

AperTO - Archivio Istituzionale Open Access dell'Università di Torino

**Didattica con i corpora di italiano per stranieri.**

**This is the author's manuscript**

*Original Citation:*

*Availability:*

This version is available <http://hdl.handle.net/2318/75633> since

*Terms of use:*

Open Access

Anyone can freely access the full text of works made available as "Open Access". Works made available under a Creative Commons license can be used according to the terms and conditions of said license. Use of all other works requires consent of the right holder (author or publisher) if not exempted from copyright protection by the applicable law.

(Article begins on next page)

## DIDATTICA CON I CORPORA DI ITALIANO PER STRANIERI

*Elisa Corino e Carla Marellò<sup>1</sup>*

### 1. INTRODUZIONE

Si è molto parlato nel ventennio appena trascorso delle applicazioni didattiche dei corpora di apprendenti (cfr. Leech, 1998 e più recentemente Granger, 2009). Tuttavia non si è fatto ancora veramente molto per convincere gli insegnanti dell'utilità dei *learner corpora* che finora sono stati consultati e apprezzati soprattutto da linguisti e da lessicografi.

I corpora d'apprendenti accessibili on line non sono molti: ICLE quello per l'inglese, ad esempio, va acquistato; quelli in rete (FRIDA per il francese e FALKO per il tedesco) si presentano con interfacce di interrogazione pensate per linguisti, non così amichevoli per insegnanti o allievi in situazione di autoapprendimento.

Nei paragrafi che seguono presentiamo il risultato di una serie di ricerche volte a (far) utilizzare davvero un corpus di apprendenti di italiano (VALICO<sup>2</sup>) a fini di didattica della morfosintassi italiana a stranieri e italiani e di autoformazione di docenti di italiano L1 e L2/LS.

#### 1.1. *Esigenze di linguistica educativa nel disegno del corpus*

Poiché tra i risultati possibili di un *learner corpus*, sia da un punto di vista teorico che applicato, c'è l'individuazione degli errori più frequenti commessi dagli apprendenti, è importante poter contare su un insieme di dati sociolinguistici che descrivano con precisione il profilo degli apprendenti. Un corpus di apprendenti slegato da una base di dati sociolinguistica è infinitamente meno utile; l'architettura di VALICO è stata appositamente studiata per dare all'utente la possibilità di interrogare i testi selezionando la L1 dell'apprendente, l'anno di studio della lingua italiana, la conoscenza di altre lingue, i contatti con l'italiano. Contiene anche informazioni sul tipo testuale e soprattutto consente di inquadrare ciascun studente in un gruppo (la classe alla quale l'apprendente appartiene o, allargando la portata della ricerca, l'insieme di tutti gli studenti con la sua

---

<sup>1</sup> Università di Torino. I §§ 1 e 3 sono di Carla Marellò, tutti gli altri sono di Elisa Corino.

<sup>2</sup> vedi [www.valico.org](http://www.valico.org) VALICO.org è un portale di ricerca linguistica e glottodidattica, di formazione per insegnanti e studenti di italiano LS nato intorno a VALICO (Varietà Apprendimento Lingua Italiana Corpus Online), il corpus di apprendenti creato presso l'Università di Torino nel 2003. Il sito si presenta come una finestra sul mondo dell'apprendimento/insegnamento dell'italiano e mira ad offrire ai linguisti un corpus annotato per parte del discorso e tipo testuale per fare ricerche sull'italiano L2, ai docenti strumenti utili per la riflessione sui modi di utilizzare nella didattica corpora in rete bilanciati e annotati, agli studenti esercizi con correzione immediata e spiegazione dell'errore commesso.

stessa lingua madre e un livello di studio comparabile che hanno svolto la stessa consegna).

VALICO quindi permette di creare dei sottocorpora di testi generati a partire dalla selezione dei dati sociolinguistici. Questi testi sono etichettati per parti del discorso e quindi interrogabili anche a partire da sequenze sintattiche, non solo a partire da lessemi. Ad esempio, è possibile analizzare l'uso del tempo imperfetto in testi di studenti germanofoni di vent'anni, che studiano italiano da tre anni, che conoscono l'inglese e che hanno scritto un testo (senza uso del dizionario) a partire da una certa vignetta (si veda Corino 2009). Ripetendo la stessa interrogazione (= *query*) su apprendenti di diversa lingua madre è possibile fare dei cauti confronti sulla varietà d'apprendimento dei due gruppi.

Da questi brevi accenni si comincia a intravedere la relazione tra corpora per apprendenti e linguistica applicata all'insegnamento e, di riflesso, tra corpora di apprendenti e didattica; si tratta di una relazione per così dire "a valle" del corpus perché presuppone che gli strumenti siano stati allestiti e che l'attività si concretizzi nel manipolare i contenuti per dare loro una forma funzionale alle esigenze della formazione di insegnanti e studenti di italiano LS/L2.

## 1.2. *Didattica universitaria nell'allestimento del corpus d'apprendenti*

Esiste una componente di formazione/didattica anche "a monte" del corpus e ha luogo durante le fasi di implementazione. VALICO in particolare si è rivelato strumento per la formazione di studenti di didattica delle lingue moderne e di didattica della lingua italiana come LS fin dai primi stadi della sua creazione. Gli studenti della Facoltà di Lingue e Letterature straniere dell'Università di Torino hanno infatti collaborato attivamente alla trascrizione dei testi, alla correzione manuale dell'attribuzione delle PoS, alla correzione degli elementi sconosciuti al programma di etichettatura, alla categorizzazione delle caratteristiche testuali dei materiali.

In questo modo gli studenti hanno avuto l'opportunità di scontrarsi "artificialmente" con problemi che dovranno affrontare quando saranno insegnanti, hanno constatato quali sono gli errori più frequenti commessi dagli apprendenti, ma sono anche stati in qualche modo obbligati a riflettere sulle cause di quegli errori, sulla variazione del lessico e delle strutture sintattiche usate per descrivere gli eventi, sulle strategie di evitamento e sul concetto di norma e accettabilità. Per questi studenti catalogare è stato un ripassare le regole della loro lingua materna guardandole attraverso gli errori dei parlanti nativi di altre lingue che loro hanno studiato come LS e conoscono piuttosto bene, e hanno così raggiunto un livello di maggiore consapevolezza e conoscenza della loro stessa L1. Hanno fatto linguistica applicata all'insegnamento in modo estremamente concreto perché, partiti dall'assunto che il contrasto con la lingua madre sia la fonte principale d'errore, hanno constatato con mano che:

- non sempre si verificano errori nei punti di maggior differenza fra lingue;
- tutti gli apprendenti di italiano fanno un certo tipo di errori indipendentemente dalla propria lingua madre;
- ci sono errori che gli apprendenti riconoscono agevolmente come errori nei test a scelta multipla, ma che in situazione di produzione continuano a fare anche a livelli molto alti di apprendimento.

## 2. ATTIVITÀ DIDATTICHE CON TESTI DEI CORPORA

Abbiamo anche fatto sperimentazione con testi di VINCA<sup>3</sup>, il corpus di italofoeni appaiato a VALICO, disordinandoli per farli poi ordinare dagli studenti sia italiani che stranieri (si veda Merlo, 2009) per la costruzione di una prova per studenti romeni di scuola secondaria superiore.

Ancora precedentemente avevamo fatto un'esperienza con universitari italiani e statunitensi, ponendo domande di comprensione e riflessione linguistica a partire da testi di VINCA abbinati ad una serie di vignette dal titolo "Amore". I testi sono stati scelti dagli studenti universitari torinesi come più rappresentativi e adatti a mostrare usi diversi dell'italiano (si veda Corino-Marelo, 2009).

Si può dire che questi sono stati tentativi di usare i corpora VALICO e VINCA nella didattica universitaria e nella formazione di docenti in servizio soprattutto a fini di attività di comprensione delle scritto e di riconoscimento della trama testuale (catene anaforiche, susseguirsi di tempi verbali, connettivi) e di riconoscimento dell'implicito (deissi personale e narratore, modi di interpretare ed esplicitare ciò che non è disegnato).

In questa sede descriveremo un'attività basata sul corpus d'apprendenti VALICO, ma di carattere più morfosintattico

### 2.1. *Quando i distrattori vengono dal corpus*

Un primo esperimento è stato svolto da Bárbara K. Pesce e Nelly M. Sirpa Caceres, (vedi Pesce, Sirpa Caceres, 2009) mentre nell'anno 2008 frequentavano corsi presso l'Università di Torino, con borse di studio "WWS - World Wide Style", cofinanziate dalla Fondazione CRT e dall'Ateneo torinese al fine di formarsi come insegnanti di italiano nei rispettivi paesi. In quanto parlanti native di spagnolo, rispettivamente argentino e boliviano, abbiamo pensato di far leva sulla loro conoscenza di tale lingua per far cercare nei testi di VALICO gli errori commessi da apprendenti l'italiano LS ispanofoni e quindi costruire dei test (si noti che per brevità chiamiamo test ogni singolo item di un test e non l'intera batteria necessaria per testare davvero un livello).

Successivamente abbiamo chiesto a studenti italofoeni della Facoltà di Lingue e letterature straniere dell'Università di Torino con una buona conoscenza del francese, dello spagnolo e dell'inglese di esaminare i testi di VALICO scritti da apprendenti francofoeni, ispanofoni e anglofoeni e di scegliere degli enunciati con errori frequenti per costruire su di essi esercizi a scelta multipla.

Questa prima fase di lavoro ha permesso di rilevare come il corpus aiuti molto a trovare errori e varianti di errori da usare per costruire i distrattori in test a scelta multipla in cui le quattro risposte sono: una costituita dall'enunciato in italiano standard corretto, una dall'enunciato scorretto tratto da VALICO e le altre due, anch'esse errate, create a partire da quello corretto e da quello errato, ma contenenti degli errori

---

<sup>3</sup> VINCA ([http://www.valico.org/vinca\\_CORPUS.html](http://www.valico.org/vinca_CORPUS.html)) è un corpus di italiano scritto prodotto da nativi. VINCA, inizialmente pensato per fungere da corpus di controllo per VALICO ha iniziato ora a vivere di vita propria, come supporto alla didattica applicata e alla riflessione sulla didattica.

leggermente diversi ispirati alla naturalezza dell'errore originario. È da questi due distrattori intermedi che si capisce se chi costruisce il test ha le idee chiare su cosa vuole davvero far notare a chi esegue l'esercizio.

Poiché avevamo chiesto di spiegare in modo semplice perché un enunciato era scorretto/parzialmente corretto/corretto, Pesce e Sirpa Caceres si sono dovute porre il problema del metalinguaggio da usare e quando si sono trovate a dover scrivere spiegazioni molto corpose si sono rese conto che era necessario alleggerire l'enunciato su cui era costruito il test a scelta multipla, perché troppo ricco di errori. È apparso chiaro che non per tutti gli errori è consigliabile dare un test a scelta multipla, e non per tutti gli errori vale davvero la pena dare spiegazioni o è facile trovare la spiegazione del perché sono degli errori.

Preparare questi esercizi, in concreto, si è rivelata una vera e propria cartina di tornasole di ciò che si ritiene importante sapere della riflessione metalinguistica dell'allievo.

## 2.2. *La costruzione degli esercizi*

Per costruire un buon test a scelta multipla bisogna enucleare il tipo di errore su cui si vuole focalizzare l'attenzione all'interno di un enunciato prodotto da apprendenti l'italiano L2 che ne contiene altri; bisogna lasciarne uno solo di tipo sintattico, al massimo insieme ad uno di tipo lessicale oppure ad uno di accordo morfologico. Non è una procedura banale costruire scelte multiple plausibili; è un percorso a ritroso nell'applicazione delle regole; è difficile per chi non è madrelingua, ma anche per un madrelingua con scarsa abitudine a valutare il rapporto forma/funzione.

È chiaro che gli errori sono elementi complessi, interpretabili da più punti di vista e a livelli diversi. Alla scelta delle frasi da inserire nelle batterie di esercizi è seguita un'opera di normalizzazione degli errori con la correzione di quelli che non erano funzionali e l'eliminazione o la sostituzione di parti troppo complesse o incomprensibili se decontestualizzate. Utilizzare enunciati che sono contemporaneamente errati sul piano ortografico, lessicale o sintattico confonde ancora di più l'apprendente, che può perdere di vista l'errore più importante "bersaglio" del test.

Focalizzato il distrattore, l'attenzione viene posta sulla costruzione di altre due risposte più o meno errate per accompagnare quella corretta. Per scegliere quale errore utilizzare in ciascuna risposta è stato interrogato nuovamente il corpus alla ricerca di errori diversi commessi in presenza di parti dello stesso distrattore che abbiamo scelto come punto di partenza per il nostro esercizio. Se, per esempio, come terreno di esercizio è stata scelta la costruzione *cercare di*, per stabilire quali risposte sbagliate aggiungere all'esercizio si è nuovamente inserito nel corpus il lemma *cercare*, per vedere in quali altri contesti appariva e quali errori ad esso correlati avevano commesso in precedenza altri studenti. Il risultato è quindi, a partire dall'enunciato:

*“Quando c'è il tempo, bisogna a cercare di approfondire gli argomenti che ci interessano”.*

la costruzione di un esercizio ((Livello A2 livello intermedio o B1), con gli item che seguono:

- A. Quando c'è tempo, bisogna cercare di approfondire gli argomenti che ci interessano.
- B. Quando c'è tempo, bisogna cercare ad approfondire gli argomenti che ci interessano.
- C. Quando c'è tempo, bisogna cercare per approfondire gli argomenti che ci interessano.
- D. Quando c'è tempo, bisogna cercare da approfondire gli argomenti che ci interessano.

Per stimolare gli studenti e i docenti in formazione a chiedersi quali errori sono più gravi di altri abbiamo chiesto di attribuire punteggi diversi a seconda della gravità degli errori contenuti nelle frasi scelte come risposte. Spesso, infatti, dietro una risposta apparentemente errata c'è comunque un lavoro di analisi e di scelta da parte dell'apprendente che, anche se non premiato, dovrebbe essere se non altro riconosciuto in modo positivo.

Un errore di accordo soggetto-verbo, non riconosciuto da studenti di livello intermedio, è sicuramente più preoccupante della scelta di un tempo sbagliato. Allo stesso modo, ci sentiamo in dovere di penalizzare più la scelta dell'ausiliare errato in verbi importanti utilizzati fin dalle prime lezioni del livello A1 (principiante) che non una forma riflessiva incompleta. L'esercizio proposto sopra è stato, ad esempio, corredato dei seguenti valori

- A. 4 punti. Risposta esatta *cercare* regge *di* + infinito.
- B. 0 punti. Al livello B1 si deve sapere che *cercare* regge *di* + infinito.
- C. 0 punti. Al livello B1 si deve sapere che *cercare* regge *di* + infinito. Interferenza di *to look for*, grave se si considera che *cercare* in questo contesto è *to try*.
- D. 0 punti. Al livello B1 si deve sapere che *cercare* regge *di* + infinito.

Nella costruzione di un test, in considerazione delle caratteristiche dell'errore "bersaglio" può essere opportuno riflettere anche su quale sia il tipo di esercizio (a scelta multipla, di completamento, di sostituzione, ecc.) più adeguato da proporre all'apprendente. Ad esempio, l'esercizio sopra riportato è costruito su un tipo di errore per cui risulterebbe adeguato anche un esercizio di completamento (senza la soluzione); gli errori per cui risulta più adeguato un esercizio a scelta multipla sono, ad esempio, quelli riguardanti la posizione dei vari sintagmi fra di loro nella frase. Ecco un esempio di test per studenti di Livello B1:

*"Tutte le persone sono spaventate perché non sapevano che cosa è successo".*

- A. Tutte le persone sono spaventate perché non sanno che cosa succede.  
(punti 4).
- B. Tutte le persone erano spaventate perché non sanno che cosa era successo.  
(punti 2).
- C. Tutte le persone erano spaventate perché non sapevano che cosa era successo.  
(punti 4).
- D. Tutte le persone sono spaventate perché non sanno che cosa successe.  
(punti 3).

In questo esercizio sono state date due soluzioni corrette, la A. e la C., motivate, una dal presente e l'altra dall'imperfetto dell'indicativo. Alla frase B., invece, abbiamo dato, nell'attribuzione del punteggio in funzione di quanto accettabile a livello B1, 2 punti e alla frase D. abbiamo attribuito ben 3 punti perché è corretta salvo che per l'ultimo verbo posto al passato remoto anziché al presente (errore in parte accettabile al livello B1).

Non c'è qui spazio per affrontare la questione se gli esercizi per “distrarre” gli apprendenti possano essere uguali per tutti gli apprendenti di un certo livello, età, ecc. o debbano essere costruiti tenendo conto delle specifiche differenze tra la loro lingua madre e l'italiano e se, quindi, ad esempio, gli esercizi per apprendenti francofoni siano somministrabili o meno ad apprendenti anglofoni e viceversa. Altro è prevedere una esercitazione intensiva e mirata per un apprendente, ad esempio francofono, che sceglie come corretta la risposta che contiene l'errore presente nell'enunciato di un francofono tratto dal corpus di VALICO: questo vuol dire che non è ancora in grado di riconoscere quale sia la forma corretta dell'enunciato proposto e che perciò necessita di ulteriori esercizi sulla forma linguistica che non percepisce come errata.

Come fa osservare Colombo (2009) spesso gli elementi che non sembrano degni di nota svelano problemi inaspettati ed evidenziano punti caldi della grammatica dell'apprendente che l'insegnante non avrebbe mai associato ad una possibile difficoltà e che possono essere argomento di una discussione, di una comune riflessione grammaticale tra l'insegnante e gli apprendenti. Si veda, ad esempio la discussione riportata da Colombo, e presente in formato audio sul sito di VALICO, che ha come centro un esercizio sul pronome complemento oggetto diretto che richiede di riflettere tanto sulla morfologia quanto sul significato. Questa discussione mostra come i test a scelta multipla possono essere utilizzati per fare riflessione grammaticale oltre la frase semplice e permette di valutare il tempo richiesto da una riflessione seria che coinvolga tutti gli studenti. Per questa ragione, pur essendo molto favorevoli a discussioni di questo tipo, pensiamo siano da fare in gruppo classe solo per argomenti veramente cruciali e ci si sta orientando a una fruizione soprattutto individuale e al computer di batterie di esercizi corredati di una spiegazione degli errori. Attualmente nel sito c'è un piccolo saggio di test a scelta multipla costruiti in modo da indurre i docenti e gli studenti a riflettere sulla grammatica italiana attraverso una visualizzazione graduale, non immediata del valore delle risposte

### 3. CONCLUSIONI IN ITINERE

Fra i vantaggi del far sperimentare ciò che si può fare con un corpus d'apprendenti annoveriamo:

- a. l'allenamento alla consultazione di corpora on line più vasti;
- b. lo stimolo al confronto fra scritti di italiani e scritti di stranieri di diversa lingua madre;
- c. l'impulso alla consultazione di grammatiche e relativa scoperta dell'assenza di certe regole;
- d. la stesura di “pillole” di grammatica che portano alla riflessione sull'uso scritto e sull'uso orale in italiano;

- e. la maggior consapevolezza dei diversi piani dell'analisi linguistica a partire dall'errore.

Quando questi test con spiegazione saranno in rete potranno servire per l'autoapprendimento di studenti e per l'autoformazione dei docenti: per il momento ne abbiamo testato l'efficacia attraverso interviste orali ad apprendenti universitari in Spagna e Francia. Ha riscosso gran favore l'idea del punteggio a scalare; variamente apprezzata è la spiegazione dell'errore, accusata a volte di essere troppo tecnica. Un passo che stiamo intraprendendo è perciò far scrivere direttamente dagli studenti le minispiegazioni dell'errore.

## BIBLIOGRAFIA

- Colombo O. (2009), "La discussione orale sui distrattori: l'esempio di studenti francofoni", in Corino E., Marelo C. (a cura di), *VALICO. Studi di linguistica e didattica*. Perugia, Guerra, pp. 217-223.
- Corino E. (2009), "L'altro giorno al lavoro mi era annoiato e così cominciavo di sognare... L'uso di passato prossimo e imperfetto in apprendenti germanofoni di italiano LS", in Atti del XLI Congresso SLI, *Alloglossie e comunità alloglotte nell'Italia contemporanea. Teorie, applicazioni e descrizioni, prospettive*. Pescara 27-29 Settembre 2007.
- Corino E., Marelo C. (2009a), "Elicitare scritti a partire da storie disegnate: il corpus di apprendenti valico", in Andorno C., Rastelli S. (a cura di), *Corpora di Italiano L2: Tecnologie, metodi, spunti teorici*, Perugia, Guerra, pp. 113-138.
- Corino E., Marelo C. (2009b) (a cura di), *VALICO. Studi di linguistica e didattica*. Perugia, Guerra.
- Leech G. (1998), "Preface: Learner corpora: what they are and what can be done with them", in Granger S. (ed.) *Learner English on Computer*, London & New York, Addison Wesley Longman, pp. xiv-xx.
- Granger S. (2009), "The contribution of learner corpora to second language acquisition and foreign language teaching: A critical evaluation" in Aijmer, K. (ed.), *Corpora and Language Teaching*, Amsterdam & Philadelphia, Benjamins, pp. 13-32.
- Merlo R. (2009), "Scienza multilingue per romeno fonici", in Marelo C. (a cura di) *Scienza multilingue. L'italiano disciplinare attraverso la lingua madre dello studente straniero*, Perugia, Guerra edizioni, pp. 159-223.
- Pesce B. K., Sirpa Caceres N. M. (2009), "Distrattori per apprendenti ispanofoni di italiano", in E. Corino C. Marelo Corino E., Marelo C. (a cura di), *VALICO. Studi di linguistica e didattica*. Perugia, Guerra, pp. 195-208.