

Introduzione dei curatori

Il presente volume raccoglie gli Atti del “Convegno dottorale internazionale LinUD”, un evento organizzato dai dottorandi dei cicli xxxv e xxxvi dei curricula di Linguistica e Umanistica Digitale, tenutosi a Firenze nei giorni 27, 28 e 29 giugno 2022 presso il Dipartimento di Lettere e Filosofia (DILEF) dell’Università degli Studi di Firenze.

LinUD è nato dall’idea di realizzare non solo un convegno accademico, ma anche e soprattutto un luogo di incontro interdisciplinare e di scambio scientifico, finalizzato a condividere le novità e i progetti che circolano nel campo della Linguistica e delle *Digital Humanities*, favorendo così un dialogo tra i partecipanti, sia dottorandi che giovani ricercatori.

Dopo esserci interrogati a lungo sul tema del convegno, la scelta conclusiva è ricaduta sui *dati linguistici*, poiché costituiscono un comune denominatore per le discipline umanistiche in questione e un oggetto esplorabile secondo molteplici sfaccettature, parametri e sistemi d’indagine: si pensi, ad esempio, alla raccolta dei dati, alla definizione di rappresentatività del campione oppure agli strumenti usati per l’estrapolazione e l’analisi dei dati; accanto a ciò – in un’ottica più squisitamente teorica – non sono da trascurare i criteri e le metodologie con cui essi vengono esaminati o ancora le modalità con cui possono essere condivisi e diventare fruibili (specie alla luce delle nuove tecnologie che consentono di accelerare i processi e le fasi di trattamento dei dati, nonché di renderli disponibili in formato digitale). In altre parole, si è lasciato spazio al dato linguistico sotto larghi punti di vista, nella

I dati linguistici. Metodologie e strumenti della ricerca

fattispecie ai modi in cui questo può essere elicitato, trattato e condiviso – a tal proposito, la diversità degli approcci riflette la vastità e la complessità degli studi linguistici e dimostra come metodi e modalità di indagine varino con il variare delle scelte e degli obiettivi di ogni singola ricerca.

A una molteplicità di analisi e di risorse corrispondono, dunque, molteplici prospettive di ricerca, in grado di stimolare nuove idee e produrre continui interrogativi. Ed è proprio in questo contesto che si inserisce LinUD, nato come occasione per caldeggiare la condivisione di pratiche, metodologie e modelli di ricerca – sia nel solco di tradizioni già consolidate sia nel panorama di frontiere sperimentali –, al fine di promuovere un'interazione proficua tra le diverse aree di ricerca. Crediamo, infatti, che idee innovative e stimolanti possano sgorgare anche dall'avvicinamento di ambiti di ricerca che di solito restano lontani fra loro, con l'auspicio di contrastare quell'isolamento che talvolta caratterizza il percorso delle singole ricerche col rischio di renderle aride: tutto ciò nell'ottica di far arricchire vicendevolmente le branche di un ventaglio composito come quello della Linguistica e dell'Umanistica digitale.

Con questa raccolta si intende offrire il contributo dei relatori che hanno preso parte alle giornate del convegno e che hanno permesso di toccare una vasta gamma di argomenti riguardanti i dati linguistici. Come anticipato, gli articoli proposti si muovono in ambiti eterogenei: si passa dal mutamento e dalla variazione nelle lingue a questioni di semantica, dai linguaggi specialistici alla fraseologia e alla lessicografia, fino ad arrivare all'analisi testuale, all'utilizzo dei corpora e ai linguaggi di mark-up.

A occuparsi di morfo-sintassi dialettale GIUSY TRUNCELLITO, che indaga la MDO (“Marca Differenziale dell'Oggetto”) nei dialetti calabresi settentrionali dell'area Lausberg, conducendo un'analisi micro-variazionista che tiene conto dei diversi parametri semantici, quali l'animità, la definitezza e la specificità dell'oggetto, col fine ultimo di chiarire quali siano i contesti in cui, nelle diverse varietà, è ammessa oppure esclusa la MDO; appartiene allo stesso filone di ricerca l'articolo di CARMELINA TOSCANO, che indaga il fenomeno della pseudo-co-

ordinazione (PseCo) nella Calabria meridionale attraverso un'analisi micro-variazionista, esaminando i contesti semantico-sintattici in cui la codifica della PseCo è esclusiva e in quali altri invece co-esiste con altre strutture di modo finito diffuse in questi stessi territori; l'obiettivo del lavoro è quello di individuare i contesti prototipici di codifica della PseCo e i fattori che governano la co-ricorrenza o l'alternanza tra questa e le altre costruzioni diffuse nella stessa area.

Spostandoci sul versante del rapporto tra lingua e dialetto, troviamo lo studio di VALERIA GAROZZO ed ENZO SANTILLI, il cui obiettivo consiste nell'individuare, categorizzare e analizzare i fenomeni di contatto italiano-dialetto: il lavoro dei due studiosi prende le mosse da un'analisi delle produzioni orali di parlanti bilingui che si avvicinano il più possibile al parlato spontaneo prototipico.

Alla fraseologia manzoniana sono dedicati due contributi: quello di ERSILIA RUSSO, che ricostruisce il ruolo rivestito dalla cosiddetta "Crusca veronese" nella formazione della fraseologia dei *Promessi sposi* (nella fattispecie, le postille apportate sul vocabolario testimoniano l'attenzione che Alessandro Manzoni ha riservato alle unità fraseologiche, in quanto tessere particolarmente rilevanti per la restituzione di una lingua "viva e vera") e quello di IRENE RUMINE, in cui si approfondisce l'aspetto paremiologico del romanzo manzoniano (in continuità con il clima romantico, i proverbi sono visti come un serbatoio di cultura popolare, in grado di conferire una molteplicità di toni e di registri al testo; in particolare, vengono messe in evidenza quelle paremie che, dal *Fermo e Lucia* alla *Quarantana*, sono segnalate da glosse metalinguistiche che ne attestano un uso consapevole da parte dell'autore). Sempre incentrato sulla fraseologia, ma in relazione alla lingua slovena, l'articolo di EVA TRIVUNOVIC, che presenta le problematicità dello studio diacronico della fraseologia, occupandosi nello specifico del materiale ricavato dal *Dizionario della lingua letteraria slovena del XVI secolo*: dall'analisi delle fonti del dizionario (53 testi di natura prettamente religiosa) si evincono due caratteristiche proprie delle unità fraseologiche, ossia la stabilità e l'espressività.

Focalizzato sulla lingua delle donne il contributo di ALESSANDRA REA, che indaga il rapporto che intercorre tra le scrittrici e la lingua,

I dati linguistici. Metodologie e strumenti della ricerca

intesa come costruzione sociale, studiando la percezione di Sibilla Aleramo riguardo a una presunta lingua delle donne; l'attenzione dell'autrice si concentra sull'analisi di più variabili, come la variazione diastratica, diafasica e di genere. L'analisi si sviluppa attraverso una griglia di riferimento contenente, da un lato, i tratti ritenuti caratterizzanti, dall'altro alcuni fenomeni linguistici attraverso cui è misurata la presenza di quei tratti.

Di linguaggi specialistici si è occupato STEFANO MIANI, in particolare della formazione del lessico italiano della psicanalisi nella prima metà del XX secolo; lo studio presenta il corpus di testi su cui sono stati condotti gli spogli lessicali, i criteri di digitalizzazione del materiale raccolto e alcuni dei dati quantitativi ricavati grazie all'utilizzo delle piattaforme *Sketch Engine* e *Nodegoat*.

Sull'interferenza tra lingua comune e lingua specialistica si concentra il saggio di MATTEO MAZZONE, in particolar modo sulla ridefinizione semantica in chiave tecnica di voci d'uso quotidiano, mediante l'analisi semantica di alcuni lemmi tratti dallo spoglio lessicale effettuato su quella che è oggi considerata come la seconda traduzione cinquecentesca del *De architectura* di Vitruvio, databile tra il 1504 e il 1511, e contenuta nel manoscritto Ottoboniano latino 1653 della Biblioteca Apostolica Vaticana.

Diverse ricerche ruotano, invece, intorno alla linguistica dei *corpora*, strumenti declinati in base alle diverse modalità di indagine, a partire da quella di SIMONA TRILLOCCO, che analizza e indaga il ruolo delle pause nel parlato schizofrenico, mediante l'analisi del *Corpus Italiano di Parlato Patologico Schizofrenico* (CIPPS) nel quadro teorico della "Teoria della Lingua in Atto", mostrando, mediante un confronto con i dati relativi al parlato normofasico, i primi risultati ottenuti.

YASMINA MOUSSAID analizza il bilinguismo di parlanti italo-arabofoni, mediante una ricerca che combina l'analisi delle informazioni ricavate da interviste semi-strutturate – rivolte a parlanti bilingui – con l'analisi di *corpora* bilingui, con l'intento di indagare alcune caratteristiche, come ad esempio le scelte linguistiche dei parlanti coinvolti, i fenomeni di contatto e ulteriori questioni di tipo sociolinguistico e socioculturale.

Nel suo saggio CATERINA CACIOLI propone una discussione dei metodi di analisi dei dati linguistici in ottica contrastiva e nel campo della semantica, focalizzandosi sull'uso di diverse fonti di dati – *corpora* e video-risorse linguistiche –; con un caso di studio sui verbi di azione in inglese, svedese e italiano si propone l'uso combinato di dati linguistici, mostrando i vantaggi dell'unione della componente visiva e d'uso nello studio dei verbi e delle azioni.

LAURA OCCHIPINTI si occupa di esplorare l'importanza della comunicazione istituzionale durante la pandemia da Covid-19, evidenziando la necessità di chiarezza nei messaggi ufficiali e introducendo, in relazione al soggetto trattato, il corpus Cov-I-Cor, appositamente creato per raccogliere i testi istituzionali prodotti in periodo di crisi e analizzare l'uso del linguaggio istituzionale durante l'emergenza sanitaria.

Ancora in relazione ai *corpora* il contributo di LUCA LO RE, il quale propone un metodo per sviluppare un corpus multimodale dell'italiano parlato, seguendo la tradizione dei *corpora* compilati dal gruppo di ricerca LABLITA dell'Università di Firenze e proponendo la creazione di un corpus pilota, CORMIP (acronimo di *Corpus Multimodale dell'Italiano Parlato*), attraverso cui testare un nuovo tipo di approccio ai dati e preparare il terreno per ricerche future e comprendere la lingua come un sistema complesso, multimodale, che utilizza gesti e linguaggio parlato per esprimere significati.

Di paesaggio linguistico parla LORENZO CAMBI, che presenta una classificazione del Paesaggio Linguistico (PL) basata sull'analisi di dati ed esempi provenienti dal mercato di San Lorenzo a Firenze; la proposta si distingue per un approccio sperimentale che integra metodologie quantitative e qualitative, enfatizzando l'importanza di coinvolgere attivamente gli attori e gli osservatori del PL, ed esaminando, al contempo, l'impatto delle politiche linguistiche e il ruolo sempre più rilevante del Paesaggio Linguistico Virtuale e dei legami tra spazio fisico e virtuale nell'ambito della società contemporanea.

A chiudere la raccolta il contributo di BARBARA PATELLA, nel quale si delineano – e si illustrano con una serie di esempi – i principali criteri adottati per marcare in XML-TEI i maggiori dizionari metodici

I dati linguistici. Metodologie e strumenti della ricerca

ottocenteschi della lingua italiana, al fine di mostrare le modalità con cui tali strumenti lessicografici (particolarmente rilevanti per la storia della lingua italiana e della lessicografia), da prodotti cartacei “statici”, siano stati trasformati in repertori digitali “dinamici” grazie a un lavoro di informatizzazione per renderli interrogabili (in funzione di una piattaforma online predisposta per consultare questo tipo di dati).

Chiudiamo la nostra premessa con l’augurio che questo volume possa essere una risorsa utile per gli studiosi, costituendo un pungolo per nuove riflessioni e nuovi spazi di ricerca.

Intendiamo rivolgere i nostri ringraziamenti a tutti i relatori che hanno partecipato con grande entusiasmo alle giornate del Convegno con i loro interventi, condividendo le loro ricerche e affidandone una parte a questo volume. Un sincero ringraziamento va poi ai relatori invitati che, con i loro contenuti, hanno impreziosito l’occasione di scambio accademico: il professor Alessandro Lenci e il professor Franco Fanciullo.

Un doveroso grazie al Dipartimento di Lettere e Filosofia dell’Università degli Studi di Firenze (DILEF), in particolare al professor Pierluigi Minari e alla professoressa Paola Manni che, nel 2022 – rispettivamente in qualità di direttore del Dipartimento e di coordinatrice del Dottorato di ricerca in “Filologia, Letteratura italiana, Linguistica” –, hanno reso possibile l’organizzazione del Convegno LinUD.

Concludiamo ringraziando i membri del Comitato scientifico, specialmente il professor Marco Biffi (referente del curriculum di Umanistica digitale) e la professoressa Monica Ballerini (referente del curriculum internazionale di Linguistica) per averci guidato e coordinato nella progettazione del Convegno, nonché durante le fasi organizzative, e infine per averci sostenuto nella realizzazione degli Atti raccolti in questo volume.