
Corpora, Google e roba simile

Per quale ragione gli studenti di una lingua straniera dovrebbero perderci tempo?

By Cesare Zanca (Università di Siena, Italy)

Abstract & Keywords

English:

This paper suggests that corpus linguistics methodology can be introduced even to students with a limited foreign language competence. It describes how learners can be guided to an inductive discovery of language patterns and usage using selected and familiar online resources – such as Google or Wikipedia - and how these resources can be explicitly related to corpus linguistics and complemented by “proper” corpora to solve linguistic problems emerging in writing or translation activities, while limiting the objections often generated by the use of corpus linguistics in the classroom. This approach, I claim, gradually develops learners’ autonomy, awareness and methodological soundness. The paper describes activities with undergraduate students using mainly online resources and corpora and with postgraduate learners who also collect their own specialized corpus. The students’ feedback seems to confirm their ability to deal with aspects of the foreign language they could hardly handle with more traditional reference tools.

Italian:

Questo studio propone la tesi che la metodologia tipica della linguistica dei corpora possa essere impiegata per favorire l’apprendimento delle lingue straniere anche con apprendenti non avanzati. Descrive percorsi che guidano ad una scoperta induttiva di elementi di uso e patterns linguistici ricorrenti a partire dall’impiego di alcune risorse online più famigliari – come Google o Wikipedia – e sostiene che queste risorse possono, da un lato, essere facilmente seguite ed integrate da corpora ‘veri e propri’ e, dall’altro, ridurre le frequenti obiezioni di studenti ed insegnanti rispetto all’impiego in classe di strumenti e modalità di indagine considerati spesso troppo tecnici e specialistici. Questo approccio, usato tendenzialmente per attività asincrone come quelle basate sulla traduzione e la scrittura, può sviluppare una maggiore autonomia, consapevolezza e sicurezza metodologica nell’apprendente. Lo studio riporta esempi tratti da corsi con studenti universitari sia di un corso di laurea triennale, per il quale vengono usati soprattutto risorse e corpora reperibili online, sia di un corso di laurea magistrale, in cui è prevista anche la compilazione di un corpus bilingue. Alcuni esempi di feedback degli studenti sembrano confermare come questo approccio consenta loro di gestire aspetti problematici dell’uso della lingua straniera che sarebbero difficilmente affrontabili con strumenti più tradizionali quali dizionari e grammatiche.

Keywords: corpus linguistics, data driven learning, web as corpus, lexical priming, cognitive linguistics, didattica delle lingue

1. Introduzione

Al centro di questo mio contributo è la descrizione di un approccio didattico mirante all’impiego da parte di apprendenti di una lingua straniera della metodologia di esplorazione propria della linguistica dei corpora. Negli ultimi decenni, questo impiego è già stato al centro di numerosi studi e viene considerato da molti come un approccio innovativo e di grandi potenzialità ma, di fatto, resta largamente sconosciuto alla grande maggioranza degli insegnanti e degli studenti, sia in ambito universitario che scolastico (si vedano ad esempio Boulton, 2010; Maunier, 2011). Questa incongruenza è stata evidenziata, tra gli altri, da Römer (2008:123): «The gap between corpus linguistics and the teaching reality described by Mukherjee (2004), is still far too wide, and the extent to which corpora and concordances have actually been used in LT classrooms is, unfortunately, as yet fairly limited». La mia personale esperienza, come utilizzatore di corpora sia per la ricerca che per l’insegnamento in ambito accademico, mi ha convinto che uno dei maggiori ostacoli a questa mancata diffusione possa essere individuata nelle modalità in cui i corpora vengono proposti agli apprendenti e che un diverso approccio e inserimento nel percorso didattico possa risultare più efficace, favorendo una maggiore autonomia e motivazione nell’apprendimento.

Esso consiste nell’inserire l’uso dei corpora in un percorso didattico *task based*, legato alla soluzione di problemi linguistici contingenti e in associazione ad altri strumenti più familiari: dai dizionari ai motori di ricerca online che arricchiscono l’esplorazione autonoma degli studenti e fungono da ponte verso un impiego più tecnico dei corpora. Gli esempi trattati e le esperienze descritte riguardano corsi universitari di primo e secondo livello e studenti con una conoscenza della lingua inglese pari almeno al B1 ma, con ogni probabilità, lo stesso approccio può essere proficuo e adattabile anche a classi di livello inferiore e ad ambiti scolastici ed educativi diversi.

1.1. L’uso diretto di corpora da parte degli apprendenti

L’utilizzo dei corpora per la didattica e l’apprendimento delle lingue è stato proposto e trattato da molti studi (si vedano ad esempio Aston, 2001; 2004; Boulton, 2010; Chambers, 2005; 2007; Frankenberg-Garcia et al., 2011; Reppen, 2010; Tribble et al., 1990) e riassunto, nelle sue linee fondamentali, in uno studio di McEnery et al. (1990) che, citando Leech (1979), propone una distinzione di base (ripresa da molti autori tra cui Römer (2008: 113)) tra un uso di corpora di tipo *diretto* e *indiretto*. Nel secondo caso sono gli insegnanti, gli studiosi o i

creatori di materiali didattici a utilizzarli per approntare testi, esercitazioni, verifiche e altre attività che si vogliano valorizzare basandosi su documenti autentici e contesti d'uso appropriati. Si va dalla elaborazione di dizionari e sussidi didattici alla creazione di esercizi che mostrano concordanze con l'omissione della parola chiave centrale che gli studenti devono, dal contesto, cercare di intuire; da esempi che illustrano la prosodia semantica di una espressione ad altri che ne indagano le collocazioni più frequenti. Gli apprendenti restano utilizzatori per lo più passivi di quanto viene elaborato per loro da altri.

Molto più rilevante per gli scopi di questo lavoro è l'approccio di tipo diretto che, al contrario, si incentra su un accesso non preordinato ai corpora ed alle informazioni linguistiche da parte degli apprendenti. Ad una fase di introduzione ai corpora (*teaching about*) che riguarda aspetti teorici e metodologici, segue una fase in cui si pone l'accento su *come* utilizzare i corpora per studiare ed esplorare i dati linguistici: «direct uses of corpora include “teaching about”, “teaching to exploit”, and “exploiting to teach”, with the latter two relating to how to use.» (McEnery et al., 1990: 370).

Durante questa esplorazione diretta di testi autentici (suggerita da approcci quali il *Corpus Based Learning*, cit: 370, ed il *Data Driven Learning* o DDL, Johns, 1991; 1994), l'apprendente viene incoraggiato a comportarsi come un ricercatore e ad utilizzare i corpora in autonomia, senza percorsi preordinati:

(1) Johns was perhaps among the first to realize the potential of corpora for language learners (e.g. Higgins & Johns, 1984). In his opinion, “research is too serious to be left to the researchers” (Johns, 1991, p. 2). As such, he argues that the language learner should be encouraged to become “a research worker whose learning needs to be driven by access to linguistic data” (McEnery et al., 1990: 370).

Ciò implica che l'elaborazione e l'interpretazione di concordanze, liste di frequenza, raffronti e indagini viene lasciato all'apprendente, che deve necessariamente assimilare, almeno in parte, le competenze tecniche e la metodologia degli utenti più esperti: ad esempio, la lettura verticale delle concordanze (Tognini Bonelli, 2001:18), l'impiego di caratteri jolly (*wildcards*) e l'importanza del valutare la qualità e rilevanza dei risultati ottenuti. Queste capacità interpretative vengono solitamente demandate alla fase del *teaching about*, il cui inserimento nel curriculum scolastico o universitario dei normali corsi di lingua straniera può, tuttavia, risultare problematico e sollevare non poche perplessità. Altrettanto problematico può essere il fatto che, se un approccio indiretto può limitarsi all'utilizzo in classe di elementi, materiali o ricerche specifiche con tempi ed esiti prevedibili e misurabili, l'indagine diretta da parte di singoli o gruppi di studenti spesso richiede tempi e percorsi di esplorazione non prefissati, risultati non sempre prevedibili e la disponibilità, non sempre scontata o permessa, di un accesso ad un computer o quantomeno ad internet. Infine, assumendo che la validità di questo metodo non risieda tanto nel raggiungimento di competenze linguistiche o grammaticali specifiche, quanto nel processo e nell'acquisizione da parte degli apprendenti di strategie, consapevolezza e capacità nel gestire problemi linguistici (di trasformarsi, come si è detto, in una sorta di ricercatore), anche la difficoltà di una verifica dei risultati può costituire un inconveniente serio per la normale programmazione didattica.

Rispetto ad altri contesti educativi, l'esplorazione basata su corpora è sicuramente più facilmente inseribile in corsi in ambito universitario, dove percorsi ed esiti dipendono maggiormente da una programmazione gestita dal docente. Diversi studi trattano casi in cui agli studenti viene chiesto di indagare o anche compilare un proprio corpus per scopi che vanno dalla traduzione all'analisi testuale o del discorso. Anche in questo ambito, tuttavia, a meno che non si tratti di corsi molto specifici e destinati a studenti con un interesse verso la linguistica, l'acquisizione degli aspetti più tecnici può rappresentare un ostacolo preliminare rilevante. Nei miei primi corsi capitava più volte di raccogliere giudizi che sottolineavano come l'attenzione degli studenti fosse rivolta principalmente all'apprendimento della lingua o a studi di linguistica, e in cui l'introduzione di tecnicismi informatici o statistici ed elementi metodologici della linguistica dei corpora veniva considerata superflua, bizzarra o irrilevante.

Questi inevitabili problemi di adattamento alla normale didattica delle lingue sono probabilmente tra gli elementi principali che hanno impedito, come rilevato in diversi studi (si vedano tra gli altri Boulton, 2010; Meunier, 2011; Römer, 2008), una maggiore diffusione dei corpora in classe. La tesi sostenuta in questo contributo, come vedremo, è che questi problemi possano essere in larga parte superati, che la prima fase di introduzione, il *teaching about* possa e debba essere ridotta al minimo e che diversi tecnicismi di ricerca linguistica necessari per poter assimilare l'apprendente ad un ricercatore siano divenuti meno ostici grazie alla crescente familiarità dei nostri studenti con internet e gli strumenti di ricerca in rete. Si propone, in altre parole, di minimizzare gli aspetti teorici o tecnici e concentrare l'attenzione sulla fase del “come” sfruttare informazioni e dati linguistici autentici. L'utilizzo di corpora, preferibilmente online, non costituisce una premessa o un fine della didattica, ma uno strumento tra gli altri disponibili in rete per l'accesso diretto ed autonomo a informazioni, testi e contesti autentici. Per facilitare la fase del “come” sfruttare al meglio queste risorse è tuttavia molto utile applicare la metodologia di base della linguistica dei corpora e capire come l'esplorazione di testi autentici permetta un uso più fluente ed appropriato della lingua straniera.

1.2 Il lexical priming e l'uso appropriato della lingua straniera

Un riferimento teorico che ritengo possa essere molto utile a capire come l'esplorazione di materiale linguistico autentico e la metodologia propria della linguistica dei corpora possa aiutare l'apprendimento è il concetto di *Lexical Priming* definito da Hoey come:

(2) the process whereby “[a]s a word is acquired through encounters with it in speech and writing, it becomes cumulatively loaded with the contexts and co-texts in which it is encountered, and our knowledge of it includes the fact that it co-occurs with certain other words in certain kinds of context” (2005: 8)

Hoey, come altri promotori della linguistica dei corpora (si vedano ad esempio Sinclair, 1991; Partington, 1998), pone l'accento sul fatto che l'apprendimento di una lingua non può prescindere dal suo utilizzo in contesti e situazioni definite, dalla tendenza delle parole a creare legami con altri elementi lessicali e ad instaurare stringhe o *patterns* linguistici ricorrenti. Tra gli esempi citati da Hoey (2009) possiamo ricordare *patterns* come *drive somebody crazy, go crazy* o *the time has come*. Altri semplici esempi sono le frasi idiomatiche ed i proverbi come *when the going gets tough* (Zanca, 2017: 117) e *hai voluto la bicicletta e adesso pedala* (Zanca, 2014), o frasi appropriate e ricorrenti in determinati contesti come *good evening ladies and gentlemen*. La capacità dei parlanti nativi di essere fluenti risiede, secondo l'autore, soprattutto nel saper utilizzare non singoli elementi lessicali

legati da regole, ma *patterns* linguistici assimilati con l'esposizione inconscia e ripetuta a interazioni linguistiche autentiche e di saperlo fare nel contesto appropriato:

(3) This process of subconsciously noticing is referred to as lexical priming. Noticing all these things is what makes it possible for a speaker to use the right phrase in the right context at the right time. Without realizing what we are doing, we all reproduce in our own speech and writing the language we have heard or read before. We use the words and phrases in the contexts in which we have heard them used, with the meanings we have subconsciously identified as belonging to them and employing the same grammar. This is how native speakers are able to be fluent and because the things they say are subconsciously influenced by what everyone has previously said to them, it also explains why they almost always sound natural (Hoey, 2009: 1)

Da questa premessa si desume che, in buona parte, l'inadeguatezza di elaborazione linguistica degli apprendenti deriva da una loro minore frequenza di esposizione al linguaggio ed al contesto in cui viene utilizzato e, in secondo luogo, dal fatto che il loro *priming* nella lingua madre fornisce contesti d'uso e scelte linguistiche che possono suggerire scelte sbagliate:

(4) The differences in practice between a native speaker and a non-native speaker are twofold. Firstly, a non-native speaker is typically exposed to less language and to a narrower range of language, and, secondly, the non-native speaker has previously been primed for another language, which initially affects the way he or she is primed in English (ibid. 2009: 1).

A differenza di quanto avviene per gli approcci comunicativi all'insegnamento linguistico, non si pone l'accento soltanto sulla capacità di veicolare il messaggio e garantire la comprensione reciproca, ma sull'appropriatezza pragmatica e contestuale. Se un apprendente utilizza l'espressione *do* somebody crazy*, il significato implicito viene in qualche modo trasmesso, ma l'esperienza dei parlanti nativi suggerisce che il verbo da usare sia invece *drive* oppure *make* con diverse preferenze contestuali e quantitative. Cosa più rilevante per i nostri studenti, una scelta "sbagliata" in questo campo sarebbe sanzionata in un test di inglese e produce ciò che spesso etichettiamo come qualcosa che 'non si dice' o 'non suona bene', anche se non è formalmente sbagliato in termini semantici, lessicali o grammaticali. Per ovviare a questo tipo di carenza linguistica Hoey suggerisce agli insegnanti di focalizzare l'attenzione su elementi che si possono collegare in maniera evidente alla linguistica dei corpora:

(5) Native speakers have acquired a large corpus of examples of the words of English in their typical contexts, and from this they learn how the words are used. [...] learners need to be exposed to as much authentic material as possible and teachers need to find ways of accelerating the priming processes by drawing attention to the patterns of use that a text (or conversation) reveals (ibid. 2009:1).

In assenza di un 'corpus interno' sufficientemente ampio di esempi linguistici proprio dei parlanti nativi, che ovviamente varia enormemente da individuo a individuo ma che permette loro di essere fluenti nella maggior parte delle situazioni, l'apprendente dovrebbe poter esplorare materiali autentici, mettendo in evidenza ricorrenze e frequenze d'uso rivelatrici di *patterns* ricorrenti. Una parte dell'insegnamento di una lingua dovrebbe quindi consistere nel mettere in grado l'apprendente di condurre questa esplorazione con gli strumenti, come viene implicitamente suggerito, propri della linguistica dei corpora. Una volta assimilata la capacità di usare questi strumenti in modo autonomo, l'apprendente dovrebbe poter compensare le proprie carenze in termini di *priming* (o *priming gap*) acquisendo consapevolezza e fiducia nelle proprie possibilità di avvicinarsi alla performance dei parlanti nativi. L'approccio descritto di seguito ha proprio questo intento.

2. Metodologia

Come accennato in precedenza, se da un lato l'impiego dei corpora in classe e nella didattica delle lingue straniere sembra poter fornire notevoli vantaggi ed opportunità, dall'altro esso fatica ad imporsi tra insegnanti e studenti. Una delle motivazioni principali per questo stato di cose, è che sovente agli apprendenti è richiesto un passaggio preliminare consistente nello studio delle peculiarità dei vari corpora a disposizione e nell'acquisizione della teoria e delle tecniche di base della linguistica dei corpora. Questo passaggio è considerato irrinunciabile in quasi tutta la letteratura su questo tema. Nelle prossime pagine verrà invece presentato un metodo utilizzato in classe negli ultimi anni sia con studenti universitari avanzati con interessi e conoscenze di linguistica^[1], sia con studenti con un livello pari ad un B1 e con percorsi accademici di tipo letterario o delle scienze politiche^[2]. In entrambi i casi si tratta di corsi universitari – e non di corsi di lingua per la comunicazione – incentrati sulla traduzione e sulla scrittura di testi accademici come strumento per analizzare il linguaggio e per acquisire autonomia e consapevolezza nell'utilizzo dell'inglese come lingua straniera. Nel dettaglio, per il loro esame finale agli studenti è stato chiesto:

- a. studenti primo anno di laurea triennale: scegliere un articolo di giornale di un minimo di 500 ed un massimo di 800 parole che non sia stato ripreso dalle agenzie di stampa italiane e tradurlo dall'inglese all'italiano;
- b. studenti primo anno magistrale: scegliere un argomento accademico di interesse, raccogliere una serie di articoli accademici in inglese e nella lingua madre relativi all'argomento scelto, scrivere una *research paper* (RP) in inglese e tradurre in inglese un testo accademico. Il RP, in genere, non implica la necessità di svolgere una vera ricerca, ma consiste nel descrivere il corpus raccolto e tracciare alcuni aspetti strutturali e/o linguistici che lo distinguono attraverso una breve indagine quantitativa.

Dal momento che agli studenti più avanzati (b) viene chiesto anche di compilare un piccolo *disposable corpus* (Varantola, 2003) e di usare un software gratuito di concordanze (normalmente ANTCONC^[3] di Lawrence Anthony), il loro corso prevede anche una introduzione più tecnica alla linguistica dei corpora. In entrambi i casi, tuttavia, e in maniera esclusiva per gli studenti (a), l'impiego dei corpora viene finalizzato al lavoro di traduzione o scrittura e considerato solo uno degli strumenti disponibili. Questo approccio è già stato oggetto di precedenti studi e pubblicazioni (Zanca, 2013; 2014; 2017) e può essere accostato, con diverse angolazioni, a studi più recenti come quello sull'uso di Google e internet per attività DDL (Boulton, 2015; Eu, 2017; Conroy, 2010; Geiller, 2014: 28; Sha, 2010; Zanettin, 2009; **questo volume**). L'idea di base consiste nell'avvicinare l'apprendente all'utilizzo dei corpora partendo da risorse online di uso quotidiano e comunque più familiari quali i motori di ricerca, i siti per il reperimento di informazioni ed i materiali di riferimento disponibili in rete. I corpora, al pari di queste risorse, devono essere percepiti come mezzo per risolvere problemi contingenti e non come fine del processo didattico. La familiarità che oggi tutti gli studenti hanno con la navigazione in internet e

con il reperimento di informazioni online utilizzando computer, tablet, telefono e TV, viene sfruttata per proporre esplorazioni e analisi dei vari contesti in cui il linguaggio si realizza o, facendo riferimento a Hoey, per ridurre il loro *priming gap*. Invece di iniziare i corsi con una introduzione teorica e/o pratica alla linguistica dei corpora e dei software correlati, come viene suggerito nella maggior parte dei casi presentati in letteratura, si parte da problemi linguistici contingenti legati ai task che gli studenti saranno chiamati a portare a termine in vista dell'esame finale, cercando di risolverli con l'impiego, in classe e a casa, di risorse normalmente bandite dai docenti e dai puristi della linguistica dei corpora come, ad esempio, Wikipedia, forum, blog, e Google o altri motori di ricerca (vedi par. 3 e seguenti). Ad esse si aggiungono strumenti più tradizionali, ma preferibilmente nella loro versione online (grammatiche, database terminologici, glossari, dizionari ecc.) e l'uso del thesaurus di sinonimi del word processor preferito. L'utilizzo di corpora, come descritto nel paragrafo 4, anche se offre modalità di esplorazione uniche, viene introdotto come uno dei tanti possibili strumenti a disposizione. Esempi specifici di questo approccio (vedi anche Zanca, 2013; 2014; 2017) saranno presentati qui di seguito e, più in dettaglio al paragrafo 5.

3. L'esplorazione, le risorse online e alcuni esempi

Lo scopo delle attività proposte agli studenti, come accennato, è l'esplorazione di esempi di linguaggio autentico e del contesto in cui questo si realizza per promuoverne la consapevolezza, la capacità e l'autonomia nel gestire situazioni di uso linguistico in cui il loro *priming* è inadeguato. Buona parte della valutazione finale, quindi, consiste nella verifica dell'impiego efficace di una metodologia di esplorazione e di strumenti appropriati più che di una produzione accurata, che comunque resta un obiettivo importante del corso. Per questo, anche se negli esempi riportati al paragrafo 5 sono evidenti alcune scelte traduttive discutibili o inappropriate, queste non diminuiscono il valore dello sforzo di esplorazione linguistica e non pregiudicano una valutazione positiva del lavoro fatto dagli apprendenti. Questo vale anche e soprattutto quando essi vengono incoraggiati ad utilizzare gli strumenti più conosciuti e familiari, ma diversi dai corpora, citati in precedenza. Agli studenti, inoltre, viene chiesto di esplicitare in classe e aggiungere al proprio elaborato finale un commento sul percorso di esplorazione e sulle scelte compiute durante il processo di scrittura e traduzione e molti degli esempi elencati di seguito, oltre che nel paragrafo 5, sono tratti da questi loro commenti.

3.1. Google e i motori di ricerca

Google – probabilmente lo strumento più noto a tutti gli utenti di internet – è normalmente il punto di partenza per l'esplorazione linguistica in classe. Tutti lo utilizziamo, ma uno degli obiettivi del corso è di incoraggiare una metodologia di ricerca più attendibile di quella messa in atto quotidianamente dalla maggior parte degli utenti di questo ed altri motori di ricerca.

3.1.1. Motori di ricerca per il controllo di spelling o sintassi

In alcuni casi, i motori di ricerca possono fornire risposte immediate ed affidabili, come nel caso di dubbi sullo *spelling* delle parole o della struttura dei *phrasal verbs*. In questi casi, è spesso sufficiente inserire in Google parole o frasi con errori, come *independent*, *pronunciation*, *I'm looking forward to see you* o *Look this picture*, perché gli algoritmi di Google rilevino il problema e suggeriscano possibili soluzioni. Come vedremo di seguito, anche in questo caso, qualora sussistano dubbi sulla qualità o l'attendibilità dei suggerimenti di Google, gli studenti vengono invitati ad attivare delle procedure di controllo con l'utilizzo di risorse o riferimenti testuali più sicuri.

3.1.2. Gli operatori avanzati per la ricerca

Questo primo impiego di Google, che tutti noi adottiamo, può essere affinato in modo semplice, ma non noto a tutti gli utenti, con il ricorso a operatori di ricerca avanzati. L'uso delle virgolette, ad esempio, ci permette di chiedere una corrispondenza esatta della espressione ricercata. Se digitiamo *I took up it* senza virgolette, otteniamo una serie di risultati relativi ai singoli componenti, mentre se digitiamo "I took up it" ci viene immediatamente segnalato che potrebbe trattarsi di un errore: le prime voci riportano infatti l'espressione tra gli errori più comuni in inglese. Anche quest'ultima ricerca, tuttavia, produce centinaia di migliaia di risultati, tra i quali molti non sono identificabili come errori, e solo l'abitudine ad una lettura critica delle occorrenze e/o un restringimento della ricerca a fonti più attendibili (vedi oltre) può permettere agli studenti di capire che non si tratta di risultati attendibili.

L'asterisco ci permette di accettare qualsiasi parola o stringa in quella posizione all'interno di una espressione tra virgolette: "I'm looking * to" ci permetterà di individuare il *pattern* più ricorrente per il quale la parola mancante è *forward*; "can he cut the*" ci offrirà la parola finale con una occorrenza più frequente, ossia *mustard*. Altri operatori utili possono essere il segno meno per escludere la parola seguente dalla ricerca, *OR* per accettare sia la parola precedente che la seguente come *alternative* ecc. Google offre anche una interfaccia apposita per ricerche 'avanzate' reperibile all'indirizzo: https://www.google.com/advanced_search.

Anche se non è strettamente un operatore ed è ormai di uso comune, è importante far rilevare che i motori di ricerca rilevano anche la co-occorrenza dei termini cercati, per cui se si cerca *collocati* si otterranno soprattutto risultati in ambito economico, mentre basta aggiungere la parola *corpora* per ottenere risultati per lo più relativi alla linguistica. Allo stesso modo, può essere utile aggiungere al termine in inglese a cui si è interessati le parole "traduzione" o "synonyms" per ottenere link a dizionari e siti dedicati alla traduzione.

3.1.3. Restringere il campo della ricerca

Oltre a questo primo utilizzo, indagato più approfonditamente in altri studi (tra gli altri, Chinnery, 2008; Conroy, 2010; Eu, 2017; Geiler, 2014; Sha, 2010; Wu et al., 2009), Google permette di ottenere risultati più simili per attendibilità a quelli dei corpora, attraverso la ricerca non dell'intero web ma di sue sezioni più limitate. Se cerchiamo la stringa "I took up it" tra virgolette in Google notizie, Google books o Google scholar^[4], ad esempio, otteniamo un numero infinitamente minore di quello ottenuto con la ricerca nell'intero web, rendendo più facile la valutazione qualitativa dei risultati. Se cerchiamo "I'm looking * to" nello stesso ambito, otterremo una serie di risultati provenienti da testi più complessi ed è quindi possibile, ad esempio, vedere che ciò che segue sono verbi al participio presente oppure pronomi o nomi e, in generale, è possibile analizzare come questa espressione 'si comporta' in testi autentici e sufficientemente affidabili, proprio come avviene per le concordanze generate da corpora. Se si cerca l'espressione *made your bed "now * in it"* (Zanca, 2014) si vedrà come in testi autentici la

parola mancante del proverbio non è sempre *lie* ma a volte viene sostituita da *sleep, you can lie, you have to lie* ecc..

Il campo di ricerca può essere limitato, ottenendo risultati più omogenei, anche utilizzando l'operatore 'site:' che indica uno specifico dominio. Ad esempio l'espressione *perform a survey* che in uno studio di Cowie (1994: 3169) veniva individuato come esempio di frase non utilizzata a favore di *conduct a survey* e che effettivamente non compare in corpora "tradizionali" come il British National Corpus (BNC) ed il Corpus of Contemporary American English (COCA), può essere ricercata in Google con la stringa "*perform a survey*" *site:.ac.uk*, "*perform a survey*" *site:.edu*, o "*perform a survey*" *site:.mit.edu*. Si otterranno non pochi risultati (nell'ultimo caso più di 3000 al momento della stesura di questo studio) provenienti da siti accademici che mettono in dubbio l'impossibilità o l'inadeguatezza sostenuta da Cowie perché ad una analisi qualitativa risultano essere prodotti da ricercatori o accademici.

Ancora, può essere utile restringere la nostra ricerca ad un periodo di tempo limitato. In uno studio precedente (Zanca, 2014) si è evidenziato come uno studente, che, dopo aver consultato un blog di traduzione, temeva che l'espressione *you made your bed, "now lie in it"* fosse obsoleta, aveva potuto appurare con una ricognizione su Google notizie come, al contrario, i giornali la utilizzassero centinaia di volte nei tre mesi precedenti.

3.1.4. Google immagini

Anche il ricorso a Google immagini può rivelarsi particolarmente rapido ed efficace per risolvere problemi di interpretazione o resa del testo. Innanzitutto rende più facile la comprensione di elementi lessicali semplici e composti. Digitando tra virgolette "*phone card*" e "*card phone*" oppure "*high window*" o "*tall window*" si ha immediatamente una idea della differenza esistente tra le diverse espressioni cambiandone l'ordine sintattico o gli aggettivi. Allo stesso modo, le immagini chiariscono in una frazione di secondo cosa sia un *Toad in the hole*. In diversi altri casi, inoltre, Google images è servito agli studenti ad individuare aspetti culturali non immediatamente percepibili, come il riferimento a Shakespeare nel titolo di giornale *The gentleman of Verona and his incredible sausage*[5]. Lo studente che doveva tradurre l'articolo si è reso conto del riferimento letterario solo digitando "*The gentleman of Verona*" in Google immagini ed ottenendo immagini riferibili alla nota commedia. Lo stesso è accaduto per il titolo *I'll happily take the credit for a dog day afternoon*[6] in cui il riferimento è, questa volta, prevalentemente cinematografico.

3.1.5. Google Books Ngram Viewer

Un utilizzo di Google leggermente più complesso è quello offerto da *Google Books Ngram Viewer*, che visualizza graficamente le occorrenze delle parole scelte nella collezione di libri in diverse lingue e varietà dell'inglese e in diversi periodi storici. Questo strumento è stato utilizzato in classe per indagare, ad esempio, la differenza nell'uso di *Health Minister* o *Minister of Health* (quest'ultimo usato prevalentemente negli Stati Uniti) nell'inglese britannico e americano e in prospettiva diacronica. Oppure per visualizzare l'impiego del già citato *perform a survey* rispetto a *conduct a survey*.

3.1.6. Wikipedia e informazioni enciclopediche

Wikipedia o altre pagine che offrono informazioni di tipo enciclopedico sono risorse molto familiari agli studenti ed anche il loro utilizzo viene incoraggiato. Lo scopo principale, in questo caso, è il reperimento di informazioni sui partecipanti coinvolti nel discorso e su elementi che possono rivelarsi problematici per la comprensione e la traduzione dei testi, come acronimi, nomi di luoghi, persone enti ecc.. Una breve ricerca, ad esempio, permette di capire cos'è e come tradurre meglio *Faculty of Public Health Medicine*[7] senza confondere questo istituto con una Facoltà universitaria. Molto utile è anche il facile accesso spesso fornito da Wikipedia al nome scientifico in latino di molti termini. Ad esempio, se in un articolo[8] troviamo *rat snakes, milk snakes, garter snakes and vipers* è possibile ricavarne la denominazione scientifica per poi ricercarla sulla edizione italiana di Wikipedia o in altri siti per ottenere almeno una ipotesi di traduzione nell'italiano corrente.

3.1.7. Blogs, forum per traduttori e domande online

Le possibilità offerte dalla rete oggi includono in maniera crescente la possibilità di interagire con altri utenti. Anche se non accade spesso, alcuni studenti mi hanno riferito di essersi rivolti a parlanti nativi o esperti di linguaggio specialistico per affrontare un determinato problema. Esistono poi blog specifici per porre domande a traduttori e parlanti nativi come avvenuto per il citato esempio di *you made your bed, now lie in it*[9] che poi restano disponibili per la consultazione online. Questa possibilità, che può coinvolgere l'apprendente in una interazione autentica nella lingua straniera, merita senza dubbio un approfondimento e comunque rientra tra le opzioni di esplorazione incoraggiate nel corso.

3.1.8. Dizionari, Glossari, Google translate

L'utilizzo di dizionari, in un corso che prevede una traduzione, è un elemento scontato. Meno scontato è il ricorso a dizionari online che spesso rendono più facile l'esplorazione del contesto linguistico ed extralinguistico in modi simili a quelli presentati in precedenza. Brevemente, oltre alla possibilità di ricercare non solo singole parole ma stringhe di diverse parole, molti dizionari online offrono esempi tratti da testi autentici simili a concordanze, immagini, link a forum interattivi, glossari, dizionari di sinonimi ecc.. Una nota a parte meritano i traduttori automatici come Google Translate. Questo tipo di risorsa, che fa comunque parte di un settore di ricerca scientifica molto importante, può essere semplicemente portata ad esempio di come non fare una traduzione accurata. Basta far inserire agli studenti una frase come la menzionata *rat snakes, milk snakes, garter snakes and vipers* per ottenere un divertente *serpenti di ratti, serpenti di latte, serpenti giarrettiera e vipere* e suggerire poi il procedimento con Wikipedia menzionato in precedenza.

3.1.9. Reverso Context, Linguee e corpora paralleli online

Nel caso di strumenti come Reverso Context o Linguee, che vengono spesso suggeriti da Google se si aggiunge la parola *translation* al termine cercato, l'accostamento ai corpora è molto evidente. Agli studenti non viene offerta una traduzione o una spiegazione predeterminata, ma una serie di concordanze nella lingua di partenza ed in quella di arrivo tratte da corpora paralleli, lasciando a loro il compito di giudicare la rilevanza dei risultati ottenuti. Il ricorso a questo strumento, quindi, è molto utile per avvicinare gli studenti all'uso dei corpora, ma anche per mettere sull'avviso riguardo ad una delle leggi fondamentali sui corpora e cioè che «the results are

only as good as the corpus» (Sinclair, 1991: 13). I risultati offerti, infatti, non sono sempre del tutto affidabili perché in questo caso provenienti da traduzioni reperite in rete e non sempre di comprovata qualità. Ad esempio, inserendo in Reverso Context *rat snakes*, si ottengono solo due occorrenze dallo stesso testo che propongono la traduzione *serpenti neri* che non è probabilmente l'opzione migliore (più idonea è la scelta di *serpente dei ratti* o il colloquiale *saettone*).

3.1.10. Altre risorse online

Naturalmente la lista delle risorse online a disposizione di una indagine linguistica non si esaurisce con quelle elencate. Si potrebbero menzionare thesauri accattivanti come “Visual thesaurus”, database terminologici, database fraseologici, glossari specifici, grammatiche, software di analisi del testo ed una miriade di altre risorse in continuo divenire.

4. Dalle risorse online ai corpora

Quelli presentati sono esempi di strumenti di facile ed immediato impiego da parte dell'apprendente che tuttavia, sono spesso sufficienti a suggerire soluzioni appropriate o quantomeno ipotesi di soluzione al problema linguistico di partenza. In pochi istanti, lo studente può iniziare un'esplorazione dell'effettivo utilizzo di determinati elementi lessicali e della loro interazione con altri elementi, del loro contesto d'uso, di aspetti extralinguistici come le immagini ad essi associate e di possibili variabili temporali, testuali o culturali di rilievo.

In mancanza di una guida ad una metodologia di ricerca e ad una valutazione critica dei risultati, tuttavia, questo tipo di esplorazione e le ipotesi interpretative e produttive che ne scaturiscono, possono essere fuorvianti. Soprattutto nel caso di risorse come Google, che hanno il grande merito di fornire una quantità sterminata di informazioni, il rischio che buona parte di esse siano inaffidabili è molto forte e può essere ridimensionato soltanto con procedure di ricerca e verifica più accurate. In molti di questi casi, l'impiego di data base testuali e/o corpora di provata attendibilità si rivela utile e viene accettata, o almeno tollerata come un male necessario, dagli studenti.

4.1 Perché i corpora

Secondo l'approccio delineato, l'utilizzo dei corpora può essere proposto agli studenti – sia nel contesto universitario descritto ma anche, con ogni probabilità, in altri ambiti di apprendimento – per due motivi principali: da un lato consente di percorrere sentieri e risolvere problemi altrimenti non affrontabili e dall'altro può costituire un valido strumento di controllo dei risultati ottenuti con risorse meno affidabili.

In quest'ultima accezione il percorso metodologico di base (vedi Zanca, 2013: 328) può prevedere la seguente routine:

a. Problema linguistico / *research question* → b. Uso di risorse online per ottenere risposte possibili → c. controllo delle risposte ottenute con l'uso di corpora e dati autentici affidabili → d. possibile rilievo di risposte diverse ed impreviste → e. controllo di queste nuove risposte tornando al punto b della routine.

Quello che è possibile fare con i corpora e non con altre risorse online si sta via via riducendo: abbiamo visto come sia i dizionari online che strumenti come Reverso Context presentino dati provenienti da corpora linguistici e come sia possibile fare ricerche con Google in contesti che, per le necessità di un apprendente medio, sono qualitativamente sufficientemente affidabili come Google Scholar o Google Books. Tuttavia per alcuni aspetti essi sono ancora insostituibili. Innanzitutto i testi che compongono i principali corpora sono stati selezionati e sono noti. Anche se questo elemento, di per sé, non è di molto interesse per l'apprendente medio, il fatto che esso garantisca una maggiore affidabilità dei risultati lo è. Concretamente, se applicando la routine precedente l'ipotesi interpretativa scaturita dalla esplorazione online viene confermata da un corpus, l'apprendente può confidare maggiormente nella correttezza della risposta.

Più rilevante e specifica all'esplorazione di corpora è la possibilità di fare ricerche complesse ed evidenziare *patterns* ed aspetti associativi del linguaggio. Utilizzando software o interfacce dedicate come ANTCOINC o il sito della Brigham Young University^[10], è possibile, ad esempio usare l'operatore asterisco per sostituire i caratteri iniziali o finali di una parola (es. *divers** per *diverse*, *diversity* e *diversities*); ricercare più parole insieme, anche se divise da altri elementi; distinguere, nei corpora annotati, tra le categorie grammaticali delle parole ricercate (es. cercare *drink* come nome e non come verbo); restringere la ricerca a parti specifiche del corpus (solo testi accademici o giornalistici o annate specifiche) ecc.. Le tipiche liste di concordanze, che è possibile ordinare a piacimento, permettono di concentrare l'attenzione su aspetti di rilievo e, a volte, di far emergere caratteristiche inattese (ad es. *can't cut it* come alternativa a *can't cut the mustard*. Le liste di collocazioni calcolate dal software facilitano l'esplorazione delle associazioni lessicali e le liste di frequenza per singole parole (*wordlists*) o per stringhe di parole (*cluster lists*), consentono di analizzare gli elementi lessicali e fraseologici e fare confronti statistici tra diversi corpora.

Alcune di queste funzioni, come il confronto tra corpora e la compilazione di *wordlists* o *cluster lists*, sono probabilmente ancora troppo tecniche per la maggior parte degli studenti non avanzati, ma di alcune altre, come le opportunità di ricerca con *wildcards*, il calcolo delle collocazioni e la visualizzazione delle concordanze, gli studenti imparano presto l'utilità e l'unicità rispetto ad altre risorse a disposizione. I corpora, in altre parole, permettono esplorazioni e possono fornire risposte che altri strumenti non sono in grado di fornire o che non possono fornire in modo sufficientemente affidabile, e la loro reperibilità online li rende uno degli strumenti di facile accesso ed utilizzo.

Considerare i corpora, come sembra avvenire nella maggior parte degli approcci tradizionali alla DDL, come una risorsa risolutiva e superiore alle altre in ogni circostanza è tuttavia un errore. Si è già accennato a ricerche che beneficiano di immagini o informazioni enciclopediche, ma un'ulteriore conferma dell'opportunità di una sinergia tra le varie risorse si ottiene quando, per alcuni problemi riscontrabili dagli apprendenti, l'uso di corpora, dizionari ed altre risorse tradizionali può essere insufficiente. Ad esempio, nel caso dell'espressione idiomatica *Then the north Londoners got the cigars out* presente in un articolo^[11] scelto da uno studente, la ricerca su vari dizionari è stata improduttiva. Lo stesso si è verificato per la ricerca su Reverso Context, su data-base terminologici o fraseologici ed anche nei maggiori corpora online (BNC, COCA e persino GloWbE, che consta di 1,9 miliardi di parole). Google News, al contrario ci offre 2 occorrenze per *got the cigars out* e 11 per *get the cigars out*. Da queste si evince facilmente che si tratta di un'espressione utilizzata nel mondo dello sport con il significato approssimativo di 'appendere gli scarpini al chiodo' (scarpini è più frequente di scarpe o scarpette in

Google News) o ‘tirare i remi in barca’. Solo un corpus enorme di notizie dal web come News on the Web (NOW), che per la sua composizione non è molto diverso da Google News e che ad oggi consta di 5,51 miliardi di parole, si avvicina a Google News, offrendo tuttavia soltanto un risultato per *got the cigars out* e 5 occorrenze per *get the cigars out*. La capacità di utilizzare in modo appropriato diversi strumenti e di esplorare il linguaggio sotto diverse angolazioni si può rivelare più utile che arroccarsi su alcune posizioni preordinate.

5. Alcuni esempi

Alcune esperienze di applicazione dell’approccio suggerito in questo studio e della routine di verifica di ipotesi interpretative e traduttive sono già stati descritti in precedenti lavori incentrati soprattutto sulle espressioni idiomatiche (Zanca, 2013; 2014; 2017). Inoltre, quasi tutti gli esempi utilizzati in precedenza provengono da elementi linguistici evidenziati nel lavoro con gli studenti. In questa sede credo possa essere utile focalizzare l’attenzione su esempi di applicazione delle strategie di esplorazione suggerite e sul feedback ottenuto dagli studenti. Come accennato in precedenza, per il lavoro di traduzione, che svolgono a casa e per il quale sono incoraggiati ad usare risorse online e corpora, gli studenti devono commentare i loro percorsi esplorativi e le loro scelte traduttive. Per farlo hanno a disposizione una tabella con tre colonne: la prima per il testo originale, quella centrale per la traduzione e quella a destra per i loro commenti. Gli esempi riportati riguardano traduzioni dall’italiano (L1) all’inglese (L2) o dall’inglese (L2) all’italiano (L1); i commenti sono solitamente, ma non obbligatoriamente, in inglese e sono stati lasciati integri, anche quando presentano errori ortografici o grammaticali.

Nel caso della figura 1, lo studente ha scelto di tradurre un articolo di cronaca sportiva dedicato al campionato di basket italiano. Questo tipo di testo è denso di termini e locuzioni specifiche che pongono diversi problemi al traduttore. Il commento dimostra che lo studente ha utilizzato la metodologia proposta per analizzare, in contesti ritrovati in Google News, la diversa frequenza delle alternative prese in considerazione (*run*, *spurt* e *stretch*) e per riscontrare che *ignited* in contesti simili sia un collocato di *run*.

Testo L1	Testo L2	Commento dello studente
Cantù ha provato a scappare con la tripla di Pilepic per il 42-35 ma in quel momento Varese ha reagito col 10-0 ispirato da Maynor (44-46).	Cantù tried to pull away on Pilepic's three-pointer at 42-35, but Varese responded right away with a 10-0 run ignited by Maynor (44-46).	Amplification: 10-0 run ('spurt' or 'stretch' are also candidates, but 'run' is far the most frequent) Collocation: In basketball news a "run" is followed by 'ignited by' in the context of somebody/thing being the reason for the "run".

Fig. 1: Frequenze e collocazioni

Il testo scelto negli esempi in figura 2 e 3, è un articolo di cronaca dedicato alla tragica morte dell’attore Memé Perlini. Il traduttore è uno studente Erasmus la cui lingua madre è diversa sia dall’italiano che dall’inglese per cui il percorso di riduzione del *priming gap* riguarda sia la lingua di partenza che quella di arrivo. Nel primo caso (Fig. 2) le risorse online (Reverso context, il dizionario online del Corriere della Sera) sono state utilizzate dapprima per identificare possibili opzioni di traduzione (*palace*, *building* o *apartment block*) e poi per attivare la routine di controllo con il ricorso, in questo caso, a Google immagini e anche a Google maps che hanno permesso di esplorare il contesto extralinguistico di riferimento. Da ultimo, Google e Google News sono stati utili per decidere se tradurre o meno *rione Esquilino* fosse parte delle convenzioni del discorso giornalistico della lingua di arrivo.

Testo L1	Testo L2	Commento dello studente
Il suo corpo è stato trovato stamane nel cortile del palazzo dove abitava con la badante in via Principe Eugenio 32, rione Esquilino a Roma.	His body was found this morning in the courtyard of the building where he lived with the caretaker, at Via Principe Eugenio (Prince Eugene Street) 32, in the Esquiline district of Rome	Alternative translation: The first translation that came to my mind when I read “palazzo” was “palace”. I wanted to check if that could be right, so I used the dictionary “Reverso” and found that the translation could be “building” or “palace”. I then tried using the dictionary “Corriere” and found that “palazzo” could also be translated to “apartment block”. I then searched for “palazzo” in “Google images” and searched for the actual address in “Google Maps, street view” for further knowledge. By looking at the google images and the actual house Perlini lived in, I decided to go with “building”, as it seems “apartment block” in English is referring to much bigger buildings that are connected. Proper noun: I chose to use the English “Esquiline district”, for the Italian “Rione Esquilino” as it is a known district in the English language and also used in English travel guides of Rome (checked on Google and Google news).

Fig. 2: Esplorazioni del contesto extralinguistico e convenzioni discorsive

Nel terzo gruppo di estratti (Fig. 3) lo studente annota invece come un'esplorazione che tenga conto della frequenza e delle collocazioni rilevate in un corpus permetta di compiere scelte più oculate e sganciate dall'influenza della L1.

Testo L1	Testo L2	Commento dello studente
Il silenzio, l'isolamento di questi ultimi anni opachi, non fanno dimenticare quello che Memè Perlini è stato,	The silence and the isolation of these past few dull years, does not make one forget who Memé Perlini was.	False cognate: "Anni opachi", directly translated would be "opaque years". It is not a term used in English and it does not convey the same meaning. Instead I chose to find an equivalent.
... negli anni Sessanta-Settanta	... in the 60s and 70s	... Alternative translation: "Sixties and seventies" (fewer collocations)
... uno che ha imposto il suo stile in tv	... someone who has made his mark on TV	... Alternative translation: "Put his mark on" was also an option, but I checked in a corpus and found that it had much fewer collocations.

Fig. 3: Falsi amici, frequenze e collocazioni

Gli esempi riportati in figura 4 sono tratti da una traduzione che uno studente italiano ha svolto dall'italiano in inglese con il supporto di un corpus di articoli accademici appositamente collezionato (vedi par 2.). In questo caso i commenti evidenziano come l'utilizzo di un corpus permetta esplorazioni più precise e quantificabili, consentendo un livello di esplorazione e scelta traduttiva più sofisticato. Ad esempio nella scelta di *problems* invece di *disorders* perché quest'ultimo è un collocato di *psychiatric* piuttosto che di *psychological*, di *type*, oppure nella conferma di *in the case of* per *nel caso degli* anche se questa scelta inizialmente sembrava un falso amico.

Testo L1	Testo L2	Commento dello studente
Il bilinguismo sottrattivo invece è quello per cui l'acquisizione della seconda lingua è avvenuta, in maniera più o meno evidente, a scapito della lingua madre.	Subtractive bilingualism, on the other hand, is when second language acquisition has happened, more or less evidently, at the expense of the first language	first language: the literal translation "mother tongue" appears only 5 times in the corpus and has been discarded. The second candidate "native language" appears 58 times, but often in the expression "non-native language". "First language" was chosen because it yields 80 hits and is surely the most wide-spread in this context, especially contrasted with "second language" above.
Ovviamente il tipo di bilinguismo da perseguire e incoraggiare è quello additivo, poiché la perdita, anche se solo parziale, della lingua madre, può suscitare problemi di ordine sia psicologico che cognitivo.	Obviously the type of bilingualism to pursue and encourage is additive bilingualism, because the loss, if even partial, of the native language can lead to both cognitive and psychological problems.	type: the expression "type of bilingualism" is found 3 times in the corpus, while both "kind" and "form" are not attested. problems: the alternative translation "disorders" was considered, but it only appears 10 times in the corpus and often together with the adjective "psychiatric", so the original "problemi" was preserved instead.
Nel caso degli studenti immigrati, ad esempio, si può riscontrare un bilinguismo sottrattivo	In the case of immigrant students, for example, subtractive bilingualism can be observed	In the case of: it was double-checked in the corpus because it sounds like a literal translation, but it appears 31 times, often at the beginning of a sentence, so it was included.

Fig. 4: impiego di un disponibile corpus

L'esempio in Figura 5, al contrario, riguarda uno studente del primo anno di un corso triennale in Lettere e Filosofia con un livello più basso di inglese. I commenti sono in italiano e non fanno alcun riferimento a corpora tradizionali, ma denotano ugualmente una esplorazione del contesto linguistico ed extralinguistico che sembra aver sostenuto il processo di comprensione del testo di partenza (*metallic blanket*, *all-terrain vehicle*, *energy supplements*) e di traduzione (coperte isotermitiche, fuoristrada, integratori energetici) con l'impiego mirato di risorse online come Google, Google immagini, Wikipedia e siti web correlati al tema.

Testo L1	Testo L2	Commento dello studente
Covered in metallic blankets and other warm materials, a disheveled	Avvolto in coperte isotermitiche e materiali caldi,	-metallic blanket: coperta isotermica, telino isotermico, un Root

<p>and red-faced Root smiled at TV reporters after authorities slowly brought him down in an all-terrain vehicle from the snowy trail in the Sierra Nevada, about 200 miles east of San Francisco.</p>	<p>scarmigliato e con la faccia rossa ha sorriso agli inviati speciali dopo che le autorità lo hanno lentamente fatto scendere dal sentiero nevoso della Sierra Nevada, a circa 320 chilometri a est di San Francisco, su di un fuoristrada</p>	<p>metallina (google, wikipedia, google immagini) -All-terrain veichle: in acronimo ATV, sono dei mezzi di trasporto speciali. Vengono chiamati ATV anche in Italia, o tradotti veicoli tuttoterreno ma, sembra, solo in una terminologia tecnica. Più generalmente si possono tradurre con “fuoristrada” (google, dizionari, g.immagini, woodsandwheelsatvclub.com) -200 miglia = 321,8 chilometri</p>
--	---	---

<p>Root said he survived on energy supplements and a small amount of water he carried with him, and when he couldn't stop shaking, he compressed and released his muscles over and over to stay warm, she added.</p>	<p>“Root ha detto di essere riuscito a sopravvivere grazie ad integratori energetici e ad una piccola quantità d'acqua che aveva portato con sé, e quando non riusciva a smettere di tremare, comprimeva e rilasciava i muscoli più e più volte per riscaldarsi” ha aggiunto.</p>	<p>-Energy supplements: ci sono vari tipi di integratori. Dato che qui si parla di corridori si fa riferimento agli integratori energetici, spesso utilizzati da chi fa sport (google, google immagini, vari siti)</p>
--	---	--

Fig. 5: esplorazione di risorse online da parte di studenti con livello più basso di inglese

6. Conclusioni

Se l'uso di corpora linguistici viene visto da molti come un mezzo molto promettente nel favorire l'apprendimento delle lingue straniere, il loro impiego da parte di insegnanti e studenti è molto limitato se non praticamente inesistente. Leggermente migliore è la situazione in ambito universitario dove, grazie all'interesse di molti ricercatori, alcuni corsi hanno introdotto l'impiego di corpora per l'analisi e l'esplorazione linguistica diretta da parte degli studenti. Il più delle volte si tratta di corsi specialistici destinati a studenti di linguistica o traduzione, ma l'approccio presentato in questo studio intende avvalorare la tesi che sia possibile ed opportuno inserire percorsi di accesso ed esplorazione diretta ed autonoma anche per studenti meno esperti e con un livello di conoscenza della lingua straniera più basso. La traduzione e la scrittura nella lingua straniera sono stati usati per evidenziare ed affrontare problemi linguistici delimitati e definiti dal contesto, come la comprensione e traduzione di termini specifici, l'individuazione di *patterns*, collocazioni, fraseologie, falsi amici ed elementi extralinguistici. In una società in cui l'accesso alla rete è divenuto ormai un'abitudine ed in cui anche il mondo della formazione si interroga su come utilizzare internet, telefonini e risorse online per la didattica, sembra inevitabile coniugare la consultazione di corpora con l'accesso a data-base testuali e multimediali che possano facilitare l'esplorazione del linguaggio e dei contesti in cui esso si realizza. Agli apprendenti è stato proposto un metodo di lavoro che consiste nel partire non direttamente dalla linguistica dei corpora e da corpora linguistici, come avviene nella maggior parte delle esperienze riportate in letteratura, ma da strumenti più famigliari e noti, come i motori di ricerca, i siti di informazione enciclopedica, i dizionari e altre risorse disponibili online. In questo modo l'esplorazione linguistica e contestuale autonoma e diretta è molto più immediata, poiché coincide con quanto gli apprendenti fanno quotidianamente per altri motivi e non richiede una introduzione teorica o tecnica, ma pochi click su un computer o un cellulare. Un'esplorazione più sofisticata e guidata di corpora linguistici veri e propri sarà demandata ad una fase successiva e/o ad esigenze specifiche. Partendo da attività molto semplici, come il controllo ortografico, la correttezza di una preposizione, la definizione di un termine o di una locuzione, il significato di un'espressione idiomatica ecc., l'apprendente inizia un percorso di esplorazione che lo conduce, quando necessario, ad utilizzare strumenti più complessi, come la ricerca avanzata in Google o in corpora online. L'obiettivo principale di questo approccio è l'acquisizione di una dimestichezza, una metodologia ed un'autonomia nell'esplorazione di dati linguistici autentici che consenta agli apprendenti di colmare, almeno in parte, il loro *priming gap* e di ottenere risultati che siano più vicini a quelli di un parlante nativo. Una volta acquisita una metodologia adeguata e l'abitudine all'esplorazione saranno in grado, si suppone, di essere più indipendenti ed efficaci nell'uso delle lingue straniere e, se utilizzata in contesti meno familiari, anche della loro lingua madre. Il tempo 'perso' nel percorso didattico, insomma, dovrebbe essere ampiamente ripagato dai risultati.

Riferimenti bibliografici

- Aston G. (ed.), (2001), *Learning with corpora*, Athelstan, Houston, TX.
- Aston G., Bernardini S. & Stewart D. (eds), (2004), *Corpora and language learners*, John Benjamins, Amsterdam.
- Boulton A. (2010), *Empirical research in data-driven learning – a summary*, <https://nancy-universite.academia.edu/AlexBoulton>.
- Boulton A. (2015), *Applying data-driven learning to the web*, in Leńko-Szymańska A. & Boulton A. (eds), *Multiple affordances of language corpora for data-driven learning*, John Benjamins, Amsterdam: 267-295.
- Conroy M. A. (2010), *Internet tools for language learning: University students taking control of their writing*, in *Australasian Journal of Educational Technology* 26(6): 861-882.

- Chambers A. (2005), Integrating corpus consultation in language studies, in *Language Learning and Technology*, 9(2): 111-125.
- Chambers A. (2007), Popularising corpus consultation by language learners and teachers, in Hidalgo E., Quereda L. & Santana J. (eds), *Corpora in the foreign language classroom: Selected papers from the sixth International Conference on Teaching and Language Corpora (TaLC 6)*, Rodopi, Amsterdam: 3-16.
- Chinnery G. (2008), You've got some GALL: Google-assisted language learning, in *Language Learning & Technology* 12(1): 3-11.
- Cowie A. (1994), Phraseology, in Asher R. E. (eds), *The encyclopedia of language and linguistics*, Pergamon, Oxford: 3168-3171.
- Eu J. (2017), Patterns of Google use in language reference and learning: a user survey, in *Journal of Computers in Education* 4(4): 419-439.
- Frankenberg-Garcia A., Flowerdew L. & Aston G. (eds), (2011), *New trends in corpora and language learning*, Continuum, London and New York.
- Higgins J. & Johns T. (1984), *Computers in language learning*, Collins, London.
- Hoey M. (2005), *Lexical priming. A new theory of words and language*, Routledge, London.
- Hoey M. (2009), *Lexical priming*, <http://www.macmillandictionaries.com/MED-Magazine/January2009/52-LA-LexicalPriming-Print.htm>.
- Johns T. (1991), Should you be persuaded: Two samples of data-driven learning materials, in Johns T. & King P. (eds), *Classroom concordancing*, *ELR Journal* 4, University of Birmingham, Birmingham: 1-16.
- Leech G. (1997), Teaching and language corpora: A convergence, in Wichmann A., Fligelstone S., McEnery T. & Knowles G. (eds), *Teaching and language corpora*, Longman, London:1-23.
- McEnery T. & Xiao R. (1990), What corpora can offer in language teaching and learning, in Tribble, C. & Jones, G. (eds), *Concordances in the classroom: A resource book for teachers*, Longman, Harlow.
- Mukherjee J. (2004), Bridging the gap between applied corpus linguistics and the reality of English language teaching in Germany, in Connor U. & Upton T. (eds), *Applied corpus linguistics: a multidimensional perspective*, Rodopi, Amsterdam: 239-250.
- Meunier F. (2011), Corpus linguistics and second/foreign language learning: exploring multiple paths, in *Revista Brasileira de Linguística Aplicada* 11(2): 459-477.
- Partington A. (1998), *Patterns and meanings: Using corpora for English language research and teaching*, John Benjamins, Amsterdam.
- Römer U. (2008), Corpora and language teaching, in Lüdeling A & Merja K. (eds), *Corpus linguistics Vol. 1: an international handbook*, De Gruyter Mouton, Berlino: 112-130.
- Reppen R. (2010), *Using corpora in the language classroom*, Cambridge University Press, Cambridge.
- Sha G. (2010), Using Google as a super corpus to drive written language learning: A comparison with the British national corpus, in *Computer Assisted Language Learning* 23(5): 377-393.
- Sinclair J. M. (1991), *Corpus, concordance, collocation*, Oxford University Press, Oxford/New York.
- Tribble C. & Jones G. (1990), *Concordances in the classroom*, Longman, London.
- Tognini-Bonelli E. (2001), *Corpus linguistics at work*, *Studies in Corpus Linguistics* 6, John Benjamins, Amsterdam.
- Varantola K. (2003), *Translators and disposable corpora*, in Zanettin F., Bernardini S. & Stewart D. (eds), *Corpora in translator education*, St. Jerome, Manchester: 55-70.
- Zanca C. (2013), Online learning and data driven learning in translation training and language teaching, in Argondizzo C. (eds), *Creativity and innovation in language education*, Peter Lang, Bern ; New York: 317-334.
- Zanca C. (2014), Developing translation strategies and cultural awareness using corpora and the web, <http://lodel.irevues.inist.fr/tralogy/index.php?id=116&format=print>
- Zanca C. (2017), When the going gets tough... the tough get googling, in Damascelli A.T. (ed.), *Digital resources, creativity and innovative methodologies in language teaching and learning*, Cambridge Scholars Publishing, Cambridge.
- Zanettin F. (2009), Corpus-based Translation Activities for Language Learners, *The Interpreter and Translator Trainer* 3(2): 209-224.

Corpora e risorse online

- [BNC] (2007) British National Corpus, Oxford University Computing Services, Oxford. www.natcorp.ox.ac.uk
- [COCA] Davies, Mark. (2008-) The Corpus of Contemporary American English: 560 million words, 1990-present. Available online at <https://corpus.byu.edu/coca/>
- [Context Reverso] <http://context.reverso.net>
- [GlowBe] Davies, Mark. (2013) Corpus of Global Web-Based English: 1.9 billion words from speakers in 20 countries (GloWbE). Available online at <https://corpus.byu.edu/glowbe/>
- [Google Immagini] <https://www.google.it/imghp>
- [Google News] <https://news.google.com/>
- [Google Scholar] <https://scholar.google.com/>
- [Google Books] <https://books.google.com>
- [Google Ngrams] <https://books.google.com/ngrams>

[Google Translate] <https://translate.google.com/m/translate>

[Linguee] <https://www.linguee.it/>

[Visual Thesaurus] <https://www.visualthesaurus.com>

Note

[1]. Negli ultimi anni gli studenti interessati erano iscritti al primo anno di una laurea magistrale ed al terzo anno di una triennale.

[2]. In questo caso gli studenti erano iscritti al primo anno di lauree triennali in Lettere e a Scienze della Comunicazione.

[3]. Reperibile all'indirizzo: <http://www.laurenceanthony.net/software/antconc/>

[4]. I siti di riferimento per queste ed altre risorse citate nell'articolo e disponibili online sono elencati dopo la bibliografia.

[5]. <https://www.thetimes.co.uk/article/the-gentleman-of-verona-and-his-incredible-sausage-9lxsnb2mwr5>

[6]. <https://www.thenational.ae/i-ll-happily-take-the-credit-for-a-dog-day-afternoon-1.241348>

[7]. <http://news.bbc.co.uk/2/hi/health/2493631.stm>

[8]. <http://www.bbc.com/news/science-environment-42428528>

[9]. <https://forum.wordreference.com/threads/hai-voluto-la-bicicletta-e-allora-pedala.714846/?hl=it>

[10]. <https://corpus.byu.edu/>

[11]. <http://www.mirror.co.uk/sport/football/match-reports/tottenham-4-3-leicester-harry-kane-5374282>

©inTRAlinea & Cesare Zanca (2018).

"Corpora, Google e roba simile Per quale ragione gli studenti di una lingua straniera dovrebbero perderci tempo?", *inTRAlinea* Special Issue: Translation And Interpreting for Language Learners (TAIL).

Stable URL: <http://www.intraline.org/archive/article/2296>