei dizionari e a loro volta li migliorano (ma anche questo è già stato detto).

Ripetuto questo, credo che un buon dizionario storico debba fornire una *rappresentazione* affidabile della stratificazione lessicale di una tradizione linguistico-culturale, che significa non di una lingua —cosa che almeno nel mondo romanzo in epoca medievale non si dà—, ma del complesso delle varietà linguistiche che hanno popolato, inizialmente su un piano di parità, l'universo della scrittura di una zona geografica che poi, per ragioni soprattutto extralinguistiche, ha visto affermarsi una varietà sulle altre. Il raccordo tra la fase anteriore alla normazione e quella posteriore è uno degli snodi concettuali più delicati, con varie ricadute pratiche sull'attività lessicografica: cosa mettere a lemma? trattare insieme o separatamente gli esiti delle stesse basi latine attestati in zone geografiche non contigue? ricostruire lemmi «potenziali» per

raccordare gli esiti di una stessa base etimologica che però non ha avuto continuatori nella varietà poi divenuta nazionale? Volendo evitare tanto un approccio «teleologico», che della fase pre-norma vede solo ciò che poi è entrato nella lingua nazionale, quanto un approccio puramente descrittivo, che frammenta la storia linguistica in tanti piccoli rivoli locali, un buon dizionario storico è bene che si interroghi su questi problemi.

Quanto poi all'affidabilità della *rappresentazione*, è un'ovvietà ribadire l'importanza della qualità della documentazione da cui si parte, qualità che in primo luogo concerne il «rigore filologico» con cui sono trattati i dati. Ma bisogna intendersi sul significato di «rigore filologico»: già nel 1941, Giorgio Pasquali metteva in luce il «circolo vizioso» che si crea con l'«aspettare che i testi antichi siano tutti pubblicati criticamente» prima di avviare il dizionario, quando proprio il «vocabolario storico [è] il sussidio più necessario per pubblicare adeguatamente i testi». Escluso quindi che «rigore filologico» significhi allestire preventivamente tutte le edizioni critiche necessarie, ritengo invece che significhi stabilire il perimetro delle certezze (almeno relative) sulle quali edificare il dizionario e degli errori (almeno allo stato attuale degli studi) che ci consegna la tradizione. In buona sostanza, come si chiama «critica» e «filologicamente rigorosa» l'edizione che fornisce al lettore i dati sui quali è basata e illustra il modo in cui li interpreta, così sarà «filologicamente rigoroso» il dizionario che dichiarerà al lettore la differente attendibilità dei dati che classifica. E ancora più in sostanza: se il «rigore filologico» si identifica in ultima analisi con la *verificabilità* dei dati e la *riproducibilità* del processo interpretativo, allora anche un dizionario storico potrà esserlo senza dover attendere la rifondazione del sapere testuale e linguistico prima di avviarsi. Un corollario pratico: il criterio della verificabilità dei dati impone la chiarezza nella citazione dei testi, che devono essere facilmente identificabili e esplicitamente datati.

In secondo luogo, qualità della documentazione significa anche continuità cronologica e varietà tipologica: la prima perché assicura l'analisi a «maglie strette» che dà continuità alla rappresentazione lessicografica, la seconda perché garantisce spessore variazionale. Certo, per le fasi antiche della documentazione bisogna fare i conti con i limiti di ciò che la storia ci ha consegnato, ma per le fasi più recenti, quando si tratta di scegliere e selezionare quali testi includere nel corpus di riferimento, penso che i due criteri vadano tenuti presenti e ragionevolmente bilanciati.

Infine, un'ultima nota sulla *rappresentazione* della stratificazione lessicale, che comporta una certa dose di astrazione: sono dell'idea che

un buon dizionario storico debba appunto essere un dizionario, non una somma di glossari o un commento ai singoli contesti. Non deve spiegare tutto ciò che è stato scritto, ma mostrare con ragionevolezza ed equilibrio l'evoluzione del lessico nella *langue*.

R.C.G.: *¿Qué beneficios le ha aportado el trabajo lexicográfico para su concepción de la historia de la lengua?*

R.CELLA: Come lessicografa, mi sono occupata soprattutto della lemmatizzazione del corpus testuale (dalle Origini al 1375) sulla base del quale è redatto il TLIO. Mettere ordine e classificare decine di migliaia di forme grafiche (che erano circa 440.000 nel 2006, cioè fino a che me ne sono occupata) per milioni di occorrenze (un po' più di 21) appartenenti a tutti i volgari italiani in epoca anteriore alla normazione del tipo linguistico comune (quando cioè ciascuna varietà peninsulare viveva di vita propria e di proprie tradizioni scritte) credo che mi abbia insegnato a discriminare tra ciò che è tecnicamente *possibile* e ciò che è *probabile* che sia avvenuto nell'evoluzione delle parole, in relazione sia alla fonomorfologia sia alla semantica. Certo, ci sono anche altri modi per impararlo, ma trattare grandi quantità di dati affina una sensibilità che chiamerei statistica, e allena a dubitare della reale consistenza storica di hapax che magari sono pure attestati, materialmente, in qualche manoscritto, ma che sono frutto di errori o raffazzonamenti dei copisti accolti tali e quali dall'editore.

R.C.G.: *¿Qué opina de la calidad de los corpus italianos para el trabajo filológico? En su opinión, ¿cuáles son los aspectos que deben mejorarse, y qué puede aportar la filología en ese esfuerzo de perfeccionamiento?*

R.CELLA: Dal punto di vista qualitativo penso che il Corpus TLIO sia quanto di meglio oggi sia possibile fare in Italia, secondo lo stato degli studi e la disponibilità delle edizioni. Aspira alla completezza dei testi, a quelle «maglie strette» della documentazione su cui ciascun filologo e storico della lingua desidera lavorare. Per far questo raccoglie ciò che è a disposizione: dalle edizioni critiche affidabilissime, sia pluritestimoniali sia mono-testimoniali, alle edizioni meno o pochissimo affidabili ma significative per l'analisi linguistica, passando per tutti i gradi intermedi dell'accuratezza critica. Nel corso degli anni i ricercatori dell'«Opera del Vocabolario italiano» hanno provveduto a ricontrollare molte edizioni monotestimoniali sui manoscritti, apportando modifiche alla versione elettronica rispetto a quella a stampa: l'elenco delle modifiche è ora dis-

ponibile anche on-line, a partire dalle voci del TLIO, cliccando sulla sigla del testo e poi sul campo «doc. filolog.» («documentazione filologica») della scheda bibliografica che appare. L'indubbia disomogeneità dei materiali è trattata con giudizio, fornendo al lettore le informazioni necessarie per ponderare i dati. Per esempio, nel Corpus TLIO la sigla TS segnala i testi altamente affidabili per l'analisi linguistica, il carattere \$ prima e dopo l'abbreviazione del titolo indica un testo di cui esiste nello stesso corpus un'edizione più recente (presumibilmente migliore, comunque basata su altri principi); nel TLIO le parentesi acute (< >) segnalano un'edizione ritenuta non soddisfacente, la sigla *F* preposta all'esempio indica un'attestazione tratta da un testimone non incluso nel corpus ma personalmente verificato dal redattore (*f* un'attestazione da un testo non incluso nel corpus e non verificato).

Qualche sforzo va invece fatto sul versante quantitativo: il Corpus TLIO comprende testi precisamente datati fino al 1375 e genericamente datati al sec. XIV. Per le fasi successive, prima di arrivare alle basi di dati di testi del Novecento, non esiste nulla di sufficientemente organico e strutturato per servire alla lessicografia storica: esistono raccolte di testi letterari o di scritture correnti, insiemi di testi specialistici o tecnici, ma non sono raccordati tra loro e, soprattutto, sono allestiti per altri fini. Credo che sotto questo profilo sia necessario uno sforzo di fantasia e un qualche investimento di energie progettuali e finanziarie: costruire qualcosa di analogo al Corpus TLIO che comprenda testi dal Quattro all'Ottocento è impossibile, ma partendo dal raccordo e dall'integrazione operativa tra quanto già esiste si potrebbe gettare una prima base, certo da integrare nel tempo, per il dizionario dei secoli successivi al Trecento.

R.C.G.: *En el simposio internacional sobre «El futuro de los diccionarios en la era digital», que se celebró en noviembre de 2014 en la Real Academia Española, usted incidió en cómo las nuevas tecnologías generan mucha información pero que «solo los redactores fiables» sabrán gestionarla de manera adecuada. ¿Cuál es su perspectiva sobre el uso de recursos digitales en lexicografía histórica, y cómo pueden los redactores gestionarlos adecuadamente?*

R.CELLA: L'evoluzione della tecnologia informatica applicata alla linguistica ha radicalmente mutato non solo le procedure di lavoro, ma la qualità stessa dell'indagine. In primo luogo è oggi possibile pensare di allestire ampie basi testuali funzionali alla lessicografia, che non devono avere necessariamente corrispettivi cartacei e che sono implementabili

nel tempo. Il vantaggio è quello di poter disporre di dati di prima mano molto più ampi e distribuiti in modo più continuo nel tempo, che non obblighino il lessicografo ai tradizionali voli di fantasia per colmare le lacune nella documentazione. Il prezzo da pagare (oltre quello strettamente finanziario) è rappresentato dalla difficoltà di gestione dell'aumentata massa di informazioni: velocità del reperimento dei dati non significa infatti facilità di gestione degli stessi, ma tutto al contrario comporta la moltiplicazione del tempo necessario per verificarli, analizzarli, interpretarli e rappresentarli in un dizionario. Arriverei a sostenere che la tecnologia allunga, non abbrevia, la catena di operazioni interpretative di varia natura che lega il dato di partenza e la sua rappresentazione. Tutto ciò richiede una migliore *qualità* del lavoro umano: solo redattori più preparati, più consapevoli della multidimensionalità del lavoro lessicografico (che mobilita considerazioni filologiche e linguistiche), potranno mettere a frutto gli straordinari avanzamenti della tecnologia.

L'aumentata massa di dati da gestire amplia le possibilità d'errore già insite nella lessicografia condotta con strumenti tradizionali; la moltiplicazione dell'informazione facilita inoltre la propagazione degli errori, per esempio delle parole fantasma nate dalla passiva assunzione –da parte dell'editore e del lessicografo– della semplice materialità dell'attestazione non sottoposta a vaglio critico (cioè, riprendendo quanto dicevo prima, non esaminata alla luce della possibilità o della probabilità di avere avuto una consistenza e una circolazione reale). Se la percentuale d'errore è inscritta in tutte le grandi opere ed è in qualche misura ineliminabile, tanto in epoca digitale quanto lo era in passato, è comunque possibile limitarla, valutando criticamente i dati e dichiarando al lettore il loro grado di attendibilità. Come ho già detto, la bontà del dizionario non starà tanto nella perfezione del materiale di partenza, ma nella capacità dei redattori di ponderarlo alla luce della filologia e della storia linguistica e di rendere consapevole l'utilizzatore della qualità del dato fornito. Al proposito, non trovo parole migliori di quelle di Pietro Beltrami, direttore del TLIO dal 1992 al 2013: «è sempre necessaria, da parte del lessicografo, una continua valutazione critica dei dati, particolarmente, s'intende, nel caso delle attestazioni più antiche o più rare» per «sapersi districare fra le edizioni, mettendo a frutto criticamente ciò che ognuna gli offre». Ne approfitto per precisare che la mia esperienza lessicografica è legata al TLIO, al quale ho lavorato dal 1999 al 2006, ed è quindi inevitabile che il mio punto di vista sia condizionato dall'impostazione genera-

le che lo sorregge, e appaia molto simile –se non identico– a quello più volte espresso da Pietro Beltrami.

Detto tutto questo, che rischia di suonare un po' astratto, scendo sulla terra e provo a formulare qualche suggerimento operativo. Le «maglie strette» della documentazione sono fondamentali per assicurare o smentire la vitalità delle accezioni a bassa frequenza, ma diventano una camicia di forza per le parole ad altissima frequenza: l'eccesso di documentazione relativa al vocabolario fondamentale impedisce anche al redattore più esperto di fornirne una corretta rappresentazione in scala. Si può allora ricorrere agli «abbattimenti selettivi» della documentazione, costruendo dei sottoinsiemi testuali sufficientemente rappresentativi ma gestibili. Il rischio di perdere un esempio eccentrico o non canonico è ampiamente compensato dalla sicurezza di poter redigere la voce: come dicevo scherzando, il miglior dizionario storico è quello che esiste.

R.C.G.: *Este número de la revista Cuadernos se centra en el uso de fuentes documentales para el estudio de la historia del léxico. ¿Qué cree que aportan las fuentes documentales al historiador de la lengua, respecto a las tradicionales fuentes literarias, o a los corpus de prensa? ¿Cuáles son las claves, en su opinión, para el refinamiento en el uso de las fuentes documentales?*

R.CELLA: Le fonti documentarie (i «testi pratici», come li chiamiamo in italiano) sono la struttura portante dell'analisi linguistica: non perché siano registrazioni fedeli del parlato (cosa che non sono), ma perché di solito sono databili precisamente, non sono sottoposte ai processi di copia e, evenienza più importante per la lessicografia, perché coprono zone del lessico escluse dai testi letterari.

L'errore da evitare è quello, istintivo, di trattarli come fedeli rappresentazioni grafiche del parlato: non lo sono, perché anch'esse, in gradi variabili, rispondono a tradizioni scritte e a formularità linguistico-testuali (particolarmente evidenti nelle tenute contabili dei mercanti, che sono quelle che conosco meglio). Che in genere non siano sottoposte ai processi di copia non significa poi che le fonti documentarie non siano prive di errori: qualsiasi scrivente nel processo di formulazione linguistica e di resa scritta commette errori, e spesso non li individua neppure dopo la rilettura. Credo quindi che anche i testi documentari vadano trattati con le stesse cautele che si usano per i testi letterari, abbandonando la cieca fiducia in ciò che è materialmente attestato e cercando invece di interpretarli alla luce del contesto funzionale e d'uso per il quale sono stati prodotti: anche per loro vale ciò che Aurelio Roncaglia scriveva a

proposito dell'operazione editoriale: «a stabilire il testo [...] non è la materialità del manoscritto, ma la razionale interpretazione del manoscritto». Senza poi dimenticare che il lessico di tipo pratico ha bisogno, per essere inteso, di informazioni contestuali che, quanto più sono lontane nel tempo, tanto più hanno bisogno di essere indagate alla luce di altre discipline: se non sappiamo come era fatto un determinato manufatto, come era tessuta una certa stoffa, quale sistema si usava per tenere i conti, quali procedure giuridiche regolavano un determinato comportamento, l'indagine lessicografica rischia di fermarsi alla forma della parola perdendone la concretezza.

Note

Giorgio Pasquali, *Per un Tesoro della lingua italiana* [1941], in Michele Barbi, Giorgio Pasquali e Giovanni Nencioni, *Per un grande vocabolario storico della lingua italiana*, Firenze, Sansoni, 1957 (ristampa anastatica Firenze, Le Lettere –Accademia della Crusca, 2012), pp. 39-107, a p. 52. La riflessione di Pasquali è stata variamente ripresa, prima da Giovanni Nencioni nel 1955 (*Relazione all'Accademia della Crusca sul Vocabolario della lingua italiana* [1955], in Michele Barbi, Giorgio Pasquali e Giovanni Nencioni, *Per un grande vocabolario*, cit., pp. 111-152, alle pp. 118-120), e, in tempi recenti, da Pietro Beltrami. Quest'ultimo afferma, con la semplicità propria del vero, che «'spezzare il circolo vizioso' significa fare ognuno il proprio mestiere, i filologi le edizioni e i lessicografi i dizionari» (*Lessicografia e filologia in un dizionario storico dell'italiano antico*, in *Storia della lingua e filologia. Atti del VII Convegno ASLI, Pisa-Firenze, 18-20 dicembre 2008*, a cura di Claudio Ciociola, Firenze, Cesati, 2010, pp. 235-248, a p. 238).

In molti hanno trattato delle «parole fantasma» della lessicografia; cito solo, tra le ultime cose, Vittorio Formentin, *Filologia e lessicografia: due discipline in contatto*, in *La nascita del Vocabolario. Atti del Convegno di Studio per i quattrocento anni del Vocabolario della Crusca* (Udine, 12-13 marzo 2013), Padova, Esedra, 2014, pp. 193-209 (con esempi di errori prodotti dagli editori), e Pietro G. Beltrami, *Il mito dell'edizione per lessicografi e il Tesoro della Lingua Italiana delle Origini*, in Anja Overbeck, Wolfgang Schweickard, Harlad Völker (Hrsg.), *Lexicon, Varietät, Philologie. Romanistische Studien Günter Holtus zum 65.*

Geburtstag, Berlin, de Gruyter, 2011, pp. 341-349 (alle pp. 344-345 un paio di esempi di errori di copista; alle pp. 347-348 la citazione riportata poco dopo nel testo).

La citazione da Aurelio Roncaglia è tratta dall'intervento alla tavola rotonda in *La filologia romanza e i codici. Atti del Convegno* (Messina, 19-22 dicembre 1991), Messina, Sicania, 1993, vol. II, pp. 775-778, a p. 778.25 de junio, 2015.

