

Proposte didattiche attorno ad un corpus di testi

Internazionale è come il maiale: non si butta via niente!

Francesco Caviglia,
Romansk Institut,
Aarhus Universitet,
Danimarca

INTRODUZIONE

La possibilità di utilizzare un computer per analizzare grandi quantità di testo è emersa fin dagli anni Settanta ed è stata una delle prime applicazioni dell'informatica in ambito umanistico. Grandi *corpora* di testi (v. riquadro 1) sono da tempo strumenti d'uso quotidiano nelle mani dei linguisti per ricerca 'pura' o applicata alla redazioni di dizionari, mentre un gruppo di ricercatori e insegnanti, prevalentemente di area anglosassone, ne ha proposto con energia fin dalla fine degli anni '80 l'utilizzo a scuola in un contesto di esplorazione del linguaggio [Johns, 1988].

Nel frattempo, grandi archivi di testi sono realmente diventati disponibili su computer a costi accessibili, in forma di grandi archivi di raccolte di giornali su CD-ROM, enciclopedie elettroniche, o siti web contenenti raccolte di testi. Queste opere elettroniche possono essere un'importante risorsa per diverse aree disciplinari, con rilievo particolare per l'educazione linguistica.

Questo articolo presenta appunto una serie di proposte didattiche che ruotano attorno all'uso di un CD-ROM di testi giornalistici in lingua italiana. Le proposte partiranno da qualche esempio di uso di articoli della raccolta come base per attività di lettura e scrittura, per passare successivamente a considerare l'intera raccolta come un ambiente per la ricerca di informazioni. Verrà infine esaminato l'uso dell'intero *corpus* di testi come base di dati linguistici da esplorare e da interrogare in un contesto di educazione linguistica in prima e seconda lingua.

Si seguirà quindi un percorso che da usi

didattici abbastanza assestati arriverà a proposte quasi 'di frontiera', seguendo un doppio filo conduttore: da un lato, uno strumento che ha il potenziale per aiutare l'insegnante a costruire un ricco ambiente di apprendimento; dall'altro, l'obiettivo di esercitare gli studenti in attività che promuovano strategie di apprendimento.

INTERNAZIONALE E INTERNAZIONALE SU CD-ROM

Internazionale è un settimanale che dal 1993 presenta in traduzione italiana articoli dalla più quotata stampa straniera su temi di società, politica, economia, divulgazione scientifica, cultura e spettacoli. Ogni numero contiene, al netto della scarsa pubblicità, una cinquantina di pagine formato A4, tratte in buona parte da quotidiani e periodici "occidentali" (da *El Pais* e *Le Monde* al *Wall Street Journal* a *New Scientist*, tra i più citati) con frequenti e interessanti incursioni nella stampa post-coloniale dei continenti africano, asiatico e sudamericano (es. *Jeune Afrique*). Le traduzioni sono molto buone, la grafica è semplice ma molto curata, con splendide fotografie, nonostante la stampa su carta economica. Insomma, un piccolo miracolo editoriale.

Internazionale su CD-ROM offre poi a un prezzo molto ragionevole (quindici-mila lire) tutti i numeri usciti dall'ottobre 1993 a fine 1998, vale a dire oltre 4000 articoli (senza contare le notizie brevi) e dieci milioni di parole.

Ma a che cosa serve o interessa tutto ciò per la didattica in generale e della lingua in particolare?

Internazionale è una miniera di materiale che si può agevolmente inserire in diversi contesti didattici.

Un primo gruppo di proposte vede *Internazionale* usato di volta in volta come materiale di base per la didattica della lettura per lo studio, come fonte per attività di scrittura documentata, come repertorio di modelli di testo, come libro di lettura per il corso di geografia del primo anno e persino come base per preparare il tema di maturità, per quanto riguarda attualità e argomenti di taglio scientifico-tecnologico. A questi usi, che si richiamano a una didattica relativamente consolidata della lettura e della scrittura, *Internazionale* ha da offrire soprattutto la ricchezza e qualità dei materiali.

Ma ci sono anche altre possibilità, legate in particolare alla disponibilità della versione elettronica: innanzi tutto, la consultazione di un archivio di medie dimensioni allena a una certa capacità di muoversi tra informazioni complesse; inoltre, l'interrogazione del *corpus* di testi alla ricerca di informazione su lessico e sintassi offre piste per esplorare il linguaggio in un ambito di prima o di seconda lingua.

A questo secondo gruppo di attività è dedicato in questo articolo uno spazio maggiore non perché più importanti delle altre, ma perché di norma meno radicate nella tradizione didattica.

Gli esempi d'uso proposti sono basati su esperienze di insegnamento di prima lingua in un Istituto Tecnico e di Italiano come seconda lingua con studenti universitari danesi.

INTERNAZIONALE E LA DIDATTICA DELLA

“LETTURA PER LO STUDIO”

Internazionale fornisce un gran numero di articoli interessanti e ben tradotti, scritti per lettori colti non specialisti, che si prestano assai bene come materiale per una didattica della “lettura per lo studio” con studenti italiani di biennio superiore. Particolarmente utili in questo contesto articoli di divulgazione scientifica oppure analisi di questioni sociali, spesso da paesi lontani (es. “fidanzamento e matrimonio in Costa d'Avorio tra tradizione e innovazione”, oppure “musica e anti-integralismo in Algeria”). Incidentalmente, alcuni degli articoli sono diventati parte inte-

RIQUADRO 1: Corpora di testi

Il sito di Michael Barlow (<http://www.ruf.rice.edu/~barlow/corpus.html>) offre una panoramica di *corpora* di testi disponibili e una ricca bibliografia sull'argomento (alcuni dei testi sono disponibili *on-line*). Una nota per chi si avventurerà in questo sito: alcuni *corpora* sono segnalati come *tagged* e ciò significa che sono corredati di marche semantiche introdotte dai redattori per individuare, ad esempio, le parti del discorso (altri testi hanno una marcatura più *leggera* che serve, ad esempio nelle trascrizioni di processi, a identificare le diverse persone che parlano).

Può essere utile segnalare alcuni *corpora* di testi particolarmente importanti e accessibili al pubblico (a pagamento, per ora; ma è possibile fare ricerche di prova).

Il *British National Corpus* (<http://info.ox.ac.uk/bnc/>) è consultabile *on-line* e offre un'utile guida al suo uso didattico: Guy Aston, *The British National Corpus As A Language Learner Resource*, 1996 (<http://info.ox.ac.uk/bnc/using/papers/astor96a.html>). Chris Tribble (un altro dei pionieri dell'uso didattico di questi strumenti) offre utili indicazioni pratiche per chi è interessato all'acquisto della versione ‘ridotta’ per PC (<http://ourworld.compuserve.com/homepages/Christopher Tribble/BNC PC.html>).

Anche un sottoinsieme della *Bank of English* gestita dall'editore Collins è accessibile sul web (<http://titania.cobuild.collins.co.uk/directinfo.html>).

Chi vuole provare un programma di concordanze può scaricare la versione *demo* di *MonoConc Pro* (<http://www.athcl.com/READ-ME.html>) e provarla su qualche testo preso, ad esempio, dal sito del Progetto Manuzio (<http://www.liberliber.it>).

Per una recensione della prima versione di *MonoConc* e un'analisi delle caratteristiche dei *concordancers* v. Vance Stevens, *MonoConc 1.2 (review)* in *CALICO Software Product Review*, 1997 (<http://www.arts.monash.edu.au/others/calico/review.monconc.html>).

Chi volesse costruire un *corpus* di testi piccolo ma relativamente omogeneo su un argomento può saccheggiare un sito web oppure, ancora più facilmente, un'enciclopedia su CD-ROM, seguendo le indicazioni di Chris Tribble, *Improvising corpora for ELT: quick-and-dirty ways of developing corpora for language teaching*, 1997 (<http://ourworld.compuserve.com/homepages/Christopher Tribble/Palc.html>).

Per quanto riguarda infine l'idea di far esplorare il linguaggio agli studenti un ottimo riferimento è *Tim Johns Data-driven Learning Page* (<http://web.bham.ac.uk/johnstf/timconc.html>).

grante nel corso di geografia nel primo anno.

L'obiettivo didattico più importante delle attività svolte era che gli studenti capissero la necessità di formulare domande – prima, durante e dopo la lettura – sugli scopi e sui contenuti del testo.

In una prima fase fornivo io stesso le domande, presentandole come un aiuto, dopo che qualche “interrogazione” pro-forma aveva mostrato la difficoltà di molti a individuare nel testo quello che è davvero importante. In seguito ho continuato a fornire io domande e chiose nel caso di testi complessi, ma ho insistito perché

fossero il più possibile gli studenti a preparare le “domande di aiuto” per sé o per altri compagni.

Sempre nella prospettiva di sviluppare questo tipo di *competenza interrogativa*, la mia spiegazione di qualche passo o concetto difficile avveniva possibilmente in un contesto in cui erano gli studenti a “interrogare” formalmente l’insegnante, cosa in sé abbastanza ragionevole (v. in proposito [Weinrich, 1983]).

Un altro obiettivo delle attività di lettura guidata era l’affinamento di tecniche per gestire i vocaboli sconosciuti o compresi in modo impreciso. A questo fine l’approccio proposto in una prima fase era di far negoziare il significato agli studenti a partire dal contesto, e di utilizzare eventualmente il dizionario come conferma o come *ultima ratio*. La possibilità, attraverso la versione su CD-ROM, di vedere la parola in *una serie* di contesti ha segnato un salto di qualità in questo approccio, che sarà comunque descritto meglio nel seguito.

INTERNAZIONALE E LA SCRITTURA DOCUMENTATA

La costante attenzione di *Internazionale* per grandi temi come l’impatto delle nuove tecnologie, la coesistenza e i contrasti tra le culture, oppure le ricerche sulla mente, si traduce in una massa critica di materiale utilizzabile per attività di scrittura documentata, dall’espansione di agenzie di stampa (di cui il sito web di *Internazionale* fornisce ogni settimana agli abbonati l’anteprima) alla preparazione al *tema d’attualità* della maturità, fino alla stesura di relazioni più impegnative, eventualmente in contesto interdisciplinare (ad es. sono serviti in questo senso una serie di articoli sui confini tra biologia e cultura).

A conferma dell’esistenza di un filo conduttore nella scelta dei materiali da parte della rivista, periodicamente *Internazionale* pubblica volumi monografici costruiti come raccolta dei migliori articoli già pubblicati su un dato argomento: sono finora apparsi tra l’altro volumi sull’ambiente, sui nuovi percorsi della scienza (in particolare a proposito delle discipline che tentano di definire la natura umana, dalle scienze cognitive alla psicologia evolutiva), sulla globalizzazione, ol-

tre a una serie di ritratti socio-politici dei continenti.

E’ innegabile che una collezione di articoli non è confrontabile in organicità con un testo pensato apposta per la didattica. Al tempo stesso, definire con una classe una piccola lista di argomenti e tornarvi periodicamente nell’arco dei cinque anni è un modo per cercare di ancorare la “cultura della classe” a temi rispetto ai quali gli studenti abbiano la percezione di crescere nella comprensione del mondo esterno. Contemporaneamente, il fatto che su questi argomenti si continui (nella classe e nel mondo) a ragionare e a scrivere, contribuisce a farne percepire la complessità e a promuovere un’idea della conoscenza come *processo*: un buon antidoto a un’immagine – ancora troppo diffusa in particolare tra gli studenti – della mente come *contenitore* da riempire con un insieme predefinito di contenuti [Bereiter & Scardamalia, 1996; Bereiter, 2000].

Internazionale come repertorio di testi da usare come modelli

Oltre che come portatori di informazioni contenutistiche su un argomento, gli articoli di *Internazionale* possono essere visti come modelli di testi appartenenti a un determinato genere: reportage, interviste, articoli di fondo di varia lunghezza, divulgazione scientifica, “ riquadri illustrativi” su un tema, agenzie di stampa, lettere ai giornali e vignette (queste ultime, solo nella versione su carta).

Imitare buoni modelli è un modo efficace per migliorare la propria scrittura, in prima e in seconda lingua, in particolare se l’imitazione comprende una messa a fuoco esplicita dei vincoli formali e contenutistici del testo-modello e del testo da produrre (v. basati sulle esperienze con *WordProf* e *English WordProf* [Ferraris et al., 1990; Caviglia et al., 1996]; v. inoltre [Geist, 1996]).

La disponibilità di un gran numero di articoli su vari argomenti rende inoltre possibile un’attività di attivazione del lessico relativo al tema di cui si deve scrivere. Piccoli gruppi di studenti possono produrre in circa un’ora un piccolo lessico specifico a un dato argomento (es. l’immigrazione, con i termini correlati di (*in*)tolleranza, *diversità culturale*, *accoglienza*, ecc.), e poi renderlo disponibile a

ed errore qualcuno ha proposto “lavoro minorile”, riducendo così il campo a 35 occorrenze distribuite su 5 articoli di un certo peso e numerose agenzie di stampa (se i numeri restano troppo grandi, è sempre possibile restringere la ricerca a una sola annata della rivista).

Come illustrato in Figura 2, a partire da ogni occorrenza è possibile visualizzare l'intero articolo, che si potrà poi copiare in un word processor e stampare (finita la fase di ricerca di informazioni, è preferibile far svolgere su carta tutte le attività di lettura accurata).

Naturalmente, se l'obiettivo didattico primario è quello di insegnare strategie di ricerca, le informazioni da trovare possono e spesso devono essere puntuali e precise, in modo da poter isolare l'abilità e farla esercitare su un buon numero di casi, eventualmente creando forme di competizione e collaborazione tra i gruppi.

Certo non è indispensabile un computer per fare didattica della ricerca di informazioni. Un orario dei treni, le *Pagine Gialle* e un qualsiasi dizionario offrono spunti altrettanto validi, che vale certamente la pena di esplorare parallelamente. Però, nel caso che abbiamo illustrato, il computer da un lato costringe a elaborare strategie abbastanza raffinate, ma contempora-

neamente fa vedere *tangibilmente* (se la parola si può usare in questo contesto) la potenza delle strategie stesse associate a uno strumento di ricerca abbastanza stupido ma velocissimo, che permette operazioni che dieci anni fa avrebbero richiesto a un esperto un tempo molto più lungo.

Internazionale come base di dati linguistici

Corpora di testi e esplorazione del linguaggio

Fino a questo punto sono stati proposti usi e abilità in certa misura “tradizionali”, che trovano in questo contesto uno spazio e una necessità più forti. L'uso didattico più innovativo di un *corpus* di testi è forse invece quello che lo vede come base di dati linguistici da far esplorare agli studenti. Si tratta - come accennato all'inizio di questo articolo - di un approccio proposto da oltre un decennio da un gruppo di ricercatori e insegnanti convinti che l'attività di ricerca sul linguaggio, resa possibile dalla tecnologia dei *corpora* di testi, sia di tale valore cognitivo che sarebbe un peccato non utilizzarla in un contesto scolastico (per una panoramica e una discussione su questo approccio v. [Murison-Bowie, 1996; Aston, 1995 e 1997]; v. per esempi la *Tim Johns Data-driven Learning Page* nel riquadro 1. In

questa prospettiva sono stati sviluppati buoni software per concordanze [MonoConc, 1997 e 2000] e altrettanto buone guide per il loro uso in classe (come [Tribble & Jones, 1990 e 1997]).

L'idea didattica alla base della loro proposta è quella di un approccio esplorativo, attraverso il quale gli studenti cercano di individuare regolarità nella lingua e possono provare a generalizzarle come regole, per poi verificarle nuovamente. Curiosamente,

Figura 2
Risultato della ricerca “lavoro minorile”.



si tratta di un'idea simmetrica e complementare a quella portata avanti alla fine degli anni '80 da Lydia Tornatore, che aveva realizzato alcuni programmi per consentire agli studenti di definire regole e di verificarle osservando le frasi generate dal programma stesso in base alla grammatica costruita dagli studenti [Tornatore, 1989; Caviglia & Sarti, 1989]. L'idea era affascinante ed è stata sperimentata con un certo successo a livello di scuola elementare e media inferiore, ma è risultata tecnicamente difficile da realizzare al di là di un livello elementare di complessità. La disponibilità di ampie basi di dati linguistici e di programmi di concordanze per analizzarle consente ora di affrontare tale complessità con qualche strumento in più, sia pure ancora di uso non banale. Infatti confrontarsi con grandi masse di dati può essere non facile per studenti e insegnanti [Aston, 1995; Stevens, 1995]: questa è certamente una ragione che spiega come sia le proposte della Tornatore che quelle, pure più fortunate, di area anglosassone siano rimaste limitate finora a una comunità abbastanza ristretta. Un *corpus* di testi si presta ad essere usato, in contesto didattico, in particolare per:

- analizzare e 'negoziare' il significato di una parola a partire dai contesti in cui essa appare;
- individuare e cercare di spiegare il perché di costrutti sintattici, sempre a partire dai contesti in cui ricorrono.

Inoltre, un *corpus* di testi di media dimensione può essere anche usato, da parte di studenti *strategicamente* attrezzati, come opera di consultazione complementare a dizionari e manuali di grammatica; in particolare, oltre a indagare su significati e costrutti sintattici, un *corpus* di testi si presta bene a scoprire le asso-

ciazioni tipiche tra parole (*collocations*, nel lessico dell'editoria).

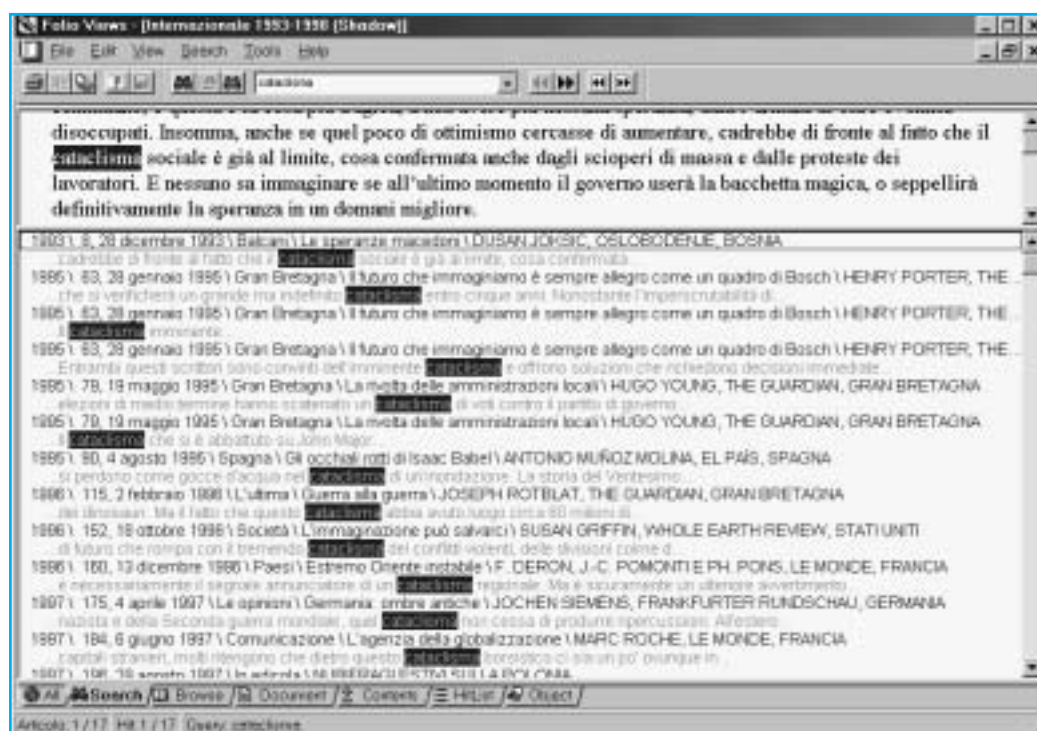
Nel caso particolare di *Internazionale*, cinque annate significano un *corpus* di circa dieci milioni di parole. Si tratta di un campione significativo di italiano scritto non letterario e non specialistico: in pratica, il livello che uno studente italiano di scuola superiore oppure uno studente universitario di italiano come seconda lingua dovrebbero padroneggiare. Il fatto che questo *corpus* sia costituito da testi tradotti non mi sembra un ostacolo a considerarlo rappresentativo. Certamente è interessante e utile anche un *corpus* basato su articoli di giornali italiani (ne esistono, per inciso, anche se non così immediatamente utilizzabili ed economici come *Internazionale*). Ma, paradossalmente, una buona traduzione (come sono quelle su *Internazionale*) è un modello particolarmente affidabile di lingua, proprio perché vincolata a uno scarto minimo rispetto agli standard linguistici. Vedremo ora qualche esempio d'uso nell'area del vocabolario e della sintassi.

Inferire significati dal contesto

La possibilità di individuare tutte le occorrenze di una parola nell'intero *corpus* della rivista e di visionarle nel loro conte-

Figura 3

Occorrenza di "cataclisma", con contesto.



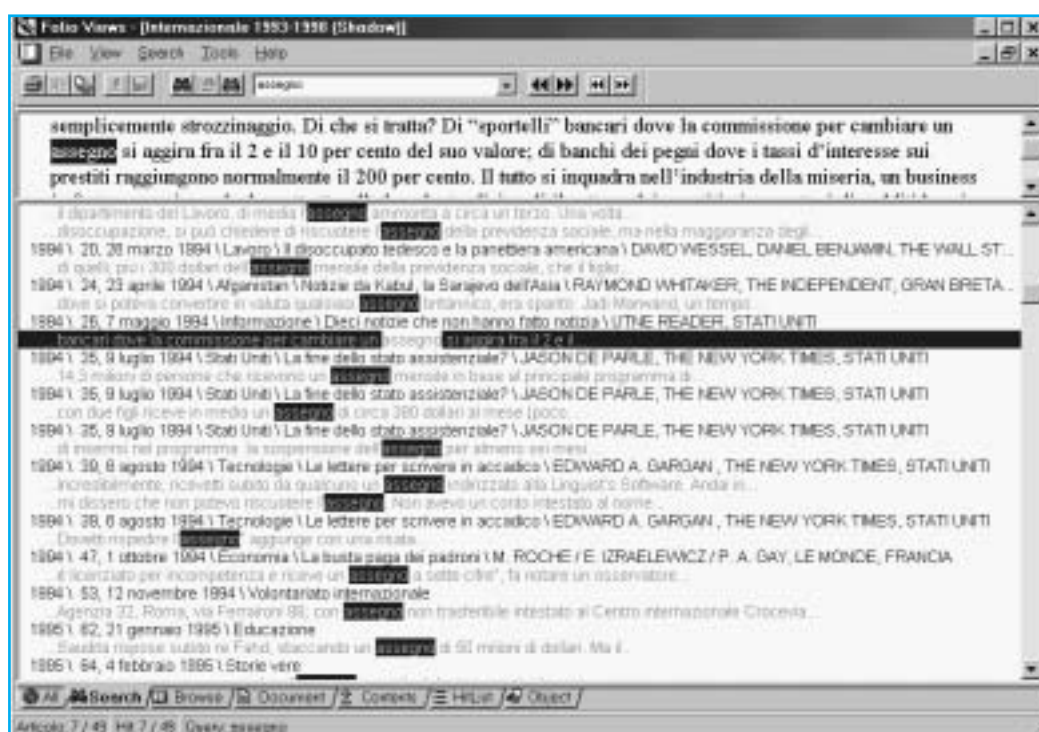


Figura 4
La parola “assegno”
in vari contesti.

sto, consente di usare il CD-ROM come una sorta di dizionario in linea. In questo senso ne è stato proposto l’uso a studenti avanzati di italiano come seconda lingua,

corrono (v. ad es. la Figura 3 con le occorrenze di *cataclisma*): si tratta di una strategia più laboriosa rispetto alla ricerca del significato nel dizionario, ma forse più

alcuni dei quali dotati di una competenza linguistica in parte comparabile - per quanto riguarda lettura e scrittura - a quella di studenti italiani di biennio superiore.

Questa funzione si presta assai bene a promuovere l’abitudine di ricavare il significato di una parola dal contesto. Il significato di parole sconosciute (o conosciute in maniera insicura, come avviene nella maggior parte dei casi) si può inferire attraverso l’analisi di una serie di contesti in cui esse ri-

RIQUADRO 2: Dizionari e limiti dell’ordine alfabetico

Lo spagnolo *Diccionario de uso del Español* di María Moliner è un esempio di dizionario che già nell’edizione cartacea raggruppa attorno al lemma definizione, esempi d’uso, indicazioni sintattiche e altre parole connesse. Per il tedesco, da molti anni la Duden offre su carta un praticissimo *Stilwörterbuch* con un set limitato di lemmi ai quali associa le espressioni di uso più tipico (ad esempio, data la parola “assegno” compariranno “emettere, incassare, girare” come pure “assegno in bianco, scoperto, postdatato, ecc.”).

Per l’inglese due strumenti particolarmente interessanti sono *WordNet* (solo elettronico, distribuito gratuitamente dalla Princeton University) e il *Longman Activator*. Per realizzare *WordNet* nomi, verbi e aggettivi della lingua inglese sono stati organizzati in insiemi di sinonimi, dove ogni insieme rappresenta un concetto lessicale: gli insiemi di sinonimi sono stati poi collegati tra loro da relazioni che definiscono rapporti di antonimia, iponimia, iperonimia, meronimia: a partire da un sostantivo *book* consentono di accedere agli equivalenti inglesi di *scritto, opera, pubblicazione* (iperonimi), oppure *catechismo, ricettario, manuale, libro scolastico, pamphlet* (iponimi), o ancora *copertina, rilegatura*, ma anche *prefazione, paragrafo, capitolo* (meronimi propri o ereditati in quanto meronimi degli iperonimi). Il *Longman Activator* (ancora cartaceo), raggruppa invece attorno a 1052 significati fondamentali (*core meanings*) 23.000 lemmi costituiti da

una parola o una perifrasi (*word- and phrase-meanings*), organizzandoli in modo da farne emergere le specificità funzionali e di registro e spesso ridefinendoli in confronto alle altre possibili scelte per quel dato *core meaning*: uno strumento prezioso basato tra l’altro su un grande *corpus* di testi contemporanei, dai quali sono ricavati i numerosi esempi d’uso. Il dizionario *Collins COBUILD*, qualche anno fa, ha direttamente offerto un *corpus* di testi accessibile con una sola ricerca dall’interno del dizionario su CD-ROM (tuttavia nell’ultima versione, il *Collins COBUILD E-dict*, l’editore ha preferito integrare il dizionario con una grammatica, e raffinare le ricerche all’interno del dizionario (è possibile richiedere un demo a <http://titania.cobuild.collins.co.uk/>).

In italiano, accanto ai vari dizionari dei sinonimi è disponibile il *Dir* “Dizionario italiano ragionato”, simile per alcuni aspetti al *Duden Stilwörterbuch*; ne esiste una versione elettronica, che ricalca sostanzialmente quella cartacea.

Quanto poi alle versioni elettroniche di dizionari tradizionali, sistemi più raffinati di ricerca (ad es. nella versione in CD-ROM del DISC della Giunti consentono esplorazioni del lessico impensabili con un dizionario cartaceo. Tuttavia non mi risultano al momento dizionari italiani pensati fin dall’inizio per il supporto elettronico.

Collins COBUILD E-dict, London: HarperCollins, 1998 (<http://titania.cobuild.collins.co.uk/>)

utile a promuovere pratiche *strategiche* di lettura e comprensione ([Haarstrup, 1991]; v. anche *Jeu du MAI*, un software didattico costruito proprio sull'idea di individuare la stessa parola 'schermata' che ricorre in contesti diversi).

La stessa strategia di analisi della parola in contesto torna poi utile in fase di scrittura, quando gli studenti possono verificare le parole che usano con minor sicurezza, per controllare che risultino davvero attestate in un dato senso e contesto.

Un'obiezione a quest'uso che è stata posta da studenti e da insegnanti è che un *corpus* di testi non offre le garanzie di completezza di un dizionario. Hanno ragione: del resto, non si tratta di strumenti alternativi: idealmente, vorrei che gli esempi proposti dal dizionario fossero all'inizio della lista di occorrenze che mi ritorna un *corpus* di testi. Tuttavia, *finché il testo che lo studente sta leggendo o scrivendo è omogeneo al corpus*, il lessico contenuto nel *corpus* rappresenta un repertorio affidabile di soluzioni possibili. Per contro non è raro – soprattutto per chi scrive in seconda lingua – usare *dopo averle controllate nel dizionario* parole che nessun parlante nativo metterebbe in quel dato contesto.

Esulerebbe tuttavia dallo spazio di questo articolo un confronto, che pure sarebbe utile, tra quello che *corpora* di testi e dizionari offrono a fronte delle esigenze di diversi tipi di utente. Mi limito qui a notare come da tempo i dizionari, cartacei non meno che elettronici, hanno provato con diverse strategie ad aggirare i vincoli dell'ordine alfabetico per rendere accessibili al lettore le parole a vario titolo associate tra loro (v. riquadro).

Proprio a questo scopo un *corpus* di testi può essere particolarmente utile.

Trovare associazioni di parole

Spesso uno studente vorrebbe sapere quale parola un parlante nativo (o chi scrive per professione) utilizzerebbe in un dato contesto. Certamente un uso mirato e creativo dei dizionari italiani disponibili fornirebbe già un certo aiuto, ma setacciare i contesti di una parola-chiave potrebbe essere un modo ancora più efficace per individuare parole che forse non conosciamo ancora, ma che sappiamo esistere in associazione con altre.



Se, per esempio, cerchiamo *assegno*, troveremo 35 occorrenze i cui contesti, parzialmente riportati nella Figura 4, contengono buona parte del vocabolario associato (*cambiare, riscuotere, emettere, staccare, non trasferibile*; oppure i più generali *mandare e ricevere*; o, ancora, espressioni metaforiche come *assegno in bianco*).

Analogamente, può essere utile trovare quelle locuzioni tipiche per cui si *commette* un crimine o si *avanza* una proposta. Anche se non sempre felici, queste espressioni rappresentano uno standard che si può benissimo rifiutare, ma di cui ha senso essere consapevoli.

Nell'esempio della Figura 5, una ricerca leggermente più complessa restituisce una lista di verbi usati in associazione con *problema*.

Una sequenza di ricerche di questo genere può servire anche per costruire uno di quei glossari da attivare in fase di prescrizione di cui si è detto sopra, nella sezione sull'uso di modelli di testo.

ESPLORARE LA SINTASSI

Un uso emerso prevalentemente in un contesto di seconda lingua è quello di utilizzare il *corpus* di testi per risolvere dubbi sintattici. Le preposizioni sono una delle aree che generano più dubbi: si dice *diverso da* o **diverso di*? *sembrare* vuole il

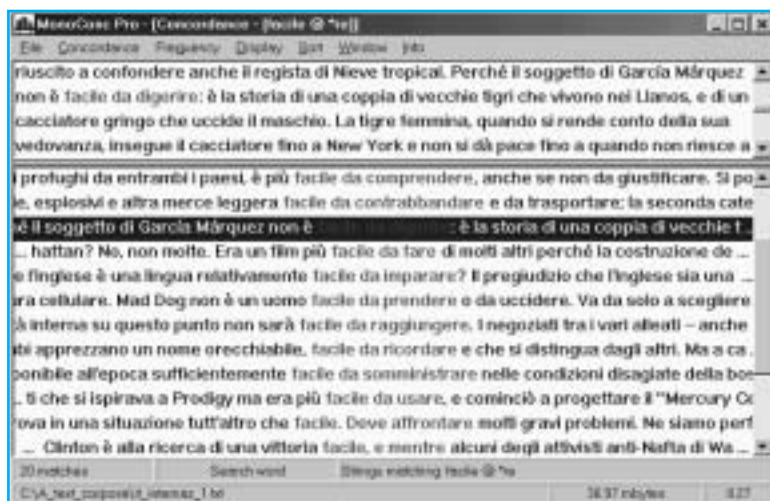
Figura 5

Alcuni verbi associati a *problema*: vedi ad esempio *risolvere, aggirare, costituire, sollevare, individuare*. La ricerca (*query*) “*re problema”/3 significa “trova tutte le occorrenze di parole che finiscono in **re* (cioè degli infiniti, probabilmente) seguite da *problema* a una distanza massima di tre parole”.



Figura 5a
Quale preposizione unisce “facile” a un infinito?

Figura 5b
Quale preposizione unisce “facile” a un infinito? Le occorrenze sono riportate in ordine alfabetico. In questo caso lo stesso corpus di testi è stato interrogato con *MonoConc Pro*, un programma per la gestione delle concordanze che consente di riorganizzare i risultati delle ricerche.



analogia, in questo caso *facile* + *una parola* + infinito (v. Figure 5a e 5b). In questo caso trovare una soluzione convincente e pressoché sicura ha permesso di percepire, a posteriori, le differenze piuttosto sottili tra *cibo buono da mangiare* e *occasione buona per risolvere un problema*.

Molte delle informazioni sintattiche che si trovano interrogando un corpus di testi sono certo presenti in un buon dizionario, ma - come nel caso del lessico - anche in questo caso partire dall'esempio per arrivare alla regola è più stimolante e verosimilmente più valido per l'acquisizione di strategie di soluzione di problemi linguistici.

Più in generale, tutte le attività finora proposte di esplorazione del linguaggio si richiamano a un'idea di interattività e di costruzione di regole e significati a partire da esempi non predefiniti, ma ricavati dallo studente: un approccio che dovrebbe favorire lo sviluppo di abilità autonome di apprendimento nel senso indicato in Oostdam & Rijlaardsam [1995].

Ma non bisogna dimenticare che fare ricerche su una grande base di dati linguistici non è facile. A detta di alcuni tra i promotori di un uso didattico dei corpora di testi [Stevens, 1995], una delle difficoltà maggiori è il lavoro di rapida selezione che lo studente deve esercitare sui risultati delle ricerche, che spesso nascondono pochi esempi decisivi in mezzo al ‘rumore’ di risultati irrilevanti. A questo proposito avrebbe senso graduare il carico cognitivo sugli studenti facendo svolgere inizialmente in prima lingua le attività esplorative del linguaggio, e aggiungere l'ulteriore difficoltà della seconda lingua solo in un secondo tempo. Inoltre, se la prospettiva didattica è quella di familiarizzare gli studenti con la ricerca su dati linguistici (e non, come negli esempi presentati finora, di avere materiale per risolvere un'ampia gamma di problemi), conviene lavorare su un corpus di dimensioni ridotte [Aston, 1995]. Con lo stesso *Internazionale su CD-ROM* è possibile, ad esempio, lavorare su una sola annata della rivista (‘appena’ due milioni di parole); ma è anche facile costruire un corpus di testi di piccole o medie dimensioni, in qualsiasi lingua, saccheggiando il sito Internet di qualche rivista, utilizzando le varie edizioni ‘nazionali’ di un'enciclopedia elettronica come *En-*

carta o, ancora, utilizzando testi on-line dei vari progetti *Gutenberg* o *Manuzio*. In questa prospettiva e per un uso non occasionale si può considerare anche l'acquisto di un programma per concordanze di facile uso, come ad esempio *MonoConc* (v. Figura 5b).

CAUTELE E CONCLUSIONI

È necessario ricordare la necessità di una certa cautela, quando si avanzano proposte didattiche legate all'uso di tecnologia, soprattutto nell'ambito di materie umanistiche: usi concettualmente complessi come quelli proposti sono produttivi solo in situazioni in cui esista già (nella classe, nella scuola, nelle case degli studenti) una cultura di uso di risorse informatiche al servizio di diverse attività. Per fare un esempio, probabilmente non ha senso proporre attività complesse di esplorazione del linguaggio a un gruppo di persone

che non ha mai usato un word processor: i malcapitati si troverebbero a lottare contemporaneamente con difficoltà di genere diverso e probabilmente rifiuterebbero dopo qualche tempo gli strumenti e, quel che è peggio, le idee didattiche che stanno dietro al loro uso.

Per richiamare brevemente in conclusione le idee didattiche in questione, il denominatore comune di tutte le attività proposte è la logica di una didattica più centrata sull'imparare a imparare che sull'insegnamento di un contenuto specifico.

Al servizio di questa logica e dell'ampia area disciplinare toccata in questa presentazione *Internazionale* - nelle due versioni, cartacea e su CD-ROM - offre una tale quantità di materiali e di possibilità d'uso da richiamare alla mente l'aspetto più nobile del maiale, del quale, diceva la saggezza contadina, *non si butta via niente*.

Riferimenti Bibliografici

Bereiter C. & Scardamalia M. (1996), *Rethinking Learning*, in Olson D. & Torrance N. (eds.), *The Handbook of Education and Human Development*, MA: Blackwell, Malden.

Bereiter C. (2000) (in progress), *Education and Mind in the Knowledge Age*, disponibile solo on-line (<http://csile.oise.utoronto.ca/ed-mind/edmind.html>)

Caviglia F., Earp J., Ferraris M. (1996), *From Italian (L1) to English (L2) WordProf: an instructional software for teaching writing*, in *1996 European Writing Conferences*. - EARLI Special Interest Group

Writing, Barcelona, 23-25 October 1996; ISBN 84-88795-36X.

Caviglia F., Sarti L. (1989), *Intelligenza artificiale: quale contributo all'educazione linguistica?*, in Calvani A. (ed.), *Scuola, computer, linguaggio*, Loescher, Turin, pp. 101-159.

Ferraris M., Caviglia F., Degl'Innocenti R. (1990), *WordProf: a writing environment on computer*, in Hartley W. (ed.), *Technology and Text: Readings in the Psychology of Written Communication*, Jessica Kingsley Ltd., London/Philadelphia, 1992 [pubblicato originariamente in *Educational Technology and Training International* (ex "PLET"), February 1990, pp. 33-42].

Geist U. (1996), *Imitation as a tool in writing pedagogy*, in Rijlaarsdam G., Van den Bergh H. & Couzijn M. (eds.), *Effective Teaching and Learning of Writing*, Amsterdam University Press, Amsterdam, pp. 51-60.

Haarstrup K. (1991), *Lexical Inferencing Procedures or Talking about Words*, Gunter Narr, Tübingen.

Johns T. (1988), *Whence and Whither Classroom Concordancing?*, in T. Bongaerts et al. (eds.)

Computer Applications in Language Learning, Foris, Dordrecht.

Murison-Bowie S. (1996), *Linguistic corpora and language teaching*, in *Annual Review of Applied Linguistics*, n. 16, pp. 182-199.

Oostdam R., Rijlaarsdam G. (1995), *Toward Strategic Language Learning*, Amsterdam University Press, Amsterdam.

Stevens V. (1995), *Concordancing with Language Learners: Why? When? What?*, in *CAELL Journal*, vol. 6, n. 2, pp. 2-10 (<http://www.ryf.ri-ce.edu/-barlow/stevens.html>).

Tornatore L., (1989), XXX, in Calvani A. (ed.), *Scuola, computer, linguaggio*, Loescher, Turin, pp. 101-159.

Tribble Ch. & Jones G. (1990), *Concordancing in the classroom*, Longman, Harlow, Essex (new edition: *Concordances in the classroom: a resource book for teachers*, Athelstan 1997).

Weinrich H. (1983), *Vers la constitution d'une competence interrogative*, in *Zielsprache Französisch*, n. 2, 1983, pp. 57-64 (tr. it. in Weinrich H., *Vie della cultura linguistica*, Il Mulino, Bologna, 1989)

Riferimenti Software

Internazionale su CD-ROM (1993-1998), Roma: Internazionale, 2000 (1. ed. 1997) (<http://www.internazionale.it>)

Jeu du mai (= **mots à indoviner**), Paris : CNDP/CREDIF, 1995 (<http://www.cndp.fr>)

MonoConc (realizzato da Michael Barlow), Houston, TX : Athelstan, 1997-2000 (versione corrente è *MonoConc Pro* v. 2.0, 2000; 1. ed. 1997) (<http://www.athel.com>)

WordProf Ferraris M., Caviglia F., Degl'Innocenti R. 1993, *Scrivere con*

WordProf, Milano: Theorema, 1993. (software + libro).

English WordProf Caviglia F., Earp J., Ferraris M., 1996, *English WordProf*, Genova: ITD/CNR, disponibile presso ITD/CNR.