

Università di Pisa

Corso di Laurea Magistrale in Letterature e filologie europee

Seminario di cultura digitale

FILOLOGIA E LINGUISTICA NELL'ERA DIGITALE

Noelia Maria Bambina

Matr. 438804

Anno Accademico 2014/2015

INDICE

1) Introduzione.	1
2) A cosa serve la filologia?	2
3) Approccio alla didattica: l'insegnamento delle lingue straniere.	4
4) Strumenti utili all'apprendimento linguistico: El molino de ideas.	7
5) Conclusioni.	9
6) Riferimenti bibliografici.	11

1) Introduzione

L'obiettivo del presente studio è quello di delineare le nuove prospettive che la recente evoluzione tecnologica offre a chi si occupa nello studio linguistico. In base a questo approccio affronterò le vie aperte dall'avvento della linguistica computazionale nell'ottica sia di un ampliamento delle possibilità professionali a disposizione del linguista-filologo, sia di un miglioramento delle risorse prettamente educative a cui un può avere accesso.

In concreto mi occuperò di presentare il progetto “El molino de ideas”, un'impresa che si occupa di creare strumenti linguistici, programmi che processino automaticamente testi e parole e risorse per gli insegnanti. Durante la mia recente esperienza Erasmus a Zaragoza ho avuto modo di entrare in contatto con il progetto “Zaragoza Lingüística”, denominazione del seminario permanente di investigazioni linguistiche promosso e organizzato dal gruppo Psylex (Psicobiolingüística del Léxico) dell' Università di Zaragoza. “Zaragoza Lingüística” si propone la divulgazione dell'investigazione linguistica, attraverso seminari e conferenze che affrontino un ampio campo tematico: dalle aree “tradizionali”, tali come la sintassi o il lessico delle lingue naturali, alle relazioni tra il linguaggio e la mente umana, la politica e la pianificazione linguistica, l'insegnamento e le dinamiche dell'apprendimento linguistico. All'interno di questo progetto ho assistito a una conferenza di Elena Álvarez (dell'impresa “El Molino de Ideas”) il cui titolo *Todo lo que un filólogo no sabe que sabe hacer* (Tutto ciò che un filologo non sa di saper fare) lascia di primo acchito quantomeno confusi. Lo scopo della linguista era quello di proporci, attraverso la sua personale esperienza, un'apertura verso le prospettive che la tecnologia poteva offrire a coloro il cui principale interesse è la lingua in tutte le sue sfaccettature e dimostrarci come la figura professionale del filologo e del linguista, aldilà dell'aula, ha ancora una sua utilità nel mondo “reale”. Al contempo il seminario di Andrea Angiolini *Vetera et nova? Profili professionali per le digital Humanities*, affrontando la nascita di nuove figure professionali all'interno del mondo editoriale mi ha offerto un diverso punto di vista sulla coniugazione tra sapere umanistico e le nuove tecnologie, portandomi a riflettere sulle implicazioni che la messa a disposizione di nuovi strumenti tecnologici potrebbe avere nel campo concreto dell'insegnamento di una lingua.

2) A cosa serve la filologia?

A cosa serve la filologia? Questa è una domanda che dobbiamo frequentemente affrontare i filologi. Tuttavia nell'epoca attuale gli studi linguistici hanno acquisito ampie prospettive grazie alla tecnologia. L'era informatica non ha comportato, come si poteva pensare, l'estinzione del ruolo dei linguisti o dei filologi, al contrario, ha reso il lavoro di questi specialisti più affascinante e necessario che mai.

(jK) la metamorfosi digitale è stata definita come la più importante svolta (dai tempi di Gutenberg) nel campo della comunicazione e quindi anche della produzione e della trasmissione dei testi letterari (jK) produce uno scenario inedito, rispetto al quale anche le scienze umane devono avviare un ripensamento del loro orizzonte epistemico. Se i filosofi della comunicazione e i psicologi cognitivi si interrogano ormai da anni sul nuovo sistema semiotico introdotto dalla diffusione dei testi digitali (elettronici, multimediali, ipertestuali) anche gli storici della letteratura contemporanea e soprattutto i filologi, sono chiamati a nuove sfide.¹

La linguistica computazionale è nata come disciplina di frontiera, ai margini sia del mondo umanistico, sia delle applicazioni informatiche

Sebbene facilitato dai progressi nel settore informatico e telematico, unitamente all'effetto catalizzante di Internet, sarebbe improprio spiegare lo sviluppo della disciplina solo in termini di fattori meramente tecnologici. In realtà, la linguistica computazionale possiede oggi una sua maturità metodologica nata dalla conquista di un preciso spazio di autonomia disciplinare anche alle sue anime originarie, l'indagine umanistico-letteraria e la ricerca informatica. Questa autonomia si contraddistingue per un nuovo e delicato equilibrio tra testo e computer. Le elaborazioni computazionali sono chiamate a rispettare la complessità, articolazione, e multidimensionalità del testo, e al tempo stesso quest'ultimo emerge come una risorsa di conoscenza che può essere gestita ed elaborata con le stesse tecniche, metodologie e strumenti che rappresentano lo stato dell'arte nella tecnologia dell'informazione.²

La disciplina nasce al principio della seconda metà del Novecento da due diversi filoni di ricerca che ne costruirono le fondamenta :

1 Alessio Giannanti, *Testi con-testi*, Cosenza, Luigi Pellegrini Editore, 2015, p. 178.

2 AA. VV, *Storia della scienza e linguistica computazionale. Sconfinamenti possibili*, Milano, Editore Franco Angeli, 2008, p. 104.

- a) gli studi di padre Roberto Busa intorno all'utilizzo del computer come mezzo di immagazzinamento e di analisi del testo (anni Cinquanta e Sessanta), il cui primo frutto fu il primo corpus elettronico dell'opera di Tommaso d'Aquino;
- b) l'applicazione di metodi formali alle suddette analisi testuali (anni Sessanta e Settanta), che procedette di pari passo ai progressi fatti nel settore dell'Intelligenza Artificiale.

Il neonato settore dell'Elaborazione del Linguaggio Naturale (in inglese Natural Language Processing), fu lungamente influenzato dai metodi deduttivi utilizzati dalla grammatica generativa di Noam Chomsky: questa teoria linguistica, prescindendo dall'uso, cerca di individuare delle regole astratte che descrivano la competenza della lingua posseduta da un parlante. Di fianco alla tradizione razionalista se ne sviluppò una seconda che dava invece la priorità alla raccolta di ingenti quantità di testi: i corpora. Questo indirizzo (prevalente in Gran Bretagna e negli Stati Uniti) sfruttava metodi statistici per estrarre dai testi regolarità linguistiche, punti di partenza per la descrizione della struttura del linguaggio. Il primo grande successo del "metodo empirista" è datato 1964, anno di comparsa del Brown Corpus di Francis e Kucera: si tratta del primo esempio di corpus nato per lo studio di una varietà linguistica contemporanea, in tal caso l'inglese americano. La crescita delle dimensioni dei corpora -sempre più fondamentali per verificare l'efficacia degli strumenti- e quella parallela del web -dalle infinite risorse testuali- hanno sancito il prevalere del secondo approccio. Infatti, con l'aumento delle dimensioni del materiale da analizzare, occorrevo tecnologie che potessero affrontare l'immensa varietà delle realizzazioni linguistiche. Oggi la Linguistica Computazionale può contare su numerosi gruppi di ricerca nel panorama scientifico internazionale; un buon numero di centri sono presenti anche sul territorio italiano (per esempio l'Istituto di Linguistica Computazionale del Consiglio Nazionale delle Ricerche, fondato da Antonio Zampolli.)³.

L'idea del filologo o del linguista come studioso barricato nel suo studio e lontano dalla vita "reale" ha subito una radicale rivoluzione, di cui la maggioranza, anche di coloro che si avvicinano a queste discipline, non è pienamente cosciente. Una delle prime aree di studio che uno studente interessato alle lingue si trova ad affrontare è la Fonetica e la Fonologia, con la conseguente memorizzazione dei tratti distintivi di vocali e consonanti, dell'alfabeto internazionale e dei diversi approcci teorici allo studio del primo livello del linguaggio. Al di là dell'interesse che queste discipline hanno in se stesse, talvolta siamo inconsapevoli del fatto che strumenti che oggi tutti utilizziamo si basano su

3 Wikipedia, l'enciclopedia libera, Linguistica computazionale, <https://it.wikipedia.org/wiki/Linguistica_computazionale>, luglio 2015

sistemi costruiti a partire da leggi fonetico-fonologiche. Un esempio molto semplice è costituito dai sistemi di sintesi vocale (che troviamo per esempio in un GPS). La sintesi vocale è una tecnica che riproduce artificialmente la voce umana. I sistemi di sintesi vocale sono noti anche come sistemi text-to-speech (TTS) per la loro possibilità di convertire il testo in parlato. Esistono inoltre sistemi in grado di convertire simboli fonetici in parlato⁴.

Un sistema o motore di sintesi vocale è composto da due parti: una front-end e una back-end. La parte front-end si occupa della conversione del testo in simboli fonetici mentre la parte back-end interpreta i simboli fonetici e li "legge", trasformandoli così in voce artificiale.

Il front-end prevede due funzioni chiave: per prima cosa, viene eseguita un'analisi del testo scritto per convertire tutti i numeri, le sigle e le abbreviazioni in parole per esteso (es. il testo '2' viene convertito in 'due'). Questa fase di pre-elaborazione viene definita come normalizzazione o classificazione del testo (in inglese: tokenization). La seconda funzione consiste nel convertire ogni parola nei suoi corrispondenti simboli fonetici e nell'eseguire l'analisi linguistica del testo rielaborato, suddividendolo in unità prosodiche, ossia in proposizioni, frasi e periodi. Il processo di assegnazione della trascrizione fonetica alle parole è chiamato conversione da testo a fonema o da grafema a fonema (in inglese text-to-phoneme, TTP). La trascrizione fonetica e le informazioni di prosodia combinate insieme costituiscono la rappresentazione linguistica simbolica che viene utilizzata dal back-end per la conversione in suoni di tali informazioni ossia per il processo di sintesi vero e proprio.⁵

La descrizione del meccanismo di base che sostiene i sistemi di sintesi vocale ci lascia largamente intuire l'importanza che lo studio linguistico può avere sulla vita pratica, concreta e quotidiana di un utente qualsiasi.

Concentrandoci sul caso concreto dell'educazione scolastica, la messa a disposizione di nuovi strumenti possono consentire all'insegnante di una lingua straniera di sviluppare nuove metodologie di approccio alla materia.

3) Approccio alla didattica: l'insegnamento delle lingue straniere.

Lo sviluppo tecnologico ha consentito una radicale trasformazione dell'editoria libraria, coinvolgendo e ampliando la figura stessa dell'editore testuale. L'oggetto libro è entrato nell'era della riproducibilità non più tecnica, ma digitale. Grazie ai media elettronici, la scrittura ha invaso gli spazi umani molto più di quanto fosse riuscito a fare il libro tradizionale. La didattica dovrebbe e

4 Jonathan Allen, *From Text to Speech: The MITalk system*, Cambridge University Press, 1987.

5 Wikipedia, l'enciclopedia libera, Sintesi vocale,

deve sfruttare le nuove possibilità aperte da questo facilitato accesso all'informazione. Nell'ambito di una riforma nel campo dell'insegnamento sarebbe sicuramente utile dotare gli stessi insegnanti della formazione necessaria per usufruire nella maniera più fruttifera possibile dei nuovi strumenti che ha a disposizione. La diffusione di Internet nelle istituzioni tradizionalmente dedicate alla formazione (scuole, università) è destinata a introdurre nel campo dell'insegnamento delle importanti innovazioni di tipo teorico, metodologico e organizzativo. In particolare nell'insegnamento delle lingue straniere, come in quello di altre discipline, la sensazione è che ci si trovi di fronte ad uno strumento in grado di modificare in profondità le strategie didattiche adottate finora, ma che non si sia ancora raggiunto un sufficiente grado di chiarezza su come sfruttarlo nel migliore dei modi.

La diffusione sempre maggiore della rete Internet in ambito accademico ha già introdotto radicali innovazioni nel campo della ricerca scientifica per ciò che riguarda la strutturazione, la conservazione, la gestione e la divulgazione delle conoscenze (World Wide Web, riviste e archivi elettronici, biblioteche, ecc.) e i contatti professionali tra gruppi di studiosi (posta elettronica, liste di discussione). Il settore che forse più di altri sta catalizzando in questo momento gli sforzi di docenti e ricercatori è quello della didattica: dopo una prima fase in cui la rete era considerata soprattutto un enorme contenitore di dati consultabili da tutto il mondo, se ne incominciano a intravedere anche le potenzialità legate al suo sfruttamento nel campo dell'apprendimento. Si moltiplicano i progetti che prevedono la realizzazione di strumenti didattici da diffondere in Internet~ l'impressione è comunque che non si sia ancora raggiunto un sufficiente grado di chiarezza sulle strategie e i metodi di apprendimento più adatti a questa nuova realtà comunicativa. L'innovatività della rete, la sua natura poliedrica e la sua globalità, intesa come estensione geografica degli utenti che può raggiungere ma anche come pluralità di strumenti di cui si serve, costituisce inizialmente un ostacolo alla formalizzazione di ipotesi didattiche teoriche da verificare poi in modo sperimentale. Tre sono le caratteristiche fondamentali della didattica fruibile attraverso la rete Internet:

- apprendimento a distanza (assenza della classe in senso fisico, mutamento del luogo e della situazione in cui l'apprendimento si produce)~
- autoapprendimento (caratteristica già sfruttata nell'insegnamento delle lingue straniere, ad esempio attraverso i laboratori linguistici)~
- mutamento dei ruoli dell'insegnante (da quello di fulcro del processo di apprendimento a quello di tutor, regista esterno della situazione comunicativa) e dello studente (che opera delle scelte, attraverso cui costruisce in prima persona il proprio apprendimento). L'evento didattico che ha luogo attraverso la rete può dunque essere definito come un autoapprendimento guidato a distanza, in cui entrano in gioco tre ruoli diversi: i docenti, gli apprendenti e i computer connessi in rete~ in esso, la funzione svolta dalla tecnologia non è solo quella di mero supporto a metodi tradizionali, ma di ristrutturazione radicale e profonda delle

strategie di apprendimento delle lingue straniere.⁶

A partire dagli anni Ottanta si è assistito a una rivalutazione dei metodi umanistico-affettivi che pongono al centro del processo didattico gli aspetti psicologici dell'apprendimento, sottolineando l'importanza del coinvolgimento degli alunni e dell'atmosfera di classe, e del pensiero di von Humboldt, secondo cui una lingua non si può insegnare, ma si possono creare le condizioni che ne favoriscono lo sviluppo spontaneo nell'apprendente. In questa prospettiva il ruolo di guida dell'insegnante si muta in quello di facilitatore dell'apprendimento dei singoli. Compito dell'insegnante di lingue è di aiutare gli allievi a sviluppare il potenziale di risorse che hanno e di condurli a una sempre maggiore autonomia, offrendo loro varie forme di supporto perché apprendano ad apprendere. Dagli anni Novanta la tendenza sempre più accentuata verso forme di autoapprendimento che tengano conto degli stili cognitivi individuali determina l'adozione di una didattica assistita dalle nuove tecnologie e basata su materiali multimediali. Si dovrebbe trattare di programmi per l'apprendimento linguistico su CD Rom che non consistano nella semplice trasposizione del libro di testo su supporto informatico, come è il caso di prodotti in commercio, in cui all'efficienza tecnica non sempre fa riscontro l'efficacia didattica.

Il computer multimediale, grazie a una delle sue prerogative principali, l'interattività, e alle molteplici modalità percettive che offre (testo, grafica, suono, animazioni, video), riassume in sé le potenzialità di più media, permettendo di svolgere attività diverse (ascoltare, parlare, leggere, scrivere, consultare banche dati in linea, quali dizionari e grammatiche) scegliendo percorsi individualizzati e seguendo i propri ritmi di apprendimento. La presenza sullo schermo di più linguaggi (iconico, filmico, linguistico) attiva i processi cognitivi di natura bimodale ed evidenzia la dimensione semiotica del linguaggio, rendendo espliciti gli aspetti dell'interazione che non possono emergere dal solo testo scritto, quali i gesti, la mimica, la postura, il ritmo, l'intonazione, il tono. Tali potenzialità promuovono una sempre maggiore autonomia nell'apprendimento e contribuiscono a sviluppare abilità che hanno una forte rilevanza sul piano cognitivo: quella euristica (soluzione di problemi per l'acquisizione di nuove conoscenze), quella esplorativa (capacità di trovare l'informazione cercata muovendosi all'interno di una rete di percorsi possibili), quella interazionale (negoiazione dei significati interagendo con l'insegnante o con i compagni) e quella metacognitiva (pervenire alla consapevolezza dei propri processi cognitivi) e di rispettare il proprio stile cognitivo.⁷

6 Stefania Spina, *L'uso di Internet nella didattica delle lingue straniere: problemi e prospettive*, <<http://elearning.unistrapg.it/dspace/bitstream/2447/57/1/spina00.pdf>>, luglio 2015, p. 1.

7 Katerin Katerinov, *L'acquisizione di una lingua seconda nell'era digitale*, <<http://www.italianistyka.uw.edu.pl/ds2180509.pdf>>, luglio 2015, pp. 2-4.

Sebbene la tecnologia offra un utilizzo sempre più semplice il rischio di una fruizione impropria o superficiale, soprattutto fra i giovani, è reale. L'uso acritico di internet provoca una sorta di analfabetismo mascherato e succube.

E' fondamentale adottare strategie che, integrate educhino alla competenza digitale e a quella linguistica e che vengano inserite in tutti i programmi scolastici e di formazione permanente. Insegnare le tecnologie attraverso le lingue straniere può diventare parte integrante dello sviluppo della conoscenza linguistica. In maniera analoga, insegnare la lingua straniera richiede lo sviluppo di competenze legate alla comunicazione digitale. Un approccio come il CLIL coniuga queste due esigenze. Il CLIL (Content an Language Integrated Learning, o "Apprendimento integrato di Lingua e contenuto) potenzia l'insegnamento/apprendimento di una seconda lingua utilizzandola come veicolo per l'apprendimento di altri contenuti. Nelle scuole che utilizzano il CLIL, una o più discipline vengono apprese impiegando una seconda lingua target, diversa da quella materna degli studenti.⁸

4. Strumenti utili all'apprendimento linguistico: El molino de ideas.

In questo capitolo mi piacerebbe soffermarmi sulla diffusione e messa a disposizione di strumenti informatici che facilitano l'insegnamento di una lingua (in particolar modo straniera) e che raramente sono presi in considerazione nel mondo scolastico odierno. In concreto presenterò il progetto del Molino de ideas, un'impresa spagnola che ha deciso di creare una pagina web con il fine di mettere a disposizione strumenti utili all'apprendimento della lingua spagnola.

Il Molino Labs lo componen unas herramientas que Molino de Ideas pone a disposición del usuario con el objetivo fundamental de crear un espacio lingüístico de investigación y aprendizaje, donde los usuarios puedan conocer el idioma a través de una perspectiva que ponga el acento en la parte lúdica, práctica e interactiva del aprendizaje de una lengua, en este caso del español. Así mismo, aunque estas aplicaciones se dirigen a cualquier persona interesada en el español y en jugar con el lenguaje, en el ámbito de la educación Molino Labs está pensado especialmente para suministrar herramientas tanto a docentes como a estudiantes, ya sea para la creación de materiales didácticos o la investigación, como para la práctica de ejercicios que sirvan de refuerzo para el estudio de la lengua española.⁹

8 Sara Ferrari, *Le tecnologie digitali per l'educazione linguistica*, Roma, EDUCatt, p. 101

9 Molino de Ideas, <<http://www.molinolabs.com/acerca.html#frecuentimetro>>, luglio 2015.

Un esempio del loro lavoro, e in generale di uno strumento utile a chi si avvicina allo studio di una lingua straniera è il “frecuentimetro”, un programma che analizza le parole secondo la frequenza della loro apparizione nel corpus testuale messo a disposizione dalle RAE (Real Academia Española) . Una parola sarà quindi considerata normale, abituale, inusuale, ecc... (diaria, normal, habitual). Come sappiamo un vocabolario può organizzarsi in base a criteri diversi: secondo un ordine alfabetico, in base ai significati, le etimologie, ecc... Uno strumento come il frecuentimetro può dare immediatamente l'idea a uno studente straniero sulla necessità di imparare prima gli usi di verbi apparentemente sinonimi (come “empezar” o “comenzar”) in base alla loro frequenza d'uso.

Un altro programma interessante è il Ciervantes¹⁰ che seguendo i livelli di riferimento per lo spagnolo fissati dall'Istituto Cervantes, secondo le indicazioni proposte dal Consiglio europeo nel sistema Europeo di Riferimento per le lingue straniere, classifica le parole secondo un livello base (A1-A2), uno intermedio (B1-B2), e uno avanzato (C1-C2). Questo è uno strumento pensato fondamentalmente per i docenti, dato che il suo obiettivo è facilitare l'acquisizione del lessico spagnolo in accordo al livello dello studente. Un'insegnante può così decidere di costruire un testo esattamente corrispondente alle necessità dei suoi alunni.

Infine un'altra risorsa fondamentale che Il Molino mette a disposizione è il Lematizador¹¹ La lemmatizzazione è un processo che consiste nell'identificazione della forma canonica di una parola, ovvero il suo lemma. Questa tecnica è molto usata per estrarre informazione da un testo in maniera automatica, dato che per il recupero dei dati è molto utile ridurre le varianti morfologiche che una parola può presentare e consentire così ai motori di ricerca di ottenere risultati migliori in maniera efficiente, permettendo una maggiore flessibilità nella ricerca e evitando l'eccessiva ridondanza nei risultati. Uno strumento come questo, nel campo dell'insegnamento della lingua straniera, permette all'alunno di scoprire e accedere nella maniera più semplice al lemma partendo dalla forma flessa con cui magari è entrato in contatto.

5. Conclusioni

L'introduzione delle nuove tecnologie pone a coloro che si interessano allo studio linguistico importanti quesiti sul loro ruolo professionale nell'era digitale e nel contempo

10 Molino de Ideas, < <http://www.molinolabs.com/ciervantes.html>>, luglio 2015.

11 Molino de Ideas, < <http://www.molinolabs.com/lematizador.html>>, luglio 2015.

stimola la creazione di strumenti educativi che supportino il lavoro dell'insegnante. Lo studente che inizia il suo approccio a queste discipline considerate da sempre prettamente umanistiche e relegate a un tipo di sapere specializzato e lontano dalla realtà di tutti i giorni, dovrebbe essere consapevole degli ampi spazi professionali offerti dall'avvento dell'epoca digitale. Nello stesso tempo l'insegnamento della lingua, in particolar modo della lingua straniera potrebbe giovare infinitamente di questa nuova realtà. La didattica dovrebbe allora aprirsi a nuove prospettive, non promuovendo un apprendimento di tipo passivo, ma al contrario sfruttando tutte le potenzialità che le risorse offerte dal Web possono dare.

6. Riferimenti bibliografici.

ALLEN, Jonathan, *From Text to Speech: The MITalk system*. Cambridge University Press, 1987.

AA. VV, *Storia della scienza e linguistica computazionale. Sconfinamenti possibili*, Milano, Editore Franco Angeli, 2008.

FERRARI, Sara, *Le tecnologie digitali per l'educazione linguistica*, Roma, EDUCatt, 2014.

GIANNANTI, Alessio, *Testi con-testi*, Cosenza, Luigi Pellegrini Editore, 2015.

KATERINOV, Katerin, *L'acquisizione di una lingua seconda nell'era digitale*, <<http://www.italianistyka.uw.edu.pl/ds2180509.pdf>> PDF.

SPINA, Stefania, *L'uso di Internet nella didattica delle lingue straniere: problemi e prospettive*, <<http://elearning.unistrapg.it/dspace/bitstream/2447/57/1/spina00.pdf>>, PDF.