

CAPITOLO 5

DIDATTICA DELL'INTERPRETAZIONE BASATA SU CORPORA

Claudio Bendazzoli

Università degli Studi di Torino

1. Nascita e sviluppo di un nuovo paradigma di ricerca: *i Corpus-based Interpreting Studies*

Per poter studiare e analizzare la comunicazione mediata da interpreti si rende necessario catturare la natura effimera del parlato (o delle lingue dei segni) in un formato stabile e archiviabile, che consenta di accedere ai dati (il testo di partenza e il testo di arrivo, oltre a tutta una serie di informazioni sul contesto di appartenenza) ogni qual volta si renda necessario e di osservarli oltre l'immediatezza in cui si sono manifestati originariamente. In generale, i metodi impiegati a tal fine sono sempre stati la registrazione (audio o anche video) e la trascrizione.

Prima dell'avvento delle tecnologie digitali e di internet, gli strumenti a disposizione degli studiosi per registrare e trascrivere dati di questo tipo presentavano importanti limiti strutturali (basti pensare alla durata del nastro di una musicassetta, alla capacità di archiviazione dei dischi portatili e fissi nei computer o alla necessità di stampare su carta le trascrizioni per analizzarle). Tuttavia, grazie ai continui progressi della tecnologia, in seguito si sono aperte nuove possibilità di analisi su scala sempre maggiore, con la creazione di banche dati le cui dimensioni potevano essere gestite efficacemente solo con mezzi informatici (non più manualmente). I corpora linguistici nascono proprio da grandi raccolte di testi in formato elettronico, organizzate in base a criteri di selezione predeterminati così da assicurare un certo grado di rappresentatività (Bowker, Pearson 2002).

In ambito traduttologico, la linguistica dei corpora trova applicazione prima negli studi sulla traduzione (Baker 1993), probabilmente per la maggiore disponibilità di testi nel formato richiesto e di programmi in grado di elaborare e analizzare il linguaggio scritto. A partire dalla fine degli anni Novanta, tale approccio richiama l'interesse anche di chi si occupa di studi sull'interpretazione e appare fin da subito come una preziosa opportunità di superare la natura spesso aneddotica di alcune ricerche che avevano esaminato la comunicazione mediata da interpreti agli albori degli studi sull'interpretazione (Shlesinger 1998). Con questo nuovo paradigma di ricerca si auspica infatti che si possano creare corpora di interpretazione, cioè grandi banche dati di trascrizioni di testi di partenza e testi di arrivo, così da poter individuare, estrarre e contare le occorrenze di vari tipi di fenomeni. Nonostante le numerose sfide insite nel processo di trascrizione e nel disporre di strumenti informatici compatibili con diverse lingue, dopo oltre due decenni dai primi appelli alla comunità scientifica, i corpora di interpretazione sono oggi una realtà.

Vi sono alcune caratteristiche particolari che distinguono i corpora di interpretazione e che vale la pena puntualizzare (Bendazzoli, Sandrelli 2009). In primo luogo, alcuni corpora di interpretazione non sono elettronici, o per lo meno non lo sono al punto da poter essere analizzati dal computer. In questi casi, il termine *corpus* viene utilizzato per campioni di dati che vengono analizzati ancora manualmente, senza l'ausilio di strumenti di linguistica computazionale, e sono di dimensioni limitate. In secondo luogo, non tutti i corpora di interpretazione sono accessibili pubblicamente – ad esempio, su piattaforme in rete. Infine, i primi studi condotti con l'approccio basato sul corpus riguardavano solo l'interpretazione di conferenza, in particolare nella modalità simultanea. Ciononostante, dalle rassegne presenti in letteratura (Bendazzoli 2018; Setton 2011) emerge chiaramente come i corpora di interpretazione siano aumentati, anno dopo anno, non solo in termini numerici ma anche in termini di dimensione, nonché di modalità (anche consecutiva) e contesti rappresentati.

Ad oggi, i corpora di interpretazione sono stati impiegati soprattutto per fini di ricerca (tanto sul processo quanto sul prodotto dell'interpretazione), ma hanno trovato impiego anche in ambito didattico (Bendazzoli, Russo, Defrancq 2018; Russo, Bendazzoli, Defrancq 2018; Straniero Sergio, Falbo 2012).

La seguente sezione offre una breve disamina della metodologia alla base della creazione di un corpus, dopodiché si presentano diverse esperienze formative e risorse didattiche attinenti a questo innovativo paradigma oramai ben consolidato in ambito traduttologico.

2. Come creare un corpus di interpretazione

La metodologia sottostante la creazione di un corpus di interpretazione attinge dalle procedure previste per i corpora orali (Thompson 2004) e comprende diverse fasi, quali la progettazione del corpus, la raccolta dei dati, la trascrizione, l'annotazione, l'allineamento, l'accesso ai dati e la distribuzione (Bendazzoli 2010; Bernardini *et al.* 2018). Mentre le prime tre fasi sono essenziali (poiché non può esistere un corpus di interpretazione senza dati trascritti), le rimanenti sono opzionali, benché contribuiscano sostanzialmente ad aumentare il potenziale di ricerca di un corpus.

2.1. Progettazione del corpus

La prima fase è determinata dal tipo di corpus che si desidera costruire, nonché dalle modalità di interpretazione e dalle lingue oggetto di indagine. Tra i principali tipi di corpora si possono annoverare i corpora paralleli (con testi di partenza e i relativi testi di arrivo) e i corpora comparabili (con testi di partenza in una determinata lingua e testi di arrivo nella medesima lingua, così da poter confrontare le caratteristiche della stessa lingua, ma prodotta in modalità differenti – ad esempio, da un eloquio spontaneo, preparato o letto per i testi di partenza, anche differenziando se si tratta di parlanti nativi o non nativi; da una interpretazione simultanea o consecutiva per i testi di arrivo). I corpora intermodali raccolgono invece più testi di arrivo prodotti sulla base del medesimo testo di partenza a partire da diverse modalità di interpretazione o di traduzione, potendo così esplorare i tratti salienti del linguaggio interpretato in simultanea rispetto al linguaggio interpretato in consecutiva o tradotto (a partire da testo scritto). In base al numero di lingue rappresentate, un corpus può dirsi monolingue, bilingue o multilingue.

Si deve inoltre prestare attenzione al tipo di modalità di interpretazione che sarà rappresentata nel corpus – se simultanea (con o senza cabina, interpretazione sussurrata, traduzione a vista), consecutiva (classica, con presa di note, o breve), o entrambe. Nell'interpretazione consecutiva, l'alternanza di brani tra testo di partenza e testo di arrivo può comportare scelte diverse a seconda del grado di interazione e delle possibili sovrapposizioni (si pensi alla consecutiva breve in scambi dialogici rispetto alla consecutiva classica in una presentazione-monologo). Le lingue di lavoro coinvolte contribuiscono parimenti alla complessità strutturale di un corpus e al grado di rappresentatività auspicato (Halverson 1998) anche in base al contesto di riferimento (ad esempio, le sessioni plenarie del Par-

lamento europeo, i convegni internazionali, le udienze di tribunale, ecc.). Tutti questi fattori influenzano in diversa misura l'impostazione che sarà necessario dare all'architettura del corpus, nonché il grado di rappresentatività degli eventi linguistici effettivamente inclusi.

Si prenda ad esempio uno dei primi corpora elettronici sviluppati in seno al Dipartimento DIT dell'Università di Bologna, ovvero EPIC (European Parliament Interpreting Corpus). EPIC è un corpus trilingue (italiano, inglese, spagnolo) di testi di partenza e testi di arrivo (in simultanea) nell'ambito delle sedute plenarie del Parlamento europeo (Russo *et al.* 2012). In realtà, considerando il tipo di testi e il numero di lingue in esso rappresentati, EPIC si compone di nove corpora, tre per ciascuna lingua dei testi di partenza e sei come risultato dell'interpretazione simultanea in tutte le combinazioni possibili tra le tre lingue coinvolte (italiano>inglese, italiano>spagnolo; inglese>italiano, inglese>spagnolo, spagnolo>italiano, spagnolo>inglese). La particolare struttura di EPIC consente di esplorarlo sia da una prospettiva parallela (cioè confrontando i testi di partenza in una lingua con relativi testi di arrivo nelle altre due lingue), sia da una prospettiva comparabile (considerando i testi di partenza in una lingua e i testi di arrivo nella medesima lingua, prodotti dai testi di partenza presentati nelle altre due lingue coinvolte, per esempio testi di partenza in italiano e testi di arrivo sempre in italiano a partire dai discorsi originariamente pronunciati in inglese e in spagnolo). Un ulteriore sviluppo di questo corpus ha portato alla creazione di EPTIC (European Parliament Interpreting and Translation Corpus), un corpus intermodale multilingue con testi di partenza e testi di arrivo non solo orali e frutto di interpretazione simultanea, ma anche scritti (i verbali delle sedute plenarie) e tradotti (le traduzioni ufficiali dei verbali) (Ferraresi, Bernardini 2019). Più aumentano le lingue rappresentate e le modalità di traduzione, più complessa diventa la struttura del corpus.

2.2. Raccolta dei dati

La seconda fase prevede la raccolta dei dati che si desidera inserire nel corpus. I corpora di interpretazione normalmente si basano su trascrizioni di dati registrati (in lingua parlata e/o in lingua dei segni). I progressi tecnologici hanno reso più facile la registrazione video di situazioni comunicative, anche se diversi progetti si sono basati solo su registrazioni audio. In entrambi i casi, la raccolta dei dati può essere realizzata sul campo (Bendazzoli 2016), o indirettamente, con mezzi quali il download di dati da internet (ad esempio, dalla videoteca del Parlamento

europeo) o la registrazione di eventi mediati da interpreti trasmessi in televisione o in streaming.

La raccolta di dati sul campo comporta diverse sfide (logistiche e tecniche), a seconda delle modalità di interpretazione e dei contesti considerati. Per esempio, nell'interpretazione consecutiva sarebbe interessante tenere traccia delle note dell'interprete, insieme alla produzione dei discorsi di partenza e di arrivo, attraverso l'uso della *digital pen technology* (Kellet Bidoli 2016). Nell'interpretazione simultanea senza cabina e nell'interpretazione sussurrata, occorre prestare particolare attenzione all'attrezzatura di registrazione per ottenere registrazioni di alta qualità dei discorsi di partenza e di arrivo (ad esempio, mediante microfoni *Lavalier*). Nell'interpretazione simultanea con cabina insonorizzata, è talvolta necessario coordinare la raccolta dei dati da postazioni situate in luoghi diversi, poiché la cabina può non essere installata nella stessa sala in cui si svolgono i lavori della conferenza (per non parlare dell'interpretazione da remoto). Inoltre, nell'interpretazione simultanea ci sono (almeno) due tracce audio da gestire allo stesso tempo, e queste devono essere tenute separate per poterle trascrivere comodamente.

Oltre alla registrazione dei testi di partenza e dei testi di arrivo, la raccolta dei dati si completa tenendo traccia delle informazioni sul contesto in cui avviene la comunicazione mediata da interpreti. In particolare, i metadati sui partecipanti (compresi gli interpreti), sulla situazione comunicativa (struttura e sezioni principali, programma, configurazione dell'ambiente, ecc.) e sui testi di partenza (stile di presentazione, uso di supporti audiovisivi, diapositive, ecc.) sono informazioni di primaria importanza per poter meglio interpretare quanto rappresentato dal corpus e, ancor prima, decidere se includere o meno un determinato testo rispetto ai criteri di rappresentatività stabiliti nella fase di progettazione.

Quale che sia il metodo di raccolta dei dati, diretto o indiretto, è inoltre necessario verificare il grado di riservatezza della situazione comunicativa oggetto di studio in modo da approntare eventuali procedure di consenso informato. Lo stesso tipo di evento comunicativo (ad esempio, una conferenza, una trattativa, un colloquio) può spaziare dall'essere di natura pubblica o totalmente riservata, con conseguenti restrizioni sull'accesso ai dati e sulla loro divulgazione al di fuori del progetto di ricerca per cui sono raccolti.

2.3. Trascrizione dei dati, annotazione e indicizzazione

La trasposizione dei dati audio/visivi in forma trascritta può essere gestita con svariati programmi informatici (Niemants 2012) e realizzata seguendo diverse

convenzioni, a seconda del tipo di analisi e del formato dell'interazione (monologica o dialogica). La trascrizione è il risultato di un processo selettivo in cui solo una (minima) parte delle caratteristiche dell'oralità viene rappresentata in forma scritta. D'altronde si tratta di un compromesso inevitabile se si vuole mantenere un certo grado di leggibilità delle trascrizioni, anche tenendo conto della possibile elaborazione informatica delle stesse.

In aggiunta al livello verbale in senso stretto (cioè la mera trascrizione delle parole dette o segnate), le trascrizioni possono essere arricchite dall'annotazione di alcuni attributi, sempre a seconda dell'analisi da effettuare. Generalmente, l'annotazione dei metadati è effettuata per tenere traccia del tipo di testo (numero totale di parole, velocità di eloquio, durata, in abbinamento a dettagli relativi al parlante e alla situazione comunicativa) e viene inserita all'inizio della trascrizione (*header*). Inoltre, ulteriori livelli (opzionali) di annotazione vanno dal *part-of-speech* (POS) *tagging* – cioè l'annotazione di categorie grammaticali per ogni parola (o *token*) – a specifiche caratteristiche verbali o non verbali (ad esempio le pause, eventuale pronuncia non standard, ecc.). Le annotazioni possono essere applicate automaticamente (come nel caso del *POS-tagging*) o manualmente.

I testi trascritti e annotati devono infine essere indicizzati per poter svolgere ricerche sistematiche e recuperare in maniera automatica le informazioni che sono state annotate. Tra i sistemi e gli strumenti usati più comunemente al fine di gestire corpus di interpretazione si annoverano il *Corpus Workbench* (Evert, Hardie 2011), EXMARaLDA (Bührig *et al.* 2012; Schmidt, Wörner 2014), *Sketch Engine* (Kilgarriff *et al.* 2014), nonché la sua versione *open source* conosciuta come *NoSketch Engine* (<https://nlp.fi.muni.cz/trac/noske>), e ELAN (<https://archive.mpi.nl/tla/elani>) particolarmente diffuso nei corpora di lingua dei segni.

2.4. Allineamento

Un ulteriore grado di elaborazione delle trascrizioni riguarda l'allineamento. Nei corpora di interpretazione, l'allineamento può essere impostato su due livelli. Si possono allineare le trascrizioni del testo di partenza con le trascrizioni del testo di arrivo per un confronto immediato. A seconda della modalità di interpretazione e della progettazione del corpus (i corpora intermodali comportano una struttura più complessa anche in termini di allineamento, Bernardini *et al.* 2018), le trascrizioni allineate possono essere visualizzate secondo un formato tabulare o seguendo una rappresentazione a spartito musicale (quest'ultimo sarebbe ideale

per poter rappresentare visivamente il *décalage* nelle modalità simultanee o le sovrapposizioni tra più parlanti nella comunicazione dialogica). Si può procedere a un allineamento manuale oppure semi-automatico.

L'altro livello di allineamento può essere realizzato tra una trascrizione e la sua registrazione audio/video (ad esempio, annotando codici temporali nella trascrizione che la colleghino alla traccia audio/video), nonché eventuali materiali di supporto (come le diapositive della presentazione, il testo del discorso, ecc.).

2.5. Accessibilità e distribuzione

L'ultimo passo nello sviluppo di un corpus riguarda la sua accessibilità e distribuzione. Alcuni corpora sono resi disponibili alla comunità scientifica senza particolari restrizioni, mentre altri sono accessibili solo ai ricercatori che li hanno sviluppati (ad esempio, per questioni di riservatezza). La fruizione aperta può avvenire tramite piattaforme online, come nel caso delle risorse nella piattaforma CoLiTec (<https://corpora.dipintra.it/>), oppure tramite distribuzione da parte dei creatori del corpus (Angermeyer, Meyer, Schmidt 2012) o di associazioni preposte (come la *European Language Resources Association* nel caso di EPIC).

3. Applicazioni didattiche: apprendere da e con i corpora

I corpora di interpretazione (e non solo) hanno avuto una graduale ricaduta in ambito formativo fin dagli albori del loro sviluppo. Oltre all'uso mirato di vari tipi di corpora nella didattica dell'interpretazione, la metodologia stessa alla base della creazione di un corpus può essere foriera di importanti riflessioni sui testi e sui contesti dell'interpretazione tra i discenti. Per quanto la creazione di un corpus rimanga un'impresa piuttosto impegnativa, la maggiore facilità di accesso ai dati in alcuni contesti, nonché la crescente disponibilità di corpora già esistenti hanno aperto la strada a esperienze collaborative di ricerca ispirate all'approccio basato sui corpora (Bendazzoli, Bertozzi, Russo 2021). Tra i fattori che hanno favorito questo processo, vanno senz'altro menzionate le tecnologie digitali che hanno reso più semplice e meno invasivo il processo di acquisizione e gestione dei dati. Oltre a questo, gli studenti hanno l'opportunità di sfruttare corpora esistenti per fare ricerche ai fini della loro tesi di laurea (Dal Fovo 2011; Russo 2010) e di dottorato (Bertozzi 2018).

La riflessione critica e metodologica rappresenta solo una parte del contributo pedagogico dei corpora nella formazione di interpreti. Sono state documentate varie iniziative a livello didattico che hanno trovato la loro propulsione a partire da raccolte ragionate di testi di partenza (e di arrivo). In alcuni casi, non si tratta di corpora di interpretazione *stricto sensu*, per lo meno stando alla definizione summenzionata, bensì di archivi contenenti materiali strutturati a fini pedagogici. Potremmo considerare queste raccolte come proto-corpora, in quanto riprendono da vicino le prime fasi metodologiche della creazione di un corpus, applicando criteri di rappresentatività e annotazioni che ne consentono una selezione mirata, pur senza fornire trascrizioni annotate e indicizzate. In altre esperienze invece viene fatto uso di corpora, anche non strettamente legati all'interpretazione, ma in grado di fornire materiali utili per attività terminologiche e di allenamento mirato. Allo stesso modo, vi sono attività didattiche sviluppate a partire da corpora di interpretazione veri e propri, sia nell'aula di interpretazione sia in moduli a latere per la preparazione linguistica degli interpreti. Infine, vi sono proposte di impiego di strumenti di linguistica computazionale che potrebbero essere di grande utilità in fase di preparazione per un determinato incarico, soprattutto per quanto concerne la terminologia settoriale.

3.1. Proto-corpora

Le istituzioni dell'Unione europea, in particolare il Parlamento europeo, sono la principale fonte di dati per le applicazioni pedagogiche legate ai corpora a partire da archivi multimediali. Uno dei primi esempi di (proto)corpora pedagogici per la formazione all'interpretazione consecutiva è dato dalla raccolta messa a punto da Dollerup e Ceelen (1996). Il loro corpus di interpretazione consecutiva in danese, olandese, inglese, francese, tedesco e italiano è in effetti una raccolta di registrazioni audio e trascrizioni, e non un corpus elettronico vero e proprio, di lezioni con testi di partenza e testi di arrivo prodotti da tirocinanti in un corso di formazione tenuto nel 1976. Uno sviluppo digitale più avanzato di questo tipo è la banca dati *Marius*, una raccolta di discorsi del Parlamento europeo e del Forum Sociale Mondiale creata da de Manuel (2003a, 2003b) e arricchita con annotazioni di metadati. Le annotazioni rendono possibile la ricerca mirata delle trascrizioni e del relativo materiale multimediale, sebbene le trascrizioni non siano perlustrabili come in un corpus elettronico standard.

Questi primi esempi assomigliano ad altre risorse ben note in ambito formativo, quali i cosiddetti *speech repositories* (ad esempio, quello della Commissione

europea o la piattaforma collaborativa *Speechpool*). Si tratta di archivi di discorsi annotati che possono essere utilizzati per allenarsi in simultanea e consecutiva. Le loro annotazioni coprono una serie di caratteristiche, dalla durata, la velocità di eloquio e l'argomento al tipo di modalità di interpretazione e il livello di difficoltà coinvolti (quest'ultimo anche determinato collaborativamente dagli utenti).

3.2. Corpora

Sempre sfruttando risorse provenienti dalle istituzioni dell'Unione europea, alcuni corpora di interpretazione hanno portato allo sviluppo di attività pedagogiche. Per l'interpretazione simultanea tra italiano e inglese, Sandrelli (2010) ha sviluppato un'attività formativa basata sull'estrazione di occorrenze dei tempi verbali in EPIC al fine di verificare come questi vengono utilizzati dagli interpreti in servizio presso il Parlamento europeo. In particolare, l'attenzione è stata posta sulla gestione di tempi verbali che si differenziano tra le due lingue (passato remoto e passato prossimo), fornendo esempi concreti di professionisti in situazioni reali.

Sulla stessa linea, Aston (2018) ha proposto attività di apprendimento autonomo a partire dall'analisi di concordanze ottenute da un corpus di trascrizioni di testi di arrivo in inglese, prodotti da interpreti del Parlamento europeo, con particolare riferimento alle unità fraseologiche formali. Spinolo (2018) ha organizzato sessioni di formazione sull'interpretazione del linguaggio figurato utilizzando il corpus IMITES (spagnolo-italiano), con approfondimenti sulle strategie adottate da interpreti professionisti della Commissione europea nel gestire le metafore in interpretazione simultanea. La gestione degli anglicismi nell'interpretazione e nella traduzione dall'italiano allo spagnolo, unitamente a schede terminologiche dettagliate, sono invece il focus della piattaforma *Anglintrad* creata da Bertozzi (2018) a partire da dati tratti da EPIC. Cresswell (2018) ha combinato diversi corpora di interpretazione simultanea (nelle sedute del Parlamento europeo e in convegni di tema medico) per studiare la frequenza dei verbi frasali nei testi di arrivo in inglese. Da questa analisi ha riscontrato una frequenza più alta nell'output di interpreti madrelingua inglese rispetto a interpreti che lavorano in inglese come lingua B (lingua straniera attiva). Tale risultato, abbinato all'estrazione del linguaggio formulaico dai corpora presi in considerazione, ha portato allo sviluppo di materiale didattico mirato per gli interpreti di conferenza, in particolare in moduli di rafforzamento linguistico. Infine, l'analisi delle collocazioni in una pro-

spettiva intermodale – ovvero, riscontrate nei testi di arrivo tradotti e interpretati a partire dallo stesso testo di partenza in EPTIC – ha portato anche a proporre attività nell’aula di traduzione: i dati dell’interpretazione di conferenza vengono confrontati con quelli della traduzione e stimolano la riflessione sulle differenze tra i due processi (Ferraresi 2016).

Oltre alle modalità simultanee, vi sono esempi interessanti di uso dei corpora anche in contesti formativi per l’interpretazione dialogica. Niemants (2013) ha realizzato un corpus di interazioni dialogiche mediate tra medico e paziente, sia reali sia da simulazioni. Tali materiali sono stati impiegati per analizzare diverse questioni cruciali in questo tipo di interpretazione (quali il ruolo dell’interprete, la gestione dei turni di parola, ecc.), coinvolgendo i discenti in una riflessione critica e sfruttando i dati per svolgere ulteriori simulazioni. A partire da un corpus di mediazioni reali in ambito giuridico, Orozco (2018) ha invece rilevato in maniera sistematica le problematicità nelle prassi messe in campo dagli interpreti (professionisti e non) e ha potuto delineare importanti necessità formative in questo ambito.

Oltre ai corpora di interpretazione veri e propri, anche i corpora non rappresentativi del linguaggio mediato sono stati utilizzati a fini pedagogici nella formazione degli interpreti. Per esempio, Bale (2013) ha sfruttato il corpus BACKBONE (Kohn 2012) per migliorare la conoscenza lessicale degli studenti come parte delle competenze linguistiche impiegate nell’interpretazione consecutiva (inglese-tedesco). BACKBONE, un corpus di interviste in diverse lingue, non è direttamente correlato alla traduzione o all’interpretazione. L’utilità di questo corpus nella formazione degli interpreti deriva dalle sue annotazioni di stampo pedagogico, che permettono di selezionare trascrizioni e video in base alla lingua, all’argomento e a particolari caratteristiche linguistiche. Gli interpreti in formazione possono quindi utilizzare questi materiali come fonte di discorsi per la pratica dell’interpretazione, nonché come testi paralleli per attività di ricerca e verifica terminologica. Una simile applicazione è stata proposta da Leeson (2008), per la formazione degli interpreti di lingua dei segni, con il corpus *Signs of Ireland*. Questo corpus può essere usato non solo come fonte di discorsi di partenza, ma anche come corpus di riferimento per verificare e apprendere collocazioni, con la possibilità di auto-analisi della propria resa interpretativa. Gli ultimi due esempi illustrano chiaramente il potenziale formativo dei corpora, anche laddove non rappresentano la comunicazione mediata, ma sono stati integrati da annotazioni di carattere pedagogico.

3.3. Terminologia

Alcuni studiosi si rifanno a strumenti di linguistica computazionale per approfondire la competenza terminologica e le abilità linguistiche (ad esempio, con ricerche di collocazioni e unità fraseologiche), a partire da corpora creati dal web o dai materiali messi a disposizione prima di un incarico di interpretazione.

La metodologia proposta da Baroni e Ueyama (2006) si basa sui cosiddetti *web-crawled* corpora, ovvero corpora i cui testi sono stati scaricati automaticamente da internet, ma in maniera mirata. La selezione avviene tramite l'impostazione di alcuni filtri (*seeds*) che riconoscono la rilevanza terminologica dei testi in rete e ne permettono una selezione tale da riuscire a ottenere un corpus specialistico in tempi rapidi. Una volta costituito, il corpus può essere esplorato attraverso ricerche sulle parole chiave, sul lessico ad alta o bassa frequenza, sulle collocazioni, con anche la possibilità di utilizzarlo in chiave comparabile e per lo sviluppo di banche dati terminologiche (Fantinuoli 2018). Un altro esempio è dato da un programma sviluppato appositamente a tal fine e noto come *InterpretBank* (Fantinuoli 2016). Questa applicazione software è corredata da strumenti per estrarre automaticamente testi e termini, gestire glossari e accedere alla terminologia in tempo reale. Risulta particolarmente interessante la possibilità di utilizzarla con i materiali forniti in preparazione per un incarico di interpretazione con scarso anticipo, specialmente nei casi in cui i materiali siano tanto estesi da non poter essere letti interamente per mancanza di tempo.

4. Considerazioni finali

Per poter allenare le abilità necessarie alla pratica dell'interpretazione professionale si rende necessario familiarizzarsi con diverse situazioni comunicative, in cui gli interpreti si trovano a gestire una vasta gamma di eventi linguistici. I corpora sono risorse linguistiche che ben rispondono a questa necessità, in quanto mettono a disposizione testi di partenza (e di arrivo) che sono stati campionati in base a criteri di rappresentatività. La disponibilità di materiali didattici è, tuttavia, solo una parte del potenziale formativo di queste risorse. Anche i risultati degli studi effettuati con l'approccio basato sui corpora contribuiscono a orientare l'offerta formativa (basti pensare ai pochi esempi citati in questo capitolo), alla luce delle prassi professionali osservate in altri interpreti professionisti. Allo stesso modo, la stessa metodologia utilizzata nella realizzazione di un corpus può offrire impor-

tanti spunti di riflessione sul processo e sul prodotto dell'interpretazione, nonché l'opportunità di svolgere approfondimenti terminologici per una preparazione mirata.

Bibliografia

- Angermeyer, P. S., B. Meyer, T. Schmidt 2012, *Sharing community interpreting corpora: A pilot study*, in T. Schmidt, K. Wörner (eds.), *Multilingual Corpora and Multilingual Corpus Analysis*, Amsterdam/Philadelphia: John Benjamins, pp. 275-294.
- Aston, G. 2018, *Acquiring the language of interpreters: A corpus-based approach*, in M. Russo, C. Bendazzoli, B. Defrancq (eds.), *Making way in Corpus-based Interpreting Studies*, Singapore: Springer, pp. 83-96.
- Baker, M. 1993, *Corpus linguistics and translation studies: Implications and applications*, in M. Baker, G. Francis, E. Tognini-Bonelli (eds.), *Text and Technology: In Honour of John Sinclair*, Amsterdam/Philadelphia: John Benjamins, pp. 233-250.
- Bale, R. 2013, *Undergraduate consecutive interpreting and lexical knowledge: The role of spoken corpora*, «The Interpreter and Translator Trainer» (VII-I), pp. 27-50.
- Baroni, M., M. Ueyama 2006, *Building general- and special-purpose corpora by Web crawling*, in *Proceedings of the 13th Nijl International Symposium. Language Corpora: Their Compilation and Application*, pp. 31-40.
- Bendazzoli, C. 2010, *Corpora e interpretazione simultanea*, Bologna: Asterisco.
- Bendazzoli, C. 2016, *The ethnography of interpreter-mediated communication: Methodological challenges in fieldwork*, in C. Bendazzoli, C. Monacelli (eds.), *Addressing Methodological Challenges in Interpreting Studies Research*, Newcastle upon Tyne: Cambridge Scholars Publishing, pp. 3-30.
- Bendazzoli, C. 2018, *Corpus-based Interpreting Studies: Past, present and future prospects of a (wired) cottage industry*, in M. Russo, C. Bendazzoli, B. Defrancq (eds.), *Making way in Corpus-based Interpreting Studies*, Singapore: Springer, pp. 1-19.
- Bendazzoli, C., M. Bertozzi, M. Russo 2021, *Du texte aux ressources multimodales : faire avancer la recherche en interprétation à partir d'un corpus déjà existant*, «Meta» (LXV-I), pp. 211-236.
- Bendazzoli, C., M. Russo, B. Defrancq 2018 (eds.), *New Findings in Corpus-based Interpreting Studies*. Special issue of «inTRAlinea», <http://www.intraline.org/specials/cbis>.
- Bendazzoli, C., A. Sandrelli 2009, *Corpus-based Interpreting Studies: Early work and future prospects*, «Tradumatica» (VII). *L'aplicació dels corpus linguistics a la traducció*, <http://webs2002.uab.es/tradumatica/revista/num7/articles/08/08art.htm>.
- Bernardini, S. et al. 2018, *Building interpreting and intermodal corpora: A how-to for a formidable task*, in M. Russo, C. Bendazzoli, B. Defrancq (eds.), *Making Way in Corpus-based Interpreting Studies*, Singapore: Springer, pp. 21-42.

- Bertozzi, M. 2018, *ANGLINTRAD: Towards a purpose specific interpreting corpus*, in C. Bendazzoli, M. Russo, B. Defrancq (eds.), *New Findings in Corpus-based Interpreting Studies*. Special issue of «inTRAlinea», <http://www.intralinea.org/specials/article/2317>.
- Bowker, L., J. Pearson 2002, *Working with Specialized Language. A Practical Guide to Using Corpora*, London/New York: Routledge.
- Bühlig, K. et al. 2012, *The corpus "Interpreting in Hospitals": Possible applications for research and communication training*, in T. Schmidt, K. Wörner (eds.), *Multilingual Corpora and Multilingual Corpus Analysis*, Amsterdam/Philadelphia: John Benjamins, pp. 305-315.
- Cresswell, A. 2018, *Looking up phrasal verbs in small corpora of interpreting: An attempt to draw out aspects of interpreted language*, in C. Bendazzoli, M. Russo, B. Defrancq (eds.), *New Findings in Corpus-based Interpreting Studies*. Special issue of «inTRAlinea», <http://www.intralinea.org/specials/article/2319>.
- Dal Fovo, E. 2011, *Through the CorIT looking-glass and what MA students found there*, «The Interpreters' Newsletter» (XVI), pp. 1-20, <http://hdl.handle.net/10077/8266>.
- de Manuel Jerez, J. 2003a (ed.), *Nuevas tecnologías y formación de intérpretes*, Granada: Comares.
- de Manuel Jerez, J. 2003b, *Medios audiovisuales e informáticos en la formación de intérpretes*, in M. A. Vega Cernuda (ed.), *Una mirada al taller de San Jerónimo: bibliografías, técnicas y reflexiones en torno a la traducción*, Madrid: Universidad Complutense, pp. 53-62.
- Dollerup, C., L. Ceelen 1996, *A Corpus of Consecutive Interpreting, Comprising Danish, Dutch, English, French and Italian*, Copenhagen: University of Copenhagen.
- Evert, S., A. Hardie 2011, *Twenty-first century Corpus Workbench: Updating a query architecture for the new millennium*, in *Proceedings of the Corpus Linguistics 2011 conference, University of Birmingham, UK*, <https://www.birmingham.ac.uk/documents/college-artslaw/corpus/conference-archives/2011/Paper-153.pdf>.
- Fantinuoli, C. 2016, *InterpretBank. Redefining computer-assisted interpreting tools*, in *Proceedings of the 38th Conference Translating and the Computer, London, UK, November 17-18*, pp. 42-54.
- Fantinuoli, C. 2018, *The use of comparable corpora in interpreting practice and training*, «The Interpreters' Newsletter» (XXIII), pp. 133-149.
- Ferraresi, A. 2016, *Intermodal corpora and the translation classroom: What can translation trainers and trainees learn from interpreting?* «Linguaculture» (VII-II), pp. 27-51.
- Ferraresi, A., S. Bernardini 2019, *Building EPTIC: A many-sided, multi-purpose corpus of EU parliament proceedings*, in I. Doval, M. T. Sánchez Nieto (eds.), *Parallel Corpora for Contrastive and Translation Studies. New resources and applications*, Amsterdam/Philadelphia: John Benjamins, pp. 123-139.
- Halverson, S. 1998, *Translation studies and representative corpora: Establishing links between translation corpora, theoretical/descriptive categories and a conception of the object of study*, «Meta» (XLIII-IV), pp. 494-514, <http://id.erudit.org/iderudit/003000ar>.

- Kellett Bidoli, C. J. 2016, *Methodological challenges in consecutive interpreting research: Corpus analysis of notes*, in C. Bendazzoli, C. Monacelli (eds.), *Addressing Methodological Challenges in Interpreting Research*, Newcastle upon Tyne: Cambridge Scholars Publishing, pp. 143-171.
- Kilgarrieff, A. et al. 2014, *The Sketch Engine: Ten years on*, «Lexicography: Journal of ASIA-LEX» (I-I), pp. 7-36, <https://doi.org/10.1007/s40607-014-0009-9>.
- Kohn, K. 2012, *Pedagogic corpora for content and language integrated learning Insights from the BACKBONE project*, «The EUROCALL Review» (XX-II), pp. 3-22.
- Leeson, L. 2008, *Quantum leap: Leveraging the Signs of Ireland digital corpus in Irish Sign Language/English interpreter training*, «The Sign Language Translator and Interpreter» (II-II), pp. 149-176.
- Niemants, N. S. A. 2012, *The transcription of interpreting data*, «Interpreting» (XIV-II), pp. 165-191.
- Niemants, N. S. A. 2013, *From role-playing to role-taking: Interpreter role(s) in healthcare*, in C. Schäffner, K. Kredens, Y. Fowler (eds.), *Interpreting in a Changing Landscape: Selected Papers from Critical Link 6*, Amsterdam/Philadelphia: John Benjamins, pp. 305-319.
- Orozco-Jutorán, M. 2018, *The TIPp project. Developing technological resources based on the exploitation of oral corpora to improve court interpreting*, in C. Bendazzoli, M. Russo, B. Defrancq (eds.), *New Findings in Corpus-based Interpreting Studies*. Special issue of «inTRAlinea», <https://www.intralinea.org/index.php/specials/article/2316>.
- Russo, M. 2010, *Reflecting on interpreting practice: Graduation theses based on the European Parliament Interpreting Corpus (EPIC)*, in L. N. Zybatow (ed.), *Translationswissenschaft – Stand und Perspektiven. Innsbrucker Ringvorlesungen zur Translationswissenschaft VI*, Frankfurt: Peter Lang, pp. 35-50.
- Russo, M. et al. 2012, *The European Parliament Interpreting Corpus (EPIC): Implementation and developments*, in F. Straniero Sergio, C. Falbo (eds.), *Breaking Ground in Corpus-Based Interpreting Studies*, Bern: Peter Lang, pp. 35-90.
- Russo, M., Bendazzoli, C., Defrancq, B. 2018, (eds.), *Making way in Corpus-based Interpreting Studies*, Singapore: Springer.
- Sandrelli, A. 2010, *Corpus-based Interpreting Studies and interpreter training: A modest proposal*, in L. N. Zybatow (ed.), *Translationswissenschaft – Stand und Perspektiven. Innsbrucker Ringvorlesungen zur Translationswissenschaft VI*, Frankfurt: Peter Lang, pp. 69-90.
- Schmidt, T., K. Wörner 2014, *EXMARaLDA*, in J. Durand, U. Gut, G. Kristoffersen (eds.), *The Oxford Handbook of Corpus Phonology*, Oxford: Oxford University Press, pp. 402-419.
- Setton, R. 2011, *Corpus-based Interpreting Studies (CIS): Overview and prospects*, in A. Kruger, K. Wallmach, J. Munday (eds.), *Corpus-based Translation Studies: Research and Applications*, London/New York: Continuum, pp. 33-75.
- Shlesinger, M. 1998, *Corpus-based Interpreting Studies as an offshoot of Corpus-based Translation Studies*. «Meta» (XLIII-IV), pp. 486-493, <http://id.erudit.org/iderudit/004136ar>.

- Spinolo, N. 2018, *Studying figurative language in simultaneous interpreting: The IMITES (Interpretación de la Metáfora entre ITaliano y ESpañol) corpus*, in M. Russo, C. Bendazzoli, B. Defrancq (eds.), *Making Way in Corpus-bases Interpreting Studies*, Singapore: Springer, pp. 133-155.
- Straniero Sergio, F., C. Falbo 2012 (eds.), *Breaking Ground in Corpus-Based Interpreting Studies*, Frankfurt: Peter Lang.
- Thompson, P. 2004, *Spoken language corpora*, in M. Wynne (ed.) *Developing Linguistic Corpora: A Guide to Good Practice*, <https://users.ox.ac.uk/~martinw/dlc/chapter5.htm>

Letture di approfondimento

- Bendazzoli, C. 2015, *Corpus-based research*, in F. Pöchhacker (ed.), *The Routledge Encyclopedia of Interpreting Studies*, London: Routledge, pp. 87-91.
- Bendazzoli, C. 2017 (ed.), *Corpus-based dialogue interpreting studies*. Special issue of «The Interpreters' Newsletter» (XXII), <https://www.openstarts.units.it/handle/10077/20733>.
- Dal Fovo, E. 2018, *The use of dialogue interpreting corpora in healthcare interpreter training: Taking stock*, «The Interpreters' Newsletter» (XXIII), pp. 83-113.

INTERPRETARE DA E VERSO L'ITALIANO

Didattica e innovazione
per la formazione dell'interprete

a cura di
Mariachiara Russo

Bononia
University Press

La versione digitale di questo volume è disponibile gratuitamente grazie al contributo dell'Alma Mater Studiorum - Università di Bologna.

Visita buponline.com/openteaching

Progetto Open Teaching Consorzio Alphabet

Bononia University Press

Via Saragozza 10

40123 Bologna

tel. (+39) 051 232882

fax (+39) 051 221019

www.buponline.com

e-mail: info@buponline.com

ISBN 978-88-6923-987-8

ISBN online 978-88-6923-988-5

Quest'opera è pubblicata sotto licenza Creative Commons BY-NC-SA 4.0

Progetto di copertina: Design People (Bologna)

Impaginazione: DoppioClickArt - San Lazzaro di Savena (BO)

Prima edizione: dicembre 2021

SOMMARIO

Un Testo per la formazione dell'interprete: Introduzione <i>Mariachiara Russo</i>	XI
--	----

PARTE I ASPETTI TEORICI E METODOLOGICI

1. L'annotazione in interpretazione consecutiva <i>Amalia Amato</i>	3
2. L'interpretazione simultanea <i>Gabriele Mack</i>	19
3. Teoria e prassi dell'interpretazione dialogica <i>Natacha Niemants</i>	41
4. L'interpretazione a distanza <i>Nicoletta Spinolo</i>	61
5. Interpretazione e ricerca su aspetti neurolinguistici e cognitivi <i>Serena Ghiselli e Mariachiara Russo</i>	79
6. Interpretazione, competenze trasversali e glottodidattica teatrale <i>María Isabel Fernández García, Ivonne Lucilla Simonetta Grimaldi, Claudio Bendazzoli</i>	97
7. Quale <i>linguistica</i> per gli interpreti in formazione? Un <i>excursus</i> sulle competenze in lingua madre <i>Cristiana Cervini</i>	115
8. Interpretazione e nuove tecnologie <i>Claudia Lecci</i>	135

PARTE II
PROBLEMI, STRATEGIE ED ESERCITAZIONI PER COPPIE
DI LINGUE: INTERPRETARE COME

1. Interpretare tra BCMS e italiano
Eleonora Bernardi e Sara Polidoro 153
2. Interpretare tra cinese e italiano
Han Wang e Riccardo Moratto 173
3. Interpretare tra francese e italiano
Vincenzo Lambertini, Lucia Baldi e Patricia Toni 191
4. Interpretare tra inglese e italiano
Veronica Irene Cioni, Ira Torresi e Christopher Garwood 211
5. Interpretare tra polacco e italiano
Mirko Coleschi 229
6. Interpretare tra portoghese e italiano
Lucia Melotti e Salvador Pippa 247
7. Interpretare tra russo e italiano
Lyubov Bezkrovna, Kristina Landa e Sara Polidoro 267
8. Interpretare tra spagnolo e italiano
Michela Bertozzi, María Jesús González Rodríguez e Mariachiara Russo 289
9. Interpretare tra tedesco e italiano
Gabriele Mack e Miriam P. Leibbrand 313

PARTE III
AMBITI PROFESSIONALI: INTERPRETARE PER CHI

1. Interpretare per la giustizia: tribunali e intercettazioni
Elio Ballardini e María Jesús González Rodríguez 337
2. Interpretare per bambini/e e ragazzi/e in contesti legali
Amalia Amato 359
3. Interpretazione umanitaria
María Jesús González Rodríguez e Maura Radicioni 373

4. Interpretare per i servizi socio-sanitari <i>Natacha Niemants e Eleonora Bernardi</i>	395
5. Educational Interpreting. Ovvero l'interpretazione simultanea delle lezioni accademiche <i>Francesco Cecchi</i>	415
6. Child language brokering e interpretazione <i>Rachele Antonini e Ira Torresi</i>	441

PARTE IV ASPETTI GENERALI

1. Interpretación: historia y profesión <i>Iciar Alonso-Araguás e Jesús Baigorri-Jalón</i>	459
2. Conference Interpreters: Role and Agency <i>Ebru Diriker</i>	477
3. Interpreting as Communication: Risks and Challenges <i>Yves Gambier</i>	495
4. Quality Evaluation in Conference Interpreting <i>Sylvia Kalina</i>	521
5. Didattica dell'interpretazione basata su corpora <i>Claudio Bendazzoli</i>	539
Indice tematico	555
Note sugli autori e sulle autrici	567