

Rassegna – UDK 81'255=163.42=131.1

81'255:004.78

DOI: 10.32728/studpol/2018.07.01.01

## GLI STRUMENTI DI TRADUZIONE E LE RISORSE LINGUISTICHE NELLE TRADUZIONI DAL CROATO ALL'ITALIANO E DALL'ITALIANO AL CROATO

Ivana Lalli Pačelat  
Università degli Studi Juraj Dobrila di Pola (Croazia)  
ilalli@unipu.hr

Petra Maras  
petra.maras@yahoo.com

### Riassunto

Gli strumenti di traduzione e le risorse linguistiche sono in continuo sviluppo e i traduttori devono essere aggiornati sulle novità nel campo delle tecnologie di traduzione per poter soddisfare le esigenze del mercato.

Lo scopo di questo saggio è di presentare generalmente le risorse linguistiche e gli strumenti di traduzione che hanno a disposizione i traduttori d'oggi. Un interesse particolare è rivolto a quelle risorse linguistiche e quegli strumenti di traduzione che possono venir usati nelle traduzioni dal croato all'italiano e dall'italiano al croato. Il saggio offre una presentazione generale di tutti gli elementi che fanno parte della stazione di lavoro del traduttore e in particolar modo, della traduzione automatica, delle memorie di traduzione, degli strumenti per la gestione della terminologia e dei corpora.

Con questa breve rassegna degli strumenti e delle risorse disponibili nel campo della traduzione moderna si è voluto sottolineare l'importanza e l'indispensabilità dell'impiego delle tecnologie della traduzione da parte dei traduttori per poter rispondere alle richieste di alta qualità, precisione e velocità, poste dal mercato globale della traduzione.

Parole chiave: CAT tools, traduzione automatica, corpora, banca dati terminologica, memorie di traduzione

## **Introduzione**

Le tecnologie linguistiche e di traduzione sono in continuo sviluppo e i traduttori hanno bisogno di essere sempre aggiornati sui nuovi sussidi che possono velocizzare il processo di traduzione e garantire un livello di traduzione di alta qualità, consentendo loro di essere concorrenti sul mercato del lavoro. Sono proprio queste le motivazioni per la stesura di questo lavoro, che potrà essere utile sia agli studenti di traduzione sia ai traduttori, per aggiornarli sulle novità nel campo delle tecnologie della traduzione. Novità che, data la velocità dei cambiamenti nel settore delle tecnologie linguistiche e della traduzione, si possono trovare difficilmente nei manuali pubblicati anche solamente un paio di anni fa. Inoltre, il lavoro, oltre a fornire una panoramica generale degli strumenti e delle risorse a supporto del traduttore, offre una presentazione dettagliata di quegli strumenti che possono venir usati per le traduzioni dal croato all'italiano e viceversa, come pure delle risorse che includono le due lingue in oggetto.

### **1. Gli strumenti e le risorse della traduzione**

Le risorse linguistiche e gli strumenti di traduzione, che i traduttori hanno a disposizione nel loro lavoro oggi, consentono loro di velocizzare notevolmente il processo di traduzione, di produrre testi di alta qualità e di mantenere la coerenza terminologica e stilistica. Il processo di traduzione diventa più rapido grazie all'uso delle memorie di traduzione, che permettono ai traduttori di salvare il lavoro fatto per poi riutilizzarlo per i lavori futuri. Le memorie di traduzione, assieme alle banche dati terminologiche, garantiscono la coerenza terminologica e stilistica dei testi tradotti. Tra gli strumenti e le risorse principali che usano i traduttori moderni si trovano non solo le memorie di traduzione e le banche dati terminologiche, ma anche i corpora e i sistemi di traduzione automatica, che sono ormai da tempo parte integrante della stazione di lavoro dei traduttori. Infatti, le piattaforme per la gestione completa dell'attività di traduzione, integrano programmi per la correzione, la revisione, la documentazione, il controllo e la garanzia di qualità, il coordinamento dei collaboratori, ecc. con strumenti e risorse per la traduzione (memorie

di traduzione, banche dati terminologiche, la traduzione automatica personalizzata). Si tratta di sistemi integrali per la gestione di progetti che facilitano e gestiscono il flusso della traduzione garantendo al contempo un alto livello di qualità e professionalità.

## 2. Traduzione automatica

La traduzione automatica è una traduzione nella quale il computer, con l'ausilio di un programma, traduce il testo originale da una lingua all'altra senza la partecipazione umana. L'uso e il risultato della traduzione automatica dipendono dal tipo di testo che si sottopone alla traduzione automatica, dalla combinazione linguistica, dalla disponibilità di corpora paralleli e dallo scopo della traduzione (cfr. Arduini e Stecconi, 2007, 172; Pavlović, 2015, 292). Particolarmente adatte alla traduzione automatica sono le lingue correlate e le coppie linguistiche che possiedono un gran numero di corpora paralleli (cfr. Pavlović, 2015, 292). Alcuni tipi di testi sono più adatti alla traduzione automatica rispetto ad altri. A essere adeguati alla traduzione automatica sono in genere i testi tecnici ripetitivi come, ad esempio, i vari manuali e le istruzioni d'uso. Grazie alle loro strutture lessicali e grammaticali semplificate, la macchina riesce a produrre traduzioni soddisfacenti (cfr. Arduini e Stecconi, 2007, 172). I risultati sono migliori se si tratta di testi specialistici di ambito molto ristretto. Comunque, dipenderà dallo scopo della traduzione se il testo dovrà essere sottoposto a una revisione e rielaborazione da parte dei traduttori umani. Se si tratta di un testo che fornisce solo informazioni di base, che riguardano il contenuto del testo, la traduzione automatica potrà venir utilizzata con modifiche minime o senza alcune modifiche (Pavlović, 2015, 292).

Nello sviluppo della traduzione automatica, iniziata negli anni Cinquanta del secolo scorso, “pur con periodi di coesistenza e sovrapposizione di tecnologie, a grandi linee si sono succedute tre fasi caratterizzate da altrettanti modelli”: la traduzione basata su regole (ingl. *Rule-based Machine Translation* - RBMT), la traduzione statistica (ingl. *Statistical Machine Translation* - SMT) e “l'ormai dominante traduzione

neurale” (ingl. *Neural Machine Translation* - NMT) (Brusasco, 2018). Per approfondimenti sullo sviluppo e sulla storia della traduzione automatica si rimanda, tra gli altri, a Hutchins e Somers (1992), Chiari (2007) e Koehn (2010) e invece per un’analisi dei vantaggi e degli svantaggi dal punto di vista linguistico dei tre modelli di traduzione automatica si consulti, ad esempio, Burchardt e collaboratori (2017).

Uno dei più famosi e dei primi servizi di traduzione automatica di tipo statistico offerti gratuitamente in Rete è sicuramente *Google Translate*, seguito da *Bing Microsoft Translator* e *Babel Fish*. Si tratta della “traduzione automatica del tipo *cloud*, a cui l’utente si collega per tradurre gratuitamente testi, documenti o pagine web in tempo reale” (Monti, 2014, 58). Le traduzioni ottenute da questi programmi di traduzione automatica, fino a qualche anno fa erano di bassa qualità, da usare a solo scopo informativo, che però rispondevano “alle esigenze degli utenti di accedere alle informazioni in una lingua che non conoscono a costo zero e in tempi rapidissimi” (Monti, 2014, 58); in tempi recenti sono diventate migliori e più adeguate al contesto e alla situazione, grazie al sistema basato sulle reti neurali. Infatti, non solo *Google*, ma anche *Systran* e *Microsoft* hanno annunciato nel 2016 il passaggio ai nuovi sistemi di traduzione automatica basata sulle reti neurali (Brusasco, 2018). La traduzione neurale diventa “concorrente diretto alla traduzione automatica statistica e in un certo senso anche a quella basata sulle regole” (Forcada, 2017). Per maggiori informazioni sulla traduzione automatica basata sulle reti neurali si vedano, ad esempio, Koehn (2017) e Forcada (2017).<sup>[1]</sup> La grande svolta nello sviluppo della traduzione automatica è strettamente collegata ai moderni successi nel campo dell’intelligenza artificiale, che sono stati uno dei temi principali, tra l’altro, anche al *Translating Europe Forum 2018*<sup>[2]</sup> nell’organizzazione della Commissione europea.

La Commissione europea, infatti, sviluppa e gestisce da anni un servizio proprio di traduzione automatica. Fino al 2017 veniva usato il servizio MT@EC. Si tratta di un servizio di traduzione automatica basata

---

[1] Per ulteriori approfondimenti si rimanda all’archivio digitale di articoli che riguardano la traduzione automatica *Machine Translation Archive* dell’Associazione Europea per la traduzione automatica (EAMT) disponibile su <http://www.mt-archive.info/> (consultato il 10/10/2018).

[2] Il programma e le presentazioni del *Translating Europe Forum 2018*, tenutosi a Bruxelles l’8 e il 9 novembre 2018, sono consultabili su: [https://ec.europa.eu/info/events/2018TEF\\_en](https://ec.europa.eu/info/events/2018TEF_en) (consultato il 10/11/2018).

su dati piuttosto che su regole e sul software statistico Moses arricchito da regole linguistiche (cfr. European Commission, 2014, 3). Sviluppato nell'ambito dei programmi di ricerca dell'UE, Moses è il sistema per la traduzione automatica *open source* più utilizzato al mondo (cfr. European Commission, 2014, 3; Koehn et al., 2007). Usando corpora contenenti traduzioni ufficiali di documenti UE in 24 lingue ufficiali, il servizio offre la traduzione automatica per più di 552 coppie di lingue (European Commission, 2014, 3). Continuando il servizio del MT@EC, nel 2017 viene lanciato il servizio di traduzione automatica *eTranslation*<sup>[3]</sup> offerto nell'ambito dei servizi dell'EU *Connecting Europe Facility* (CEF) che la Commissione europea mette a disposizione di tutte le istituzioni amministrative degli stati membri dell'Unione Europea.

È, appunto, questo tipo di traduzione automatica adattata e sviluppata su testi specialistici di un determinato settore e per determinati scopi ad offrire i migliori risultati. L'industria delle tecnologie linguistiche offre, perciò, negli ultimi tempi, lo sviluppo e la gestione di sistemi di traduzione personalizzati e adattati ai bisogni di un determinato tipo di traduzione o di un cliente. Esistono progetti scientifici, come ad esempio *LetsMT!*<sup>[4]</sup> (Vasiljevs et al., 2011), che offrono il servizio gratuito on-line di traduzione automatica personalizzata basata su corpora paralleli e banche dati terminologiche consegnate dal cliente. Numerosissimi sono, invece, i servizi che vengono offerti per sviluppare sistemi personalizzati di traduzione automatica a pagamento, di cui la maggior parte non è limitata a un determinato settore, come ad esempio *Globalese*, *Tauyou*, *Tilde MT* e altri, mentre alcuni sono specializzati solo per un settore, come ad esempio *Lingua Custodia* per il settore finanziario. L'adattamento della traduzione automatica ai propri testi viene offerta anche da piattaforme per la gestione completa dell'attività di traduzione come ad esempio *Wordbee*, *My Acclaro*, *XTM Cloud*, *Memsources Cloud*, *SDL Trados Studio*, *memoQ*, *WordfastPro*, *Smartcat* ecc.

I recenti successi nello sviluppo della traduzione automatica basata sulle reti neurali ha sollevato nuovamente la questione del futuro dei

---

[3] Per più informazioni si rimanda a: <https://ec.europa.eu/cefdigital/wiki/display/CEFDIGITAL/What+is+eTranslation++Overview> (consultato il 10/10/2018).

[4] Per più informazioni si veda: <http://www.letsmt.eu/Login.aspx> (consultato il 10/10/2018).

traduttori umani. Tema discusso sia al *Translating Europe Forum 2016*<sup>[5]</sup> che a quello di quest'anno. La professione dei traduttori è cambiata e i traduttori devono adeguarsi ai nuovi cambiamenti sul mercato dovuti allo sviluppo delle tecnologie linguistiche. Un'attività molto frequente, che negli ultimi anni viene inserita anche nei corsi universitari per la formazione dei futuri traduttori, è quella del *pre- e post-editing*, cioè la preparazione di testi per la traduzione automatica e la loro revisione. Secondo Forcada (2017), proprio perché la maggior parte dei traduttori lavora su testi prodotti dalla traduzione automatica basata sulle reti neurali è necessario che i traduttori siano a conoscenza dei nuovi approcci nella traduzione automatica. Una spiegazione semplice della traduzione neurale adattata ai bisogni dei traduttori viene offerta, ad esempio, da Forcada (2017).

### **2.1. Pre e Post-editing**

Il traduttore umano “interviene in aiuto al sistema di traduzione automatica prima, durante e dopo il processo di traduzione” (Scarpa, 2008, 304). Nella fase che precede la traduzione automatica, nel cosiddetto *pre-editing*, bisogna “assicurarsi che il testo originale sia ben scritto, senza errori di ortografia o di punteggiatura e senza ambiguità” (TAUS 2010)<sup>[6]</sup>. Il testo da tradurre viene, infatti, “preparato per la macchina per migliorare le prestazioni del sistema tramite il correttore ortografico, grammaticale e stilistico, al fine di ottenere una lingua controllata” (Scarpa, 2008, 304). Nella fase di postrevisione o *post-editing*, il testo prodotto dalla traduzione automatica viene sottoposto alla revisione da parte di un traduttore umano.

Esistono varie linee guida che aiutano i traduttori nel lavoro di pre- e postrevisione (cfr. O'Brien, 2010; Massardo et al., 2016). Per uno studio comparato delle linee guida per la postrevisione si rimanda a Hu e Cadwell (2016). Gli autori delle linee guida generalmente concordano sul fatto che l'impegno richiesto nella fase di postrevisione dipenda da due fattori. Da un lato, l'impegno viene determinato dallo scopo della traduzione e dal pubblico di riferimento cioè dalla qualità della traduzione che viene

---

[5] Per più informazioni sul *Translating Europe Forum 2016* si consulti: [https://ec.europa.eu/info/events/2018TEF\\_en](https://ec.europa.eu/info/events/2018TEF_en) (consultato il 6/9/2018).

[6] Le linee guida sono disponibili su: <https://www.taus.net/academy/best-practices/evaluate-best-practices/linee-guida-per-la-postrevisione-di-traduzioni-automatiche> (consultato il 14/9/2018).

richiesta e dall'altro lato, dalla qualità della traduzione automatica grezza (cfr. O'Brien, 2010; TAUS, 2010; Massardo et. al., 2016). Normalmente, si distinguono due livelli di revisione: la revisione completa e la revisione leggera (Massardo et al. 2016, 16). La prima serve a “raggiungere una qualità simile ad una traduzione umana di alta qualità altrimenti detta qualità pubblicabile” (TAUS, 2010; Massardo et al., 2016, 17). La revisione completa ha come risultato testi comprensibili e grammaticamente corretti con una sintassi normale e scritti in uno stile appropriato, “anche se lo stile può non essere all'altezza di quello prodotto da un traduttore umano madrelingua” (TAUS, 2010; Massardo et al., 2016, 18). Invece, secondo le linee guida per ottenere una qualità di livello più basso, spesso definita come *sufficientemente buona*, di solito viene raccomandata una postrevisione leggera. Dopo una revisione leggera si hanno testi comprensibili, che trasmettono il messaggio fedelmente, ma con una sintassi insolita e una grammatica e uno stile poco convincente (cfr. TAUS, 2010; Massardo et al., 2016, 17).

Il lavoro del *pre* e *post-editing*, come abbiamo precedentemente spiegato, è un'attività molto frequente nelle grandi organizzazioni e istituzioni che hanno sviluppato i propri sistemi di traduzione automatica e perciò Diadori (2012, 333) sostiene che ai traduttori umani rimane “la sola funzione di revisori di testi”, sempre con lo scopo di ridurre i costi e i tempi di traduzione.

### 3. Traduzione assistita

Il termine traduzione assistita da macchina si riferisce alla situazione in cui il traduttore lavora con l'aiuto di uno o più programmi informatici specializzati per la traduzione. Tutti i programmi e ausili tecnici che aiutano i traduttori sono spesso chiamati anche in italiano i *CAT tools*, dall'inglese *Computer Assisted Translation* o *Computer Aided Translation* (tutti e due abbreviati in CAT). I *CAT tools* comprendono memorie di traduzione, strumenti per la gestione della terminologia, strumenti per la localizzazione, strumenti per il sottotitolaggio, strumenti per il controllo della qualità, strumenti per la redazione dei testi ecc. (Pavlović, 2015, 292; Scarpa, 2008, 305; e altri). Oltre ad “aumentare la produttività dei traduttori

i CAT sono sistemi integrati che, come precedentemente già nominato, offrono un'ampia gamma di altre funzionalità che trovano impiego in tutte le fasi del processo di traduzione” (Scarpa, 2008, 306). Normalmente, la stazione di lavoro di un traduttore comprende le memorie di traduzione, strumenti di gestione della terminologia, risorse terminologiche, corpora e sistemi di traduzione automatica, sistemi per la documentazione e redazione dei testi ecc. Sono tutto strumenti e risorse fondamentali che vengono incorporati nei CAT *tools* disponibili oggigiorno sul mercato.

Il CAT *tool* più famoso è *SDL Trados Studio*<sup>[7]</sup>, usato anche dai traduttori dell'EU e presente sul mercato da oltre trent'anni. Un altro tra i CAT *tools* più usati dai traduttori professionisti è *Wordfast*<sup>[8]</sup>, che accanto alla versione a pagamento, mette a disposizione degli utenti anche una versione gratuita. Oltre ai programmi a pagamento (*SDL Trados*, *Wordfast*, *memoQ*, *Memsources*, *MemoAcross*, *Déja Vu* ecc.) che usano i traduttori professionisti, esistono anche programmi disponibili gratuitamente in Rete. Sandrini (2012, 1) sostiene che tutti i programmi disponibili gratuitamente possono aiutare un traduttore a creare una stazione di lavoro indipendente e gratuita. Tra i servizi gratuiti accessibili on-line, bisogna menzionare *OmegaT*<sup>[9]</sup>, *Pootle*<sup>[10]</sup>, *MateCat*<sup>[11]</sup>, *WF Anywhere*<sup>[12]</sup> e molti altri.

Tutte le istituzioni multilingui, per le quali la traduzione ha un ruolo importante e dove vengono richieste traduzioni di alta qualità, hanno sviluppato da tempo strumenti e risorse per facilitare le attività traduttive e per ridurre i costi e i tempi di traduzione. Un ottimo esempio dell'organizzazione delle attività traduttive è stato presentato dalla Commissione europea e spiegato nel dettaglio in una delle sue pubblicazioni sulla traduzione (European Commission, 2016). Oltre ai classici strumenti di traduzione e alle risorse linguistiche, a disposizione dei traduttori della Commissione troviamo tantissimi altri strumenti e risorse che rendono

---

[7] Secondo i risultati di una ricerca sull'uso dei CAT tools condotta da ProZ.com e pubblicata su prozcomblog nel 2013 il CAT-tool preferito è SDL Trados, segue Wordfast e memoQ. Gli altri programmi come SDLX, Across, Déja Vu e OmegaT sono usati da meno di venti per cento degli intervistati. Per più informazioni si veda: <https://prozcomblog.com/2013/03/28/cat-tool-use-by-translators-what-are-they-using/> (consultato il 15/6/2018).

[8] Per maggiori informazioni si veda: [https://wordfast.com/about\\_whatwedo](https://wordfast.com/about_whatwedo) (consultato il 14/7/2018).

[9] Per più informazioni si veda: <http://omegat.org/> (consultato il 14/7/2018).

[10] Per più informazioni si veda: <http://pootle.translatehouse.org/> (consultato il 14/7/2018).

[11] Per più informazioni si veda: <https://www.matecat.com/> (consultato il 14/7/2018).

[12] Per più informazioni si veda: <https://www.freem.com/> (consultato il 14/7/2018).



più efficace il loro lavoro. *Poetry* è un portale per i traduttori che serve per ricevere le richieste per le traduzioni dagli altri dipartimenti della Commissione, mentre *ManDesk* serve a gestire le richieste di traduzione all'interno della Direzione generale della traduzione (DGT) (cfr. European Commission, 2016). Da menzionare sono assolutamente *Euramis*, cioè la memoria di traduzione centrale della DG Traduzione, *TraDesk*, l'interfaccia per la gestione dei documenti di traduzione e per l'accesso agli archivi delle traduzioni fatte in passato, *DGT Vista*, un motore di ricerca specializzata nell'archivio delle traduzioni, *EUR-Lex*, una banca dati on-line di tutti i testi di legge dell'Unione Europea e *IATE*, la banca dati terminologica interistituzionale dell'UE (European Commission, 2016).

### 3. 1. Corpora

I corpora vengono generalmente definiti come collezioni di testi autentici prodotti in contesti comunicativi reali, salvati in formato digitale e “spesso corredati di strumenti di consultazione informatici” (Baroni, 2010).

I corpora sono usati nell'attività di traduzione da anni sia in modo diretto sia in modo indiretto. Essendo utili nel processo di traduzione, i corpora sono oggi molto spesso parte integrante della stazione di lavoro del traduttore. L'impiego indiretto dei corpora si ha, invece, nella lessicografia e nella linguistica computazionale. Infatti, i corpora svolgono un ruolo importantissimo nello sviluppo delle opere lessicografiche e delle tecnologie di traduzione.

Generalmente, gli strumenti di gestione e interrogazione dei corpora offrono all'utente la possibilità di analizzare un corpus mediante le concordanze, le collocazioni, le liste di frequenza, le parole chiave, gli *wordsketch*, ecc. Le concordanze vengono definite come “una lista delle occorrenze di una o più forme, ciascuna mostrata all'interno del contesto in cui compare nel corpus” (Spina 2001, 127). Gli esempi d'uso di una o più parole nei testi autentici prodotti in contesti comunicativi reali possono aiutare il traduttore nella ricerca della parola o l'espressione giusta per la sua traduzione. Oltre al vantaggio dell'autenticità dei testi inseriti nei corpora, un altro vantaggio per la traduzione presenta pure

l'informazione sulla frequenza di una determinata parola o gruppo di parole, che può influenzare la scelta del traduttore. Un aiuto nella ricerca e nella scelta può venir offerto anche dalle collocazioni, cioè dalla “co-occorrenza statisticamente significativa di due o più parole” (Lalli Paćelat, 2016). Inoltre, grazie alle liste di frequenza e le liste di parole chiave, i traduttori “con un paio di click del mouse possono individuare le principali caratteristiche lessico-grammaticali di una varietà linguistica (terminologia di base, espressioni frequenti, strutture grammaticali tipiche, ecc.” (Lalli Paćelat, 2016).

Esistono diversi tipi di corpora. Per una presentazione dettagliata della tipologia dei corpora si rimanda, ad esempio, a McEnery e Hardie (2012). Per quanto concerne l'attività di traduzione, bisogna distinguere i corpora monolingui da quelli multilingui.

Tra i corpora monolingui, per i traduttori rivestono un ruolo importante i corpora generali o di riferimento e quelli specialistici. Tra i corpora generali o di riferimento per la lingua croata e italiana vanno menzionati il *Corpus nazionale croato (HNK)*<sup>[13]</sup> (Tadić, 2009) e il *Corpus di italiano scritto CORIS/ CODIS*<sup>[14]</sup> (Rossini Favretti et al., 2002) e il *Perugia Corpus (PEC)*<sup>[15]</sup> (Spina, 2001). Bisogna specificare che “i corpora generali di riferimento contengono centinaia di milioni di parole e vengono considerati rappresentativi per una lingua in generale” (Lalli Paćelat, 2016), d'altra parte esistono corpora di dimensioni maggiori formati da testi scaricati automaticamente da Internet. Si tratta, tra gli altri, dei *Web as a corpus*, come lo sono ad esempio i corpora *hrWaC*<sup>[16]</sup> (Ljubešić e Erjavec, 2011) e *itWaC*<sup>[17]</sup> che sono accessibili on-line e hanno un'interfaccia d'interrogazione semplice per gli utenti.

I corpora specialistici sono invece di minor dimensione e vengono compilati per uno scopo determinato (cfr. Koester, 2010). Sono molto utili nella traduzione, perché consentono di individuare la terminologia

---

[13] Per più informazioni si veda: <http://www.hnk.ffzg.hr/> (consultato il 17/6/2018)

[14] Il corpus è interrogabile da questo sito: <http://corpora.ficlit.unibo.it/> (consultato il 18/6/2018)

[15] Il corpus è interrogabile da questo sito: <https://www.unistrapg.it/perugiacorporus/index.html> (consultato il 18/6/2018)

[16] Il corpus è interrogabile da questo sito: [http://nl.ijs.si/noske/all.cgi/first\\_form?corpname=hrwac;align=](http://nl.ijs.si/noske/all.cgi/first_form?corpname=hrwac;align=) (consultato il 19/6/018)

[17] Il corpus è interrogabile da questo sito: <https://www.sketchengine.eu/itwac-italian-corporus/> (consultato il 19/6/2018).

specialistica, le espressioni frequenti e le strutture grammaticali tipiche di quel linguaggio settoriale e servono per la compilazione di banche dati terminologiche (cfr. Bowker e Pearson, 2002; Lalli Paćelat, 2016; Zanettin, 2014; e altri).

Oltre ai corpora monolingui esistono i corpora bilingui o multilingui.

Se si tratta di corpora in due o più lingue diverse composti da testi originali paragonabili per quanto riguarda, ad esempio, l'argomento, il genere testuale o il periodo di stesura, allora si parla di corpora comparabili (cfr. McEnery e Xiao, 2007; Kenning, 2010; e altri). I corpora comparabili trovano un impiego nell'attività di traduzione, perché permettono al traduttore di individuare "le norme e convenzioni redazionali" usate nei testi originali, testi autentici scritti da parlanti nativi, sia della lingua di partenza che d'arrivo e possono essere uno strumento utile "nella revisione e nella produzione di traduzioni più naturali" (cfr. Scarpa, 2008, 315).

Invece, se i corpora contengono testi originali e le rispettive traduzioni, allora vengono chiamati corpora paralleli. Per una descrizione dettagliata dei tipi di corpora paralleli esistenti e della rispettiva terminologia si rimanda, tra gli altri, a McEnery e Hardie (2012, 19). I corpora paralleli devono venir allineati per poter essere usati nelle memorie di traduzione. L'allineamento viene fatto da un programma automatico che abbina i segmenti che sono l'uno la traduzione dell'altro. Il traduttore verifica l'esattezza dell'abbinamento delle unità di traduzione proposte dal programma automatico e può accettarle, annullarle, unirle, dividerle o apportare altre modifiche. Dopo aver effettuato la revisione, si ha un file bilingue, che può venir inserito in una memoria di traduzione (cfr. McEnery e Hardie, 2012, 20; Tufiş, 2007, 190; e altri).

Tra i corpora paralleli allineati bisogna menzionare quelli formati da testi prodotti nell'ambito dell'UE, come il corpus multilingue *EUROPARL*, un corpus composto da testi paralleli in undici lingue contenente più di venti milioni di parole per ogni lingua (cfr. Koehn, 2002, 2), e il corpus *JRC Acquis*<sup>[18]</sup> (Steinberger et al., 2006), una collezione di testi dell'ambito legale che include testi scritti dal 1950 fino ad oggi in ventidue lingue ufficiali dell'UE. Per una rassegna completa dei corpora multilingui

---

[18] Per ulteriori informazioni si rimanda al sito: <https://ec.europa.eu/jrc/en/language-technologies/jrc-acquis> (consultato il 19/7/2018).

dell'Unione Europea si rimanda a Steinberger e collaboratori (2014).

I corpora paralleli che includono invece sia la lingua italiana sia quella croata e che si possono consultare tramite un'interfaccia di interrogazione semplice per gli utenti sono: l'*OPUS2 parallel corpus* (Tiedemann, 2012), l'*EUR-Lex Corpus*<sup>[19]</sup> (Baisa et al., 2016), l'*Eur-Lex judgments corpus*<sup>[20]</sup> (Baisa et al., 2016) e la *DGT-Translation Memory*<sup>[21]</sup> (Steinberger et al., 2012; 2014). I corpora menzionati si trovano su *Sketch Engine*<sup>[22]</sup> che permette ai traduttori di osservare, oltre alle concordanze semplici, alle liste di parole, alle liste di n-grammi, alle parole chiave, agli *Word Sketch* e al tesaurus, anche le concordanze parallele.

Si conclude questo capitolo riportando la spiegazione di Kübler e Aston (2010) sull'utilità d'uso dei corpora in ogni fase del processo di traduzione. Nella fase di documentazione, i corpora comparabili possono offrire un punto di partenza per i traduttori, in particolare se si tratta di nuovi domini tematici o generi testuali, per cui un paragone dei testi autentici redatti nella lingua di partenza e di quelli redatti nella lingua d'arrivo può aiutare il traduttore a formarsi una base concettuale e terminologica del nuovo dominio tematico e a osservare le peculiarità e le differenze interlinguistiche dei tipi di testi in questione (cfr. Kübler e Aston, 2010, 501). Nella fase di redazione possono essere utili sia i corpora paralleli che quelli comparabili. I corpora paralleli permettono di osservare le traduzioni fatte da altri traduttori, mentre quelli comparabili aiutano nella ricerca della parola o dell'espressione adeguata al contesto d'uso. Questi ultimi trovano impiego anche nella fase di revisione della traduzione, in quanto consentono di verificare l'uso della parola o espressione scelta nei testi autentici dello stesso dominio tematico redatti da parlanti nativi della lingua d'arrivo (cfr. Kübler e Aston, 2010, 502).

---

[19] Il corpus è interrogabile da: [https://app.sketchengine.eu/#dashboard?corpname=preloaded%2Feurlex\\_ita](https://app.sketchengine.eu/#dashboard?corpname=preloaded%2Feurlex_ita) (consultato il 18/6/2018).

[20] Il corpus è interrogabile da: <https://www.sketchengine.eu/eurlex-judgments-corpus/> (consultato il 20/6/2018).

[21] Il corpus è interrogabile da: <https://www.sketchengine.eu/dgt-translation-memory/> (consultato il 21/6/2018).

[22] Per più informazioni si veda: <https://www.sketchengine.eu/> (consultato il 21/6/2018).

### 3.2. Memorie di traduzione

La memoria di traduzione (ing. *Translation Memory* - TM) è una base di dati in cui vengono salvati i segmenti allineati del testo originale e del testo tradotto per poi venir riutilizzati per delle nuove traduzioni (cfr. Arduini e Stecconi, 2007, 173; Azzano, 2008, 104; e altri). Di solito, i segmenti, denominati unità di traduzione, sono composti da una frase intera, ma possono essere composti anche da frammenti più brevi, ad esempio da un nome, da un titolo e via dicendo. Azzano (2008, 104) aggiunge che ciascuna unità di traduzione è composta “da un segmento di partenza (SP) e un segmento di arrivo (SA)” e da eventuali altre informazioni quali data di creazione, nome del traduttore, ultima modifica, lingue, argomento del testo, ecc. La memoria di traduzione si può ottenere in due modi: il primo modo è con l'utilizzo di corpora paralleli esistenti e del processo di allineamento, che produce coppie di segmenti delle traduzioni già esistenti e i loro testi originali. Nel secondo caso, il traduttore durante il processo di traduzione crea da solo le memorie di traduzione. Infatti, durante il processo, ogni segmento della traduzione viene mandato nella memoria di traduzione dove viene salvato assieme al segmento equivalente del testo originale, formando in questo modo nuove unità di traduzione (cfr. Pavlović, 2015, 290). Quando si vuole tradurre un secondo documento, il programma divide il nuovo testo in segmenti e li confronta uno ad uno con i segmenti salvati precedentemente o esistenti nella memoria di traduzione (Arduini e Stecconi, 2007, 175). Se il programma trova un'unità di traduzione uguale o simile, “il programma richiama automaticamente la traduzione già presente nella base di dati”, spiegano gli autori, e “a questo punto il traduttore può accettare la vecchia soluzione così com'è”, accettarla parzialmente e modificarla, o rifiutarla completamente e scriverne una nuova. La percentuale di somiglianza necessaria per richiamare un'unità di traduzione non deve essere sempre la stessa, ma si può impostare per ogni nuovo progetto di traduzione o testo, anche se di solito si tratta del 75%, precisa Pavlović (2015, 293).

Le memorie di traduzione, che sono particolarmente adeguate e utili nella traduzione di testi tecnici e ripetitivi, fanno aumentare la produttività dei traduttori e garantiscono una maggiore coerenza nell'uso

della terminologia, “assicurando in questo modo una maggiore qualità del prodotto finale” (cfr. Monti, 2014, 59). Ovviamente, più ampie sono le memorie di traduzione, migliori sono i risultati. Una delle memorie di grandi dimensioni contenente traduzioni di buona qualità e resa disponibile gratuitamente on-line e quindi accessibile a tutti i traduttori è la memoria di traduzione della DGT<sup>[23]</sup> della Commissione europea. Si tratta di una memoria di traduzione che comprende testi in ventiquattro lingue, inclusi il croato e l’italiano, segmentati in unità di traduzione (Steinberger et al., 2012; 2014).

### 3.3. Strumenti per la gestione della terminologia

Gli strumenti per la gestione della terminologia servono ad archiviare dati in una base terminologica che poi verranno riutilizzati (Pavlović, 2015, 294), facilitando in questo modo il lavoro dei traduttori e garantendo la coerenza terminologica, in particolar modo, quando si tratta di grandi progetti di traduzione o della traduzione in gruppo.

Le banche dati terminologiche sono spesso parte integrante delle memorie di traduzione nei CAT *tools*. Infatti, quando il traduttore incontra nel testo un segmento che equivale a un termine già esistente nella banca dati terminologica attivata, il sistema gli offre l’equivalente nella lingua d’arrivo. Il traduttore può accettarlo e continuare la traduzione senza dover fare una ricerca terminologica che spesso richiede molto tempo (cfr. Pavlović, 2015, 296).

Le banche dati terminologiche hanno sostituito i dizionari di una volta, “risolvendo molte difficoltà nella traduzione settoriale, specialmente quando si tratta di banche dati continuamente aggiornate e accessibili on line” (Diadori, 2012, 333). Oltre a quelle commerciali, ci sono delle banche dati gratuite che sono accessibili in Rete e i traduttori possono consultarle facilmente e velocemente, fatto molto importante se si pensa che i traduttori hanno spesso delle scadenze molto brevi per la consegna delle traduzioni.

*IATE (Inter-Active Terminology for Europe)*, è la banca dati terminologica interistituzionale dell’UE, in uso nelle istituzioni e agenzie dell’UE dal 2004 per la raccolta, la diffusione e la gestione condivisa della

[23] Per più informazioni si veda: <https://data.europa.eu/euodp/it/data/dataset/dgt-translation-memory> (consultato il 15/7/2018)

terminologia specifica dell'UE<sup>[24]</sup>. L'inizio di questo progetto terminologico risale al 1999 e aveva come obiettivo la formazione di un'infrastruttura in Rete per dare maggior disponibilità e accessibilità alle informazioni terminologiche. L'ultima versione di IATE, lanciata nel novembre del 2018, incorpora tutte le banche dati terminologiche esistenti dei servizi di traduzione dell'UE in un'unica banca dati interistituzionale nuova, altamente interattiva e moderna, e accessibile a tutti<sup>[25]</sup>.

Tra le banche dati terminologiche croate va menzionata *Struna*. Si tratta di un grande progetto coordinato dall'*Istituto di lingua e linguistica croata* che ha come obiettivo la raccolta sistematica, la normizzazione, la formazione e l'elaborazione della terminologia in lingua croata<sup>[26]</sup>. Tutti i risultati del progetto vengono inseriti nella banca dati e sono accessibili on-line. Si tratta di una banca dati terminologica che attualmente contiene la terminologia di più di trenta settori diversi, di cui una ventina è consultabile in Rete.

Un sito utile dove è possibile trovare moltissime informazioni che riguardano la terminologia croata è il *Hrvatski terminološki portal*<sup>[27]</sup>. Lo scopo principale del sito è quello di raccogliere la terminologia in un unico posto e renderla accessibile agli utenti grazie a un semplice motore di ricerca.<sup>[28]</sup>

Una delle banche dati che include anche la terminologia italiana è *bistro*<sup>[29]</sup>. Si tratta di una banca dati sviluppata da *EURAC Research* una decina di anni fa e contenente la terminologia legale in italiano e tedesco.

Un'altra banca dati terminologica interessante per le traduzioni in italiano è *TERMDAT*<sup>[30]</sup>. Si tratta di una banca dati multilingue in lingua italiana, francese, tedesco-svizzera, inglese e romanico che contiene termini dell'ambito legale, dello sport, dell'educazione, della medicina, della politica e di altri settori.

[24] Preso da: [http://iate.europa.eu/about\\_IATE.html](http://iate.europa.eu/about_IATE.html) (consultato il 15/11/2018).

[25] Per maggiori informazioni si veda: [http://iate.europa.eu/about\\_IATE.html](http://iate.europa.eu/about_IATE.html) (consultato il 15/11/2018).

[26] Per maggiori informazioni si veda: <http://struna.ihj.hr/page/o-struni/> (consultato il 18/6/2018).

[27] Disponibile su: <http://nazivlje.hr/stranica/o-portalu/2/> (consultato il 09/6/2018).

[28] Per maggiori informazioni si veda: <http://nazivlje.hr/stranica/o-portalu/2/> (consultato il 12/6/2018).

[29] Per maggiori informazioni si veda: <http://bistro.eurac.edu/it/informazioni-su-bistro/#bancadati> (consultato il 19/7/2018).

[30] Per ulteriori informazioni si veda: <https://www.bk.admin.ch/bk/it/home/documentazione/lingue/termdat.html> (consultato il 19/7/2018).

Alla fine bisogna menzionare anche *TERMit*<sup>[31]</sup>, una banca dati multilingue che si sta sviluppando presso l'Università di Trieste ed è configurata come un insieme di raccolte terminografiche plurilingui.

#### **4. Altre risorse e fonti utili per le traduzioni dal croato all'italiano e viceversa**

Oltre agli strumenti e alle risorse finora presentate, esistono altre risorse che devono venir menzionate, perché aiutano i traduttori a migliorare la qualità delle traduzioni. Si tratta non solo di varie risorse linguistiche (come, ad esempio, vari dizionari on-line, glossari, tesauri, ecc.), ma anche di fonti informative (vari siti, linee guida, manuali, ecc.). Si ricorda, infatti, che le traduzioni prodotte devono essere di alta qualità, dal punto di vista linguistico, che include sia la grammatica, il lessico (terminologia), ma anche lo stile.

Proprio per questo motivo i siti che offrono servizi di consulenza linguistica sono particolarmente interessanti per i traduttori. Sia *L'Accademia della Crusca*<sup>[32]</sup> per l'italiano sia *l'Institut za hrvatski jezik i jezikoslovlje*<sup>[33]</sup> per il croato mettono a disposizione un servizio di consulenza rivolto a tutti coloro che cercano informazioni e chiarimenti grammaticali e lessicali per le rispettive lingue.

Molte informazioni utili per la lingua croata si trovano sul sito dell'*Istituto per la lingua e linguistica croata*<sup>[34]</sup> dove è possibile consultare l'ortografia croata, la grammatica croata, il corpus diacronico per la lingua croata e trovare consigli linguistici di vario genere. La maggior parte delle risorse dell'Istituto esiste in forma digitale ed è accessibile on-line, consentendo così un accesso immediato alle informazioni e una soluzione veloce dei dubbi linguistici dei traduttori. Tra le risorse disponibili, quella che va senz'altro nominata è il *Hrvatski pravopis*<sup>[35]</sup>, molto utile per risolvere i dubbi che riguardano l'ortografia croata.

---

[31] Per più informazioni si veda: <http://lextrain.units.it/?q=termit> (consultato il 20/7/2018).

[32] Per *L'Accademia della Crusca* si veda: <http://www.accademiadellacrusca.it/it/lingua-italiana/consulenza-linguistica> (consultato il 18/6/2018).

[33] Per *l'Istituto per la lingua e linguistica croata* si veda: <http://ihj.hr/stranica/jezicni-savjeti/27/> (consultato il 18/6/2018).

[34] Disponibile su: <http://ihj.hr/stranica/mrezn-resursi-instituta/28/> (consultato il 09/6/2018).

[35] Disponibile su: <http://pravopis.hr/> (consultato il 09/6/2018).



Un altro sito molto utile, questa volta per la lingua italiana, è quello dell'*Associazione Italiana per la Terminologia* (ASS.I.TERM) che esiste da più di trent'anni. Quest'associazione, tra l'altro, promuove un programma per la gestione e raccolta di termini tecnico scientifici della lingua italiana<sup>[36]</sup> e contiene informazioni utili sulla terminologia.

Le *Direzioni generali della traduzione* (DGT) presso la Commissione europea offrono informazioni utili sulle risorse per i traduttori in generale, ma anche per ogni lingua a parte<sup>[37]</sup>. Va nominato *EuroVoc*<sup>[38]</sup> un tesoro multilingue e pluridisciplinare in ventitré lingue, tra cui l'italiano e il croato, comprendente la terminologia dei settori d'attività dell'Unione Europea. Questo tesoro viene utilizzato dalle istituzioni dell'Unione Europea, ma anche da amministrazioni nazionali e utenti privati, inclusi i traduttori. Da menzionare è anche *Eur-LEX*<sup>[39]</sup>, un portale che contiene informazioni legali e permette l'accesso al diritto dell'UE nonché interessante per i traduttori, anche perché offre una consultazione parallela multilingue di uno stesso testo.

Tra le risorse che si trovano on-line ad accesso gratuito, bisogna ricordare *BabelNet*<sup>[40]</sup>, il più grande dizionario multilingue on-line. Non si tratta solo di un dizionario multilingue, ma anche di una rete semantica e banca dati enciclopedica che offre informazioni di tipo enciclopedico, lessicografico, semantico e altre informazioni ricavate da varie risorse come ad esempio *WordNet*, *Wikipedia*, *OmegaWiki*, *Wiktionary*, *Wikidata*, *Wikiquote*, *VerbNet*, *Microsoft Terminology*, *GeoNames* e molte altre. *BabelNet* si aggiorna di continuo da tutte le sue risorse e rappresenta una fonte molto preziosa anche per i traduttori. Per maggiori informazioni si rimanda a Navigli e Ponzetto (2012).

Tra le risorse incorporate in *BabelNet* va nominato in particolare *Open Multilingual WordNet*, una collezione di *WordNet* in diverse lingue, inclusi anche il *WordNet croato*<sup>[41]</sup> e *l'ItalWordNet*<sup>[42]</sup>. Dato che si tratta di

---

[36] Più informazioni su ASS.I.TERM sono disponibili su questo sito: <http://www.assiterm91.it/>. (consultato il 22/8/2018).

[37] Ad esempio: [https://ec.europa.eu/info/resources-partners/translation-and-drafting-resources/guidelines-translation-contractors/guidelines-contractors-translating-croatian\\_hr](https://ec.europa.eu/info/resources-partners/translation-and-drafting-resources/guidelines-translation-contractors/guidelines-contractors-translating-croatian_hr) (consultato il 20/7/2018).

[38] Disponibile su: <http://eurovoc.europa.eu/drupal/?q=it/node> (consultato il 21/10/2018).

[39] Per più informazioni consultare: <https://eur-lex.europa.eu/homepage.html?locale=it> (consultato il 20/7/2018).

[40] Accessibile da: [www.babelnet.org](http://www.babelnet.org) (consultato il 22/10/2018).

[41] Per maggiori informazioni si veda Raffaelli et al. (2008). Disponibile su: <http://compling.hss.ntu.edu.sg/omw/> (consultato il 20/10/2018).

[42] Per ulteriori informazioni si rimanda al sito: <http://www.ilc.cnr.it/it/content/italwordnet> (consultato il

database semantico-lessicali che organizzano, definiscono e descrivono i concetti espressi dai vocaboli, i traduttori possono trovare informazioni utili che possono aiutarli nella scelta del vocabolo da usare.

Bisogna menzionare anche *Glosbe*<sup>[43]</sup>, uno dei tanti dizionari multilingui on-line, che è però molto utile anche per le traduzioni dall'italiano al croato e viceversa, poiché include oltre a numerosissimi dizionari anche delle memorie di traduzione contenenti le due lingue in questione.

Oltre alle risorse nominate, è importante dare spazio ad alcune fonti d'informazione utili ai traduttori.

Esistono moltissimi siti che offrono elenchi di collegamenti a risorse terminologiche, opere enciclopediche, linguistiche, guide di traduzione, manuali di stile, motori di ricerca, ecc., come ad esempio il sito *Term-minator*<sup>[44]</sup> e la *Sitoteca a cura di EDIGEO realizzazioni editoriali*<sup>[45]</sup>.

Una fonte d'informazione interessante, che riguarda solo la lingua italiana, è la *Rete REI - Rete per l'eccellenza dell'italiano istituzionale*, creata nel 2005 su iniziativa dei traduttori italiani della Commissione europea ed altri esperti. Alla *Rete REI* danno il loro contributo, con le loro risorse linguistiche, molte delle maggiori istituzioni delle Comunità europee, come pure esponenti della pubblica amministrazione, delle università e dei centri di ricerca dell'Italia, della Svizzera, della Croazia e della Slovenia. All'interno della *Rete REI* esiste un gruppo di lavoro, *l'Osservatorio dell'italiano istituzionale fuori d'Italia* (OIIFI)<sup>[46]</sup> che, tra l'altro, serve come punto di riferimento per dubbi, difficoltà, soluzioni terminologiche o linguistiche che riguardano le traduzioni in lingua italiana sul territorio croato.

Tanti consigli pratici e informazioni utili riguardanti, tra l'altro, anche le risorse e gli strumenti di traduzione, possono venir trovati sui siti delle varie associazioni di traduttori italiani e croati come pure sui siti delle grandi agenzie di traduzione e altri siti specializzati. Le novità riguardanti le nuove tecnologie vengono presentate regolarmente anche in varie

---

20/10/2018).

[43] *Glosbe* si può consultare su: <https://hr.glosbe.com/hr/it> (consultato il 22/10/2018).

[44] Per maggiori informazioni consultare: [http://www.term-minator.it/corso/mod2\\_4.html](http://www.term-minator.it/corso/mod2_4.html) (consultato 19/6/2018).

[45] Per maggiori informazioni consultare: <http://www.edigeo.it/Sitoteca/sitoteca.php> (consultato 19/6/2018).

[46] Per maggiori informazioni consultare: <https://www.bk.admin.ch/bk/it/home/documentazione/lingue/conferenze-e-seminari/rete-per-l-eccellenza-dell-italiano-istituzionale.html> (consultato 19/6/2018).

conferenze specializzate<sup>[47]</sup>, tra cui vanno nominati il *Translating Europe Forum* e i vari *Translating Europe Workshops*, organizzati gratuitamente dalla Commissione europea.

## Conclusioni

Questo contributo, oltre a offrire una panoramica generale delle tecnologie di traduzione, presenta, in modo più dettagliato, quegli strumenti di traduzione e quelle risorse linguistiche che possono essere utilizzate nelle traduzioni dall'italiano al croato e viceversa. Sono presentati, in particolar modo, i corpora monolingui della lingua croata e italiana, i corpora paralleli, le memorie di traduzione e le banche dati terminologiche, che includono le due lingue in questione. Inoltre, sono state presentate anche delle risorse e delle fonti che, offrendo informazioni, consigli e servizi di consulenza linguistica, aiutano il traduttore nella produzione di testi scritti di alta qualità, non solo per quanto riguarda la terminologia, ma anche per quanto riguarda lo stile e la lingua, croata e italiana, in generale.

Con la presente rassegna degli strumenti e delle risorse linguistiche che hanno a disposizione i traduttori d'oggi si sono volute sottolineare l'importanza e l'indispensabilità dell'impiego delle tecnologie della traduzione nel lavoro del traduttore. Tecnologie che gli permettono di soddisfare le esigenze e gli alti criteri di qualità, precisione e velocità richiesti dal mercato globale nell'ambito della traduzione. Data la grande velocità con cui si sviluppano le tecnologie di traduzione, è necessario che i traduttori siano sempre aggiornati. Si spera, pertanto, che questo contributo possa servire ai (futuri) traduttori come punto di partenza nel loro continuo processo di aggiornamento.

---

[47] Altre conferenze interessanti per le tecnologie della traduzione sono ad esempio: *ProZ.com conference*, *TeTra*, *The International Translation Technology Terminology Conference (TTT)* e molte altre.

## Bibliografia

Arduini, Stefano, Stecconi, Ubaldo (2007). *Manuale di traduzione. Teorie e figure professionali*. Milano: Carocci editore.

Azzano, Dino (2008). La traduzione assistita: un'introduzione. *Rivista Internazionale di Tecnica della Traduzione = International Journal of Translation*, 10, 101–122. Disponibile su: <https://www.openstarts.units.it/handle/10077/3379> (consultato il 22/6/2018).

Baisa, Vít, Michelfeit, Jan, Medved, Marek, Jakubiček, Miloš (2016). European Union Language Resources in Sketch Engine. In *The Proceedings of tenth International Conference on Language Resources and Evaluation (LREC'16)*. European Language Resources Association (ELRA). Portorož, Slovenia.

Baroni, Marco (2010). *Corpora di italiano*. In Raffaele Simone (a cura di), *Enciclopedia dell'italiano*, vol. 1, (303–303). Istituto della Enciclopedia italiana: Roma. Disponibile su: [http://www.treccani.it/enciclopedia/corpora-di-italiano\\_\(Enciclopedia-dell'Italiano\)/](http://www.treccani.it/enciclopedia/corpora-di-italiano_(Enciclopedia-dell'Italiano)/) (consultato il 14/7/2018).

Bowker, Lynne, Pearson, Jennifer (2002). *Working with specialized language: a practical guide to using corpora*. London: Routledge.

Brusasco, Paola (2018). La traduzione automatica. *Tradurre-pratiche teorie strumenti*, 14. Disponibile su: <https://rivistatradurre.it/2018/05/la-traduzione-automatica/> (consultato il 12/7/2018).

Burchardt, Aljoscha, Macketanz, Vivien, Dehdari, Jon, Heigold, Georg, Jan-Thorsten, Peter, Williams, Philip (2017). A Linguistic Evaluation of Rule-Based, Phrase-Based, and Neural MT Engines. *The Prague Bulletin of Mathematical Linguistics*, 108, 159–170. Disponibile su: <https://ufal.mff.cuni.cz/pbml/108/art-burchardt-macketanz-dehdari-heigold-peter-williams.pdf> (consultato il 15/10/2018).

Chiari, Isabella (2007). *Introduzione alla linguistica computazionale*. Bari: Laterza.

Diadori, Pierangela (2012). *Teoria e tecnica della traduzione*. Milano: Mondadori Education.

European Commission (2014). *MT@EC. Secure Machine Translation for the*

*European Union*. Luxemburg: Publication Office of the European Union. DOI: 10.2799/23526.

European Commission (2016). *Translation tools and workflow*. Luxemburg: Publication Office of the European Union. DOI: 10.2782/703257  
Disponibile su: <https://publications.europa.eu/en/publication-detail/-/publication/00e51a8e-9c50-11e6-868c-01aa75ed71a1/language-en> (consultato il 12/11/2018).

Forcada, Mikel L. (2017). Making sense of neural machine translation. *Translation Spaces*, 6:2, 291–309. Disponibile su: <http://www.dlsi.ua.es/~mlf/docum/forcada17j2.pdf> (consultato il 15/10/2018).

Hu, Ke, Cadwell, Patrick (2016). A Comparative Study of Post-editing Guidelines. *Baltic Journal of Modern Computing*, 4, 2, 342–353. DOI: 10.13140/RG.2.1.2253.1446.

Hutchins, John W., Somers, Harold L. (1992). *An introduction to machine translation*. London: Academic Press.

Kenning, Marie-Madeleine (2010). What are parallel and comparable corpora and how can we use them? In Annie O’Keeffe & Michael McCarthy (a cura di), *The Routledge handbook of corpus linguistics*, (487–500). London: Routledge.

Koehn, Philipp (2002). *Europarl: A Multilingual Corpus for Evaluation of Machine Translation*. Marina del Rey: University of Southern California.

Koehn, Philipp (2010). *Statistical Machine Translation*. Cambridge: Cambridge University Press.

Koehn, Philipp et al. (2007). Moses: Open source toolkit for statistical machine translation. In *Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics Companion Volume Proceedings of the Demo and Poster Sessions*, (177–180). Prague: Association for Computational Linguistics. Disponibile su: <http://www.aclweb.org/anthology/P/P07/P07-2045> (consultato il 19/6/2018).

Koester, Almut (2010). Building small specialised corpora. In Annie O’Keeffe & Michael McCarthy (a cura di), *The Routledge handbook of corpus linguistics*, (66–79). London: Routledge.

Kübler, Natalie, Aston, Guy (2010). Using corpora in translation. In Annie O’Keeffe & Michael McCarthy (a cura di), *The Routledge handbook of corpus linguistics*, (501–515). London: Routledge.

Lalli Pačelat, Ivana (2016). *I corpora nell’apprendimento e nell’insegnamento delle lingue*. In *L’apprendimento delle lingue straniere*. Bilingualism Matters@Rijeka. Disponibile su: <http://www.bilingualism-matters-rijeka.ffri.hr/it/bilinguismo/l-apprendimento-delle-linguestraniere.html?tmpl=component&print=1&page=> (consultato il 19/6/2018).

Ljubešić, Nikola, Erjavec, Tomaž (2011). hrWaC and slWaC: Compiling Web Corpora for Croatian and Slovene. In Ivan Habernal & Vaclav Matousek (a cura di), *Text, Speech and Dialogue / Proceedings of the 14th International Conference* (TSD 2011), LNCS 6836, (395–402). Heidelberg/Berlin: Springer Verlag.

Massardo, Isabella, van der Meer, Jaap, O’Brien, Sharon, Hollowood, Fred, Aranberri, Nora, Drescher, Katrin (2016). *MT Post-editing Guidelines*. Amsterdam: TAUS Signature Editions. Scaricabile da: <https://info.taus.net/post-editing-guidelines?submissionGuid=75f4cbeb-ec9c-44cb-8214-d3ac0adbbebe> (consultato il 20/11/ 2018).

McEnery, Tony, Xiao, Richard (2007). Parallel and comparable corpora: What are they up to? In Gunilla Andermann & Margaret Rogers (a cura di), *Incorporating Corpora: Translation and the Linguist (Translating Europe)*. Clevedon, UK: Multilingual Matters.

McEnery, Tony, Hardie, Andrew (2012). *Corpus linguistics: method, theory and practice*. Cambridge: Cambridge University Press.

Monti, Johanna (2014). Le tecnologie per la traduzione nell’era del cloud computing: stato dell’arte e prospettive future. *Testi e linguaggi*, 8, 55–66. Disponibile su: [https://www.academia.edu/27945846/Le\\_tecnologie\\_per\\_la\\_traduzione\\_nellera\\_del\\_cloud\\_computing\\_stato\\_dellarte\\_e\\_prospettive\\_future](https://www.academia.edu/27945846/Le_tecnologie_per_la_traduzione_nellera_del_cloud_computing_stato_dellarte_e_prospettive_future) (consultato il 16/7/2018).

Navigli, Roberto, Ponzetto, Simone Paolo (2012). BabelNet: The Automatic Construction, Evaluation and Application of a Wide-Coverage Multilingual Semantic Network. *Artificial Intelligence*, 193, 217–250.

O'Brien, Sharon (2010). Introduction to Post-Editing: Who, What, How and Where to Next?. *AMTA*. Disponibile su: <http://amta2010.amtaweb.org/AMTA/papers/6-01-ObrienPostEdit.pdf> (consultato il 15/10/2018).

Pavlović, Nataša (2015). *Uvod u teoriju prevođenja*. Zagreb: Leykam International.

Raffaelli, Ida, Tadić, Marko, Bekavac, Božo, Agić, Željko (2008). Building Croatian WordNet. Atila Tanacs et al. (a cura di), *Proceedings of the Fourth Global WordNet Conference*, (349—359). Szeged: Global Global WordNet Association.

Rossini Favretti, Rema, Tamburini, Fabio, De Santis, Cristiana (2002). CORIS/CODIS: A corpus of written Italian based on a defined and a dynamic model. In Andrew Wilson, Paul Rayson, & Tony McEnery (a cura di), *A Rainbow of Corpora: Corpus Linguistics and the Languages of the World*, (27–38). Munich: Lincom–Europa.

Sandrini, Peter (2012). Tecnologia FLOSS per la traduzione: Disponibilità, applicazione e problematiche. *inTRAlinea. Special Issue: Specialized Translation II*. Innsbruck: University of Innsbruck. Disponibile su: [https://www.academia.edu/3764221/Tecnologia\\_FLOSS\\_per\\_la\\_traduzione.\\_Disponibilit%C3%A0\\_applicazione\\_e\\_problematiche](https://www.academia.edu/3764221/Tecnologia_FLOSS_per_la_traduzione._Disponibilit%C3%A0_applicazione_e_problematiche) (consultato il 15/7/2018).

Scarpa, Federica (2008). *La traduzione specializzata. Lingue speciali e mediazione linguistica*. Milano: Editore Ulrico Hoepli.

Spina, Stefania (2001). *Fare i conti con le parole*. Perugia: Guerra edizioni.

Steinberger, Ralf, Eisele, Andreas, Klocek, Szymon, Pilos, Spyridon, Schlüter, Patrick (2012). DGT-TM: A freely Available Translation Memory in 22 Languages. *Proceedings of the 8th international conference on Language Resources and Evaluation (LREC'2012)*, (21–27). Istanbul.

Steinberger, Ralf, Ebrahim, Mohamed, Poulis, Alexandros, Carrasco-Benitez, Manuel, Schlüter, Patrick, Marek, Przybyszewski, Patrick, Gilbro, Signe (2014). An overview of the European Union's highly multilingual parallel corpora. *Language Resources and Evaluation Journal (LREC)*, 48, 4, 679–707. DOI: 10.1007/s10579-014-9277-0.

Steinberger, Ralf, Pouliquen, Bruno, Widiger, Anna, Ignat, Camelia, Erjavec, Tomaž, Tufis, Dan, Varga, Daniel (2006). The JRC-Acquis: A multilingual aligned parallel corpus with 20+ languages. In *Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC 2006)*. Genova: European Language Resources Association (ELRA).

Tadić, Marko (2009). New version of the Croatian National Corpus. In Dana Hlaváčková et al. (a cura di), *After Half a Century of Slavonic Natural Language Processing*, (199–205). Brno: Masaryk University.

TAUS (2010). *MT Post-editing Guidelines*. Disponibile su: <https://www.taus.net/academy/best-practices/postedit-best-practices/machine-translation-post-editing-guidelines> (consultato il 15/10/2018).

Tiedemann, Jörg (2012). Parallel Data, Tools and Interfaces in OPUS. In *Proceedings of the 8th International Conference on Language Resources and Evaluation (LREC'2012)*. European Language Resources Association (ELRA).

Tufiş, Dan, Radu, Ion (2007). Parallel Corpora, Alignment Technologies and Further Prospects in Multilingual Resources and Technology Infrastructure. In Corneliu Burileanu & Horia Nicolai L. Teodorescu (a cura di), *Advances in Spoken Language Technology: Proceedings of the 4th International Conference on Speech and Dialogue Systems* (183–195). Bucharest: Romanian Academy Publishing House.

Vasiļjevs, Andrejs, Skadiņš, Raivis, Tiedemann, Jörg (2011). LetsMT!: cloud-based platform for building user tailored machine translation engines. In *Proceedings of the Thirteenth Machine Translation Summit (MT Summit XIII)* (507–511). Xiamen, China.

Zanettin, Federico (2014). Corpora in Translation. In Julianne House (a cura di), *Translation: A Multidisciplinary Approach. Palgrave Advances in Language and Linguistics*, (178–199). London: Palgrave Macmillan. DOI: 10.1057/9781137025487\_10.



## **TRANSLATION TOOLS AND LINGUISTIC RESOURCES USED IN THE TRANSLATION FROM CROATIAN TO ITALIAN AND FROM ITALIAN TO CROATIAN**

Translation tools and linguistic resources are constantly developed and translators have to keep up with the updates in the field of translation technologies in order to meet the market needs. The aim of this paper is to present the linguistic resources and the translation tools used nowadays by translators. Special attention is given to the linguistic resources and the translation tools used in the translation from Croatian to Italian and from Italian to Croatian.

Besides a general presentation of the linguistic resources and the translation tools used by translators at present, this paper offers a detailed presentation of all elements, tools and resources that are part of the translator's workbench, more precisely, machine translation, translation memory, terminology management tools and corpora.

This brief review of the translation tools and the linguistic resources available to the modern translators emphasizes the importance and the indispensability of the translation technology in the translator's everyday work. They allow translators to respond to requests of high quality, precision and speed imposed by the global translation market.

Keywords: CAT tools, machine translation, corpora, termbase, translation memory

## **UPORABA PRIJEVODNIH TEHNOLOGIJA U PREVOĐENJU S HRVATSKOGA NA TALIJANSKI JEZIK I S TALIJANSKOGA NA HRVATSKI JEZIK**

Da bi mogli zadovoljiti zahtjeve globalnoga tržišta i pratiti trendove na području prevođenja, prevoditelji moraju poznavati novosti i promjene u jezičnoj industriji.

Cilj je ovoga rada predstaviti prijevodne tehnologije odnosno alate i jezične resurse koji današnjim prevoditeljima stoje na raspolaganju, a poseban je naglasak na alatima i resursima koji su korisni za prijevode s hrvatskoga na

talijanski te s talijanskoga na hrvatski jezik. Detaljno su opisani svi resursi i alati koji su dio prevoditeljeve radionice, a posebna je pozornost posvećena strojnome prevođenju, prijevodnim memorijama, alatima za upravljanje terminologijom i korpusima.

Ovim se radom i prikazom alata i resursa kojima raspolažu današnji prevoditelji želi naglasiti važnost i nužnost uporabe tehnologije u svakodnevnome radu prevoditelja, a sve kako bi se moglo odgovoriti na zahtjeve visoke kvalitete, preciznosti i brzine koje postavlja globalno tržište prijevoda.

Ključne riječi: prevođenje, CAT alati, strojno prevođenje, strojno potpomognuto prevođenje, korpus, terminološka baza podataka, prijevodna memorija