

Francesca Biagini, Tatsiana Maiko
Valentina Noseda

Il corpus parallelo italiano-russo del *Nacional'nyj Korpus russkogo jazyka* (НКРЈа). Evoluzione e applicazioni nelle ricerche della slavistica italiana

1. *Introduzione*¹

Scopo dell'articolo è offrire una rassegna degli studi condotti in Italia con l'ausilio del Corpus parallelo italiano-russo (CP) del *Nacional'nyj Korpus russkogo jazyka* (НКРЈа)², nato nel 2013 come piccolo corpus pilota e successivamente ampliato grazie alla collaborazione fra l'Istituto di Lingua Russa di Mosca (*Institut Russkogo Jazyka imeni V.V. Vinogradova*), la Cattedra di lingua e letteratura russa dell'Università Cattolica del Sacro Cuore di Milano e il Dipartimento di Interpretazione e Traduzione dell'Università di Bologna, Campus di Forlì³.

Ad oggi la disponibilità di corpora liberamente accessibili e supportati da motori di ricerca più o meno sofisticati è vastissima, e fra gli svariati tipi di corpora, una posizione di grande importanza è occupata proprio dai corpora paralleli (Borin 1999), utili per l'analisi linguistica, la didattica delle lingue (Sinclair 2004) e negli studi traduttologici (Baker 1993, Granger 2003). Tuttavia, attualmente, sono pochissimi i corpora paralleli per la coppia russo-italiano: oltre al CP, esistono alcune collezioni di testi paralleli in queste due lingue – accessibili da portali quali Opus⁴ e InterCorp⁵ o dal software per la gestione di corpora

¹ Le autrici hanno lavorato in stretta collaborazione. A fini accademici, i §§ 1, 2, 2.1 e 3 sono da attribuire a Valentina Noseda, i §§ 2.2 e 2.6 sono da attribuire a Tatsiana Maiko, mentre i §§ 2.3, 2.4 e 2.5 sono da attribuire a Francesca Biagini.

² <www.ruscorpora.ru> (ultimo accesso: 09.02.24).

³ Allo stato attuale (gennaio 2024) il CP conta quasi 5 milioni di *token* (4.930.970) e comprende 126 testi così suddivisi: 74 in direzione russo-italiano e 48 in direzione italiano-russo. Come in tutti i corpora paralleli, la prosa letteraria è il genere testuale prevalente (con opere che coprono un periodo che va dal XVIII al XXI secolo). Si contano tuttavia anche 12 opere di saggistica e diversi articoli di giornale. Infine, il corpus è annotato secondo i criteri del НКРЈа e presenta dunque tre tipi di *tagging*: metatestuale, morfologico (per entrambe le lingue) e semantico (solo per il russo). Si rimanda a (Biagini et al. 2019, Noseda 2017, 2018) per un resoconto dettagliato sulle caratteristiche e i criteri di compilazione.

⁴ <<http://opus.lingfil.uu.se/>> (ultimo accesso: 24.01.24).

⁵ <https://kontext.korpus.cz/first_form> (ultimo accesso: 24.01.24).

Sketch Engine⁶ – le cui caratteristiche, nondimeno, si allontanano da quelle dei corpora più tradizionali, rendendo queste risorse spesso inadatte alla ricerca linguistica⁷.

Da qui muove l'esigenza di proporre un resoconto ragionato degli studi condotti con l'ausilio del CP fino ad oggi, valutandone potenzialità e limiti in vista di un suo futuro ampliamento e miglioramento.

2. *Il CP ita-ru nelle ricerche della Slavistica italiana*

Nei paragrafi a seguire presenteremo i risultati della ricerca condotta con il corpus parallelo negli ultimi nove anni suddivisi per categoria: gli studi sul verbo (§ 2.1), gli studi su lessico e fraseologia (§ 2.2), gli studi sui segnali discorsivi (§ 2.3), gli studi sulle relazioni transfrastiche (§ 2.4), quelli sulla traduzione (§ 2.5) e, infine, quelli sull'uso del CP nella didattica (§ 2.6)⁸.

2.1. *Gli studi sul verbo*

I contributi relativi al verbo possono essere suddivisi in tre sotto-gruppi:

- studi relativi a prefissi e circonfissi verbali (Biagini 2020, Noseda 2015, Ruvoletto 2020, Slavkova 2019);
- studi sui verbi causativi (Noseda 2019a, 2019b);
- studi relativi all'aspetto del verbo e, più in generale, alle strutture tempo-ASPETTUALI (Benigni, Gebert 2019, 2023; Fici 2020; Fiči 2020; Leone 2019; Noseda 2020; Biagini 2023, Biagini, Gebert in stampa).

2.1.1. *Modalità di impiego del corpus.* In alcuni degli studi sopraelencati (Leone 2019, Fiči 2020) il corpus parallelo è usato meramente come database per attingere a esempi e non si tratta dunque di ricerche propriamente *corpus-based* o *corpus-driven*⁹. Nei restanti

⁶ <<https://www.sketchengine.co.uk/>> (ultimo accesso: 24.01.24).

⁷ Sulla distinzione tra corpora tradizionali e collezioni di testi, cfr. Noseda (2017: 24).

⁸ Precisiamo che in molti degli studi presentati in questa rassegna gli esempi del CP sono affiancati da esempi tratti da altre risorse: altre sezioni del NKRJA, il corpus dei sottotitoli "Open-Subtitles" in OPUS2, InterCorp, alcuni corpora comparabili della famiglia Aranea (accessibili da Sketch Engine), i corpora della famiglia TenTen (Sketch Engine), articoli di giornale e altre opere letterarie disponibili in rete. Per il confronto con l'italiano, in un caso è stato usato il corpus CORIS / CODIS (<https://corpora.flclit.unibo.it/coris_ita.html> [ultimo accesso: 24.01.24]).

⁹ Per ricerche *corpus-based* intendiamo tutte quelle indagini empiriche i cui risultati sono il frutto dell'analisi condotta interamente sui dati estratti da corpora a partire da un'ipotesi o da studi precedenti, che vengono in questo modo confutati, confermati o arricchiti da nuove osservazioni. Con il metodo *corpus-driven*, al contrario, i dati estratti da corpora vengono analizzati in assenza di ipotesi o teorie a priori, al fine di ottenere "new insights into language" (Granger 2003: 133). Entrambi i metodi vanno distinti dagli studi "*corpus aware*" (quali Leone 2019 e Fiči 2020), che si

casi le indagini sono basate interamente sui corpora e il CP è usato per indagare le forme italiane corrispondenti a determinate strutture linguistiche.

La maggior parte dei saggi ha come scopo l'analisi contrastiva tra russo e italiano (Nozeda 2015, 2019a; Slavkova 2019; Ruvoletto 2020; Benigni, Gebert 2019, 2023; Fici 2020; Biagini 2023; Biagini, Gebert in stampa), mentre in alcuni di essi le traduzioni italiane sono usate per approfondire lo studio di una data struttura della lingua russa (Noseda 2019b, Biagini 2020).

Otto contributi presentano un'analisi di tipo quantitativo (Benigni, Gebert 2019, 2023; Biagini 2020, 2023; Noseda 2015, 2019a, 2019b, 2020): sebbene non si tratti mai di analisi statistica, spesso resa impossibile dalla scarsa quantità di esempi raccolti, vi si propone un conteggio dei dati estratti dal corpus.

Infine, non si predilige una lingua di partenza per l'interrogazione del CP, che viene usato nella sua interezza, senza condurre ricerche separate su uno dei due sottocorpora (ru-it e it-ru).

2.1.2. Risultati delle indagini. Negli studi sopraelencati, il CP si è rivelato utile per:

- identificare le strutture preferenziali da un punto di vista contrastivo. Ad esempio, in Nozeda 2015 e Slavkova 2019 vengono individuate le strategie più ricorrenti per rendere in italiano la semantica dei prefissi verbali. Dai dati del corpus emerge come spesso venga tradotta solo la base verbale, trascurando la semantica del prefisso, e come in altri casi, invece, in traduzione venga data più importanza a quest'ultimo, tanto che diverse basi con il medesimo prefisso presentano la stessa resa. Ruvoletto (2020) verifica in che modo russo e italiano codificano la semantica dell'intensificazione e identifica, grazie al CP, una serie di strategie attraverso cui l'italiano sopperisce all'impossibilità di ricorrere alla morfologia derivazionale del sistema verbale russo. In Noseda 2019a si rileva che, nell'esprimere una situazione causativa, il russo molto spesso non marca l'opposizione tra causativo diretto e indiretto ed è per questo più ambiguo rispetto all'italiano. Fici (2020) determina le realizzazioni più frequenti in russo del piucche-perfetto italiano. Infine, in Biagini 2023 e Biagini, Gebert (in stampa) si individuano le strategie di realizzazione in russo delle perifrasi verbali *stare per* (Biagini 2023), *(in) cominciare/iniziare a, continuare a, cercare/tentare di* (Biagini, Gebert in stampa). Nel secondo contributo, in particolare, viene identificata una stretta correlazione tra le perifrasi e le forme aspettuative marcate del russo, ossia gli imperfettivi dei verbi telici e i perfettivi degli atelici.
- precisare la semantica e le funzioni delle unità linguistiche prese in esame all'interno di una sola lingua, come in Biagini (2020), che riesce a dare una caratterizzazione più dettagliata dei verbi *naest'sja, napit'sja e doždat'sja* grazie al confronto con le traduzioni italiane.

servono sporadicamente di esempi estratti da corpora con l'unico scopo di illustrare un uso, un caso specifico o una teoria (Gilquin 2010: 6).

- Colmare le lacune del materiale lessicografico con cui le occorrenze del corpus vengono messe a confronto, rintracciando nuovi significati o identificando traducanti più appropriati nel passaggio da russo a italiano (Biagini 2020, Nozeda 2015).
- Verificare i contesti in cui si realizza una data struttura: Noseda (2019b) determina, grazie ai dati del corpus, che il causativo indiretto russo, definito dall'autrice "ellittico", si realizza in qualsiasi contesto purché l'iniziatore si trovi nelle condizioni di poter affidare l'azione a un intermediario (tramite ordine o richiesta).
- Verificare la validità di un'ipotesi, come fanno Benigni e Gebert (2019, 2023) relativamente alla correlazione tra aspetto e referenzialità del nome oggetto portata all'attenzione da Leiss: in particolare, in (2019) l'ipotesi non trova conferma nei dati analizzati, mentre in (2023), entro alcuni limiti e in riferimento alla costruzione in esame, l'ipotesi può dirsi confermata.

2.2. *Gli studi su lessico e fraseologia*

Il corpus parallelo è stato usato in diversi studi che trattano il lessico e la fraseologia.

2.2.1. *Modalità di impiego del corpus.* Negli studi di questo gruppo il corpus parallelo è usato sia per approfondire un fenomeno della lingua russa (Benigni 2017, 2019), sia per condurre analisi contrastive russo-italiano (di Filippo 2020, 2023; Benacchio, Biasio 2023; Badolati 2021, 2022). Il CP viene interrogato in direzione russo-italiano (Benigni 2017, 2019) o in entrambe le direzioni (di Filippo 2020, 2023; Benacchio, Biasio 2023; Badolati 2021, 2022).

La maggioranza dei contributi presenta un'analisi qualitativa. Solo due studi (Benacchio, Biasio 2023 e di Filippo 2023) riportano dati quantitativi.

Cinque contributi su sette utilizzano, oltre al CP, altre risorse, con lo scopo di ottenere dati più rappresentativi.

2.2.2. *Risultati delle indagini.* Il CP, insieme ad altri strumenti, ha favorito le ricerche finalizzate a:

- indagare il funzionamento di un fenomeno nelle due lingue in chiave contrastiva. Ad esempio, di Filippo (2020) integra il confronto teorico dei crononimi alterati che esprimono indeterminatezza in russo e italiano con l'analisi della loro frequenza d'uso, del significato e delle equivalenze traduttive costanti. L'analisi condotta da Benacchio e Biasio (2023) per verificare l'isomorfismo distribuzionale tra *sejčas e te-per'* e le corrispettive forme italiane *ora* e *adesso* ha confermato le ipotesi proposte in letteratura (Mel'čuk 1985) e ha rivelato l'interscambiabilità tra i due deittici italiani come corrispondenti paralleli dei due deittici russi, indipendentemente dalla loro posizione sintattica o dalle loro specifiche proprietà semantiche. Badolati (2021, 2022) stabilisce convergenze e divergenze tra le unità fraseologiche di origine biblica con un componente onomastico in russo e italiano; le differenze riscontrate sono di carattere semantico (polisemia, poli-corrispondenza), stilistico e d'uso. Infine, nell'analisi

dell'uso e della combinabilità dei numeri tondi all'interno dei costrutti ottenuti per inversione sintattica in russo, di Filippo (2023) riesce a dimostrare che le inversioni approssimative mostrano stretti rapporti di solidarietà semantica con i numeri tondi, individuando le strategie preferenziali della loro resa in italiano.

- Analizzare le strategie, i meccanismi e le funzioni di alcune strutture della lingua russa: Benigni (2017) indaga l'intensificazione dei nomi non scalari esaminando le diverse strategie impiegate (morfologiche, lessicali, sintattiche) e i meccanismi semantico-cognitivi alla base dell'intensificazione, nonché le sue funzioni pragmatiche. Benigni (2019) esplora invece la trasformazione delle costruzioni con *minimizer* volgari ed eufemistici in marche a polarità negativa. I dati del russo vengono successivamente confrontati con quelli dell'italiano.
- Colmare alcune lacune del materiale lessicografico, come nel caso dell'analisi dei fraseologismi di origine biblica condotta da Badolati (2021, 2022).

2.3. *Gli studi sui segnali discorsivi*¹⁰

Il CP è stato utilizzato anche in studi sui segnali discorsivi¹¹ (di seguito SD).

2.3.1. Modalità di impiego del corpus. In questi lavori è stato studiato un SD italiano mediante la comparazione con i suoi equivalenti nei testi paralleli russi (Bonola 2017, 2020, 2022; Levontina, Denisova 2017), o viceversa (Bonola 2016; Zaliznjak *et al.* 2018; Biagini, Bonola 2019; Benigni 2020; Nosedà 2021; In'kova 2023). In Bonola, Nosedà 2020 sono riportati due *case study*, uno per ciascuna direzione (sui SD *ved'* e *allora*). Alcuni contributi analizzano un SD russo attraverso i dati provenienti anche da altre lingue oltre all'italiano: francese (In'kova 2023)¹² e francese e inglese (Zaliznjak *et al.* 2018).

Sette studi presentano un'analisi di tipo quantitativo, seppur non statisticamente significativa (Biagini, Bonola 2019; Bonola 2017, 2020, 2022; Bonola, Nosedà 2020; In'kova 2023; Nosedà 2021), in quattro lavori invece non sono riportati, se non sporadicamente, dati numerici (Benigni 2020; Bonola 2016; Levontina, Denisova 2017; Zaliznjak *et al.* 2018).

Infine, il CP è usato in entrambe le direzioni con l'eccezione di In'kova 2023 in cui l'analisi viene svolta utilizzando le traduzioni in francese e in italiano di testi di partenza russi.

¹⁰ Per completezza, in questo elenco sono inclusi alcuni contributi di autori non afferenti ad atenei italiani (Levontina, Denisova 2017; Zaliznjak *et al.* 2018; In'kova 2023).

¹¹ Con SD si fa riferimento a un insieme di parole (avverbi, connettori avverbiali, particelle) accomunate dalla possibilità di assumere in certi contesti valore testuale e discorsivo. Lo studio su *uže* nelle predicazioni copulative del tipo [*x èto uže y*] (Benigni 2020), è stato incluso in questo gruppo poiché *uže*, come molti SD, svolge più funzioni, talvolta anche nello stesso enunciato (avverbio fasale, intensificatore, funzione pragmatico-discorsiva di valutazione, uso discorsivo-procedurale).

¹² In questo contributo, in particolare, si utilizza il Supracorpora database dei connettori (In'kova 2018) (<<http://a179.frccsc.ru/RSCF10004/main.aspx>> ultimo accesso: 28.01.24).

2.3.2. *Risultati delle indagini.* Nei lavori dedicati ai SD il CP ha permesso di:

- raccogliere in breve tempo, nella fase euristica, una quantità di dati linguistici significativa per precisare o modificare la propria intuizione riguardo al quesito di ricerca (Bonola, Nosedà 2020).
- Definire i valori di un SD chiarendoli mediante la comparazione linguistica, eventualmente colmando le lacune del materiale lessicografico (tutti i contributi).
- Stabilire l'eventuale presenza di equivalenti prevalenti e, in caso di analisi quantitative, descrivere i mezzi preferenziali di ogni lingua per esprimere determinate funzioni, evidenziando le differenti strategie applicate dalle diverse lingue (tutti i contributi). In questo modo vengono forniti anche strumenti utili per una maggiore consapevolezza durante il processo traduttivo.
- Individuare nei testi tradotti alcune tendenze legate al fenomeno degli universali traduttivi (Nosedà 2021).
- Evidenziare attraverso il confronto interlinguistico i meccanismi di estensione semantica (Benigni 2020).

2.4. *Gli studi sulle relazioni transfrastiche*

Un altro proficuo ambito di applicazione del CP è costituito dallo studio delle relazioni transfrastiche¹³.

2.4.1. *Modalità di impiego del corpus.* In questi contributi vengono indagati: i mezzi di espressione di una determinata relazione transfrastica (quella ipotetica, rivolgendo l'attenzione a tempo e aspetto verbale in Biagini 2020, quella preconcessiva in Biagini, Mazzoleni 2017 e 2018 e B'jadžini 2019 e quella di concomitanza in In'kova 2020); uno specifico connettivo (*tuttavia*) e i suoi corrispettivi russi (Pecorari, Pinelli 2022) e l'ordine sequenziale di reggente e subordinata in italiano e in russo (Biagini, Mazzoleni 2023).

Tre lavori presentano un'analisi di tipo quantitativo (Biagini, Mazzoleni 2023; Pecorari, Pinelli 2022; In'kova 2020), mentre negli altri quattro non vengono forniti dati relativi alle frequenze d'uso (Biagini, Mazzoleni 2017, 2018; Biagini 2020; B'jadžini 2019).

Il CP è usato nel suo insieme (Biagini, Mazzoleni 2017, 2018; Biagini 2020; B'jadžini 2019; in In'kova 2020 attraverso il Supracorpora database dei connettori), oppure con i due sottocorpora it-ru e ru-it interrogati separatamente selezionando i testi successivi a una certa data (Pecorari, Pinelli 2022; Biagini, Mazzoleni 2023), infine ricercando solo nei testi di partenza russi (In'kova 2020).

2.4.2. *Risultati delle indagini.* Il CP si è rivelato utile per:

- tracciare un inventario, raccogliendo un ampio ventaglio di forme e tipi di costrutto esprimenti una data relazione transfrastica nelle due lingue (ad es. preconcessiva in

¹³ Relazioni concettuali che unificano almeno due processi semplici e virtualmente indipendenti in un processo complesso (Prandi *et al.* 2005: 37).

Biagini, Mazzoleni 2017, 2018 e B'jadžini 2019) ed evidenziare le principali differenze interlinguistiche.

- Indagare la frequenza d'uso dei costrutti nelle due lingue (ad es. il frequente utilizzo in russo di costrutti asindetici con il perfettivo futuro della protasi o di protasi e apodosi in riferimento a eventi presenti, futuri e passati per esprimere la relazione condizionale in russo in Biagini 2020).
- Mostrare come varia la frequenza d'uso di forme italiane equivalenti a diversi connettori in russo, ad es. *pri ètom*, *pritom* e *pričem* mostrano lo stesso ventaglio di equivalenti, ma la loro distribuzione varia in base al connettore russo (In'kova 2020).
- Mettere in evidenza come una lingua possa assegnare un maggiore rilievo semantico e testuale a una relazione transfrastica, codificandola in maniera semanticamente più ricca o trattandola come una componente fondamentale dell'architettura del testo (come nel caso di quella concessiva in russo in Pecorari, Pinelli 2022).
- Studiare l'ordine dominante di comparsa di frasi sovraordinata e subordinata nel periodo complesso nelle due lingue e la relativa funzione a livello testuale (ad es. protasi e apodosi in Biagini, Mazzoleni 2023).

2.5. *Gli studi sulla traduzione*

Negli ultimi anni l'uso del CP ha iniziato a prendere piede anche negli studi sulla traduzione.

2.5.1. *Modalità di impiego del corpus.* Nei contributi dedicati alla traduzione il CP viene utilizzato per analizzare le strategie traduttive applicate in italiano o in russo. In particolare in Artoni 2021a e 2021b vengono esaminati i risultati della pratica traduttiva alla luce delle nozioni teoriche di pragmatica interculturale, mentre in Nosedà 2019c viene sviluppato lo studio della resa di due fenomeni linguistici, la particella *togo* e l'ellissi sintattica, in 15 traduzioni italiane de *Il cappotto* di Gogol'.

I tre contributi presentano tutti analisi di tipo quantitativo. Nei due lavori di Artoni (2021a e 2021b) il corpus viene interrogato in entrambe le direzioni ma viene presentato anche un confronto tra testi tradotti e non (Artoni 2021a e 2021b), in Nosedà 2019c viene invece creato un sottocorpus costituito da un unico testo di partenza russo e dalle 15 traduzioni.

2.5.2. *Risultati delle indagini.* Grazie all'utilizzo del CP è stato possibile mostrare come determinate forme linguistiche (i diminutivi temporali *attimino*, *minutino* e *minutku*, *minutočku* in Artoni 2021a e l'imperativo in Artoni 2021b) abbiano una distribuzione differente nelle due lingue, con una maggiore frequenza in russo, e come la consapevolezza dei meccanismi pragmatici si può riflettere sulle scelte traduttive.

Il CP consente inoltre di far emergere come si differenziano le strategie traduttive a seconda che si tratti di rendere una differenza interlinguistica generale o riprodurre il procedimento artistico specifico adottato da uno scrittore (Nosedà 2019c). Nel secondo caso si

osserva una maggiore varietà nelle scelte dei traduttori e vengono evidenziati alcuni limiti dell'utilizzo del CP negli studi di linguistica qualora i testi letterari presentino forme linguistiche molto diverse da quelle standard.

2.6. *Gli studi sulla didattica*

Infine, il CP si presta a ricerche nell'ambito della didattica del russo L2.

2.6.1. Modalità di impiego del corpus. Anche negli studi dedicati alla didattica il CP è stato usato sia per approfondire dei fenomeni della lingua russa (Slavkova 2017; B'jadžini 2017; Biagini, Mazzoleni 2019; Gasanova-Mijat 2019; Nosedà 2020), sia per l'analisi contrastiva russo-italiano (Benigni, Ruvoletto 2019), ma con il fine ultimo di applicare i risultati delle indagini all'insegnamento del russo come L2, proponendo anche delle modalità di utilizzo dei dati estratti dal corpus all'interno delle classi. L'analisi condotta è prevalentemente di tipo qualitativo.

Anche in questo caso, oltre al CP, sono state integrate diverse risorse. Inoltre, in due studi i dati forniti dagli apprendenti vengono impiegati come punto di partenza per l'analisi: Benigni e Ruvoletto (2019) utilizzano le produzioni di studenti di russo italo-foni e di parlanti russo-foni che studiano l'italiano per individuare i problemi più frequenti nella codifica della deissi nelle rispettive L2; Nosedà (2020) per mezzo dei dati elicitati attraverso un test a inserimento verifica il livello di acquisizione dell'uso fattivo risultativo dell'imperfettivo.

2.6.2. Risultati delle indagini. I contributi sopraelencati dimostrano che il CP può rivelarsi uno strumento molto utile per la didattica del russo L2, poiché l'osservazione delle strutture più complesse da acquisire all'interno dei contesti d'uso accompagnati dal testo parallelo in italiano facilita la comprensione del loro funzionamento. In particolare, il CP è stato utile per:

- precisare la semantica e la combinabilità di alcune unità linguistiche, come nel caso dei cosiddetti 'falsi amici' nel contributo di Slavkova (2017).
- Identificare strutture preferenziali da un punto di vista contrastivo. Ad esempio, Biagini (B'jadžini 2017) utilizza il CP per presentare agli apprendenti italo-foni le forme di codifica della relazione concessiva in russo partendo dalle congiunzioni concessive tipiche dell'italiano e analizzando i loro equivalenti in russo. Biagini e Mazzoleni (2019) mostrano come l'utilizzo del CP consenta agli studenti di ricavare, attraverso le loro *query*, un ampio ventaglio di mezzi di espressione della relazione pre-concessiva in russo e come certe forme italiane presentino degli equivalenti analoghi in russo, mentre ad altre corrispondano costrutti diversi (come le marche di carattere avverbiale quali *možet byt'* e *požaluj* a fronte del futuro concessivo italiano). Gasanova-Mijat (2019) analizza le strategie di traduzione delle formule di etichetta dall'italiano al russo e viceversa, individuando equivalenti sistematici e occasionali per colmare alcune lacune nella presentazione di queste strutture linguistiche offerta dai manuali.

- Analizzare le asimmetrie relative al modo in cui l'italiano e il russo codificano l'informazione deittica di persona, spazio e tempo: Benigni e Ruvoletto (2019), individuati i contesti più problematici per gli apprendenti, identificano, in base a criteri strutturali e funzionali, forme di asimmetria di diverso tipo (ipercodifica vs. ipocodifica, diversa articolazione interna della deissi, diversa prospettiva nella concettualizzazione del riferimento deittico).
- Verificare i contesti in cui si realizza una data struttura: Nosedà (2020) propone di spiegare l'uso fattivo risultativo dell'aspetto imperfettivo sulla base dell'analisi dei contesti in cui si realizza. L'uso del CP aiuta gli studenti a cogliere equivalenze e differenze tra perfettivo e imperfettivo in russo e tra Tempi perfetti e imperfetto in italiano.

3. Conclusioni

La rassegna di questi studi ci fa comprendere come il CP italiano-russo sia un ausilio prezioso in svariati ambiti, dalla ricerca linguistica alla didattica del russo come lingua seconda, seppur emergano, allo stesso tempo, i punti deboli dello strumento, tra cui, in particolare: i) le dimensioni, e la conseguente impossibilità di analizzare un numero cospicuo di dati e ottenere risultati statisticamente rilevanti, fattore che costringe spesso gli studiosi a rivolgersi ad altre fonti per reperire esempi; ii) la prevalenza di testi narrativi e il gran numero di edizioni molto datate, che spesso non riflettono l'uso contemporaneo; iii) il ruolo del traduttore, di cui bisogna tenere conto nell'interpretazione dei dati, come pure i possibili errori di traduzione, che si riscontrano specialmente nelle traduzioni più datate. A questi si aggiunge l'attuale sbilanciamento del corpus a favore della sezione russo-italiano, nettamente più ampia rispetto al sotto-corpus in direzione italiano-russo¹⁴.

Ad eccezione del ruolo del traduttore, che caratterizza qualsivoglia corpus parallelo e ai cui limiti si può far fronte in parte attraverso un'accurata selezione dei testi, è possibile e necessario ovviare ai restanti problemi assicurando un ampliamento del CP costante e bilanciato, che prediliga l'inserimento di testi originali italiani, così come di testi non letterari.

La rassegna qui proposta ha messo anche in evidenza alcuni aspetti metodologici dell'applicazione del CP che possono essere migliorati: per facilitare il lettore, ma soprattutto per rendere uno studio comparabile, è importante dare informazioni sul numero di dati estratti e quantificare i risultati ottenuti, esplicitare sempre la direzione della ricerca e, in caso di confronto con altri corpora, normalizzare i dati fornendo non solo quantità assolute, ma anche relative. Infine, sono apprezzabili gli studi che sfruttano la bidirezionalità del CP, verificando che una struttura ricorrente nelle traduzioni sia presente anche nei testi originali in quella stessa lingua e non rappresenti, dunque, una tendenza dei testi tradotti ascrivibile al fenomeno degli universali traduttivi (Baker 1996).

¹⁴ Cfr. Nosedà (2018: 118-122) per una spiegazione dei motivi di tale sbilanciamento.

Bibliografia

- Artoni 2021a: D. Artoni, *Quando minutku non è un minutino: tradurre la pragmatica dei diminutivi nel corpus parallelo russo-italiano*, in: M. Boschiero, R. Faggionato, S. Zinato (a cura di), *Lezioni. Per Cinzia De Lotto*, Città di Castello 2021, pp. 23-32.
- Artoni 2021b: D. Artoni, *A Pragmatic-Based Approach to Translation: Mi Dai?/ Dammi! And Да! In the Parallel Russian-Italian Corpus*, "Translation Studies: Theory and Practice", 1, 2021, 1, pp. 46-53.
- Baker 1996: M. Baker, *Corpus-based translation studies: the challenges that lie ahead*, in: H. Somers (a cura di), *Terminology, LSP and Translation: Studies in language engineering in honour of Juan C. Sager*, Amsterdam 1996, pp. 175-186.
- Badolati 2022: M.T. Badolati, *Le unità fraseologiche bibliche con un componente onomastico in russo e italiano: un'analisi contrastiva*, in: C. De Giovanni (a cura di), *Fraseologia e paremiologia tra lingua e discorso*, Roma 2022, pp. 403-418.
- Badolati 2021: M.T. Badolati, *Adamo ed Eva nella fraseologia russa e italiana. Le lingue slave: sviluppi teorici e prospettive applicative*, in: E. Gherbezza, V. Laskova, A.M. Perissutti (a cura di), *Atti dell'VIII incontro di linguistica slava (Udine, 10-12 settembre 2020)*, Roma 2021, pp. 145-167.
- Benacchio, Biasio 2023: R. Benacchio, M. Biasio, *Alcune note sulla distribuzione di rus. sečas/ teper' e it. ora/adesso*, "MediAzioni", XXXVI, 2023, 1, pp. A255-A276.
- Benigni 2017: V. Benigni, *Strategie di intensificazione in russo: i nomi non scalari tra semantica e pragmatica*, in: M. di Filippo, F. Esvan (a cura di), *Studi di Linguistica Slava. Volume dedicato a Lucyna Gebert*, Napoli 2017, pp. 15-34.
- Benigni 2019: V. Benigni, *Ni chrena v ètom ne razbirajus'*, in: I. Krapova, S. Nistratova, L. Ruvoletto (a cura di), *Studi di linguistica slava. Nuove prospettive e metodologie di ricerca*, Venezia 2019, pp. 37-52.
- Benigni 2020: V. Benigni, *Èto uže meloči. Da avverbio fasale ad intensificatore: il caso di uže*, "Studi Slavistici", XVII, 2020, 2, pp. 131-155.
- Benigni, Gebert 2019: V. Benigni, L. Gebert, *Determinatezza nominale e aspetto verbale: il caso dei verbi supporto*, in: O. Inkova, D. Mancheva (a cura di), *Contrastes. Études de linguistique slavo-romane*, Alessandria 2019, pp. 41-62.
- Benigni, Gebert 2023: V. Benigni, L. Gebert, *Determinatezza nominale e aspetto verbale nella codifica della referenzialità: polisemia e ambiguità interpretative*, "MediAzioni", XXXVI, 2023, 1, pp. A23-A45.
- Benigni, Ruvoletto 2019: V. Benigni, L. Ruvoletto, *Asimmetrie nella codifica dell'informazione deittica: italiano vs russo*, "Italica Wratislaviensia", X, 2019, 1, pp. 31-58.

- Biagini 2020: F. Biagini, *Il periodo ipotetico all'indicativo: tempo e aspetto in italiano e in russo*, in: O. Inkova, M. Nowakowska, S. Scarpel (éds.), *Systèmes linguistiques et textes en contraste. Études de linguistique slavo-romane*, Kraków 2020, pp. 46-65.
- Biagini 2020: F. Biagini, *I verbi naest'sja, napit'sja e doždat'sja e i loro equivalenti italiani nel corpus parallelo italiano-russo: interazione tra semantica della base verbale, valore dei circonfissi e contesto*, "Studi Slavistici", XVII, 2020, 2, pp. 77-98.
- Biagini 2023: F. Biagini, *Gli equivalenti russi della perifrasi verbale imminente italiana stare per*, in: F. Biagini, O. Inkova (a cura di), *Lingue slave e lingue romanze a confronto: dalla frase al testo*, "MediAzioni", xxxvi, 2023, 1, pp. A46-A64.
- Biagini, Bonola 2019: F. Biagini, A. Bonola, *Descrizione semantico-funzionale delle particelle russe e corpora paralleli. Un'analisi contrastiva (italiano-russo) corpus-based di ved'*, in: I. Krapova, S. Nistratova, L. Ruvoletto (a cura di), *Studi di linguistica slava. Nuove prospettive e metodologie di ricerca*, Venezia 2019, pp. 53-77.
- Biagini et al. 2019: F. Biagini, A. Bonola, V. Nosedà, *Il Corpus parallelo italiano-russo e russo-italiano del NKRJA: progetto di ampliamento, applicazioni e sviluppi*, in: M. C. Bidovec, M. C. Bragone (a cura di), *Il mondo slavo e l'Europa: contributi presentati al VI Congresso Italiano di Slavistica: Torino, 28-30 settembre 2016*, Firenze 2019, pp. 34-45.
- Biagini, Gebert in stampa: Biagini F., L. Gebert, *Gli equivalenti russi delle perifrasi verbali a valore aspettuale in italiano*, in: H. Bažec e J. Cergol (a cura di), *University of Primorska Press*, in stampa.
- Biagini, Mazzoleni 2017: F. Biagini, M. Mazzoleni, *I costrutti preconcettivi in italiano e in russo: primi risultati di una ricerca sul corpus parallelo del NKRJA*, "Supostavitelno ezikoznanie", XLII, 2017, 4, pp. 77-88.
- Biagini, Mazzoleni 2018: F. Biagini, M. Mazzoleni, *I costrutti preconcettivi in italiano e in russo: uno studio sul corpus parallelo del NKRJA*, "Italica Belgradensia", 1, 2018, pp. 27-47.
- Biagini, Mazzoleni 2019: F. Biagini, M. Mazzoleni, *L'utilizzo del corpus parallelo italiano-russo del NKRJA per la didattica del russo L2 ai discendenti italiani: il caso dei costrutti preconcettivi*, "H2D | Revista De Humanidades Digitais", 1 2019, 1, DOI: <https://doi.org/10.21814/h2d.233>.
- Biagini, Mazzoleni 2023: F. Biagini, M. Mazzoleni, *Sull'ordine sequenziale di protasi e apodosi in italiano e russo nel corpus parallelo del NKRJA*, in: A. De Cesare, A. Ferrari, L. Lala, F. Pecorari, (a cura di), *Forme della scrittura italiana contemporanea in prospettiva contrastiva. La componente testuale*, Firenze 2023, pp. 19-30.

- B'jadžini 2017: F. B'jadžini, *Vyraženie ustupitel'nogo mežfrazovogo otnošenija na prodvinutom etape obučenija ruskomu jazyku kak inostrannomu pri pomošči Russko-ital'janskogo parallel'nogo korpusa*, in: E. Jasin (pod red.), *XVII Aprel'skaja meždunarodnaja naučnaja konferencija po problemam razvitija ekonomiki i obščestva*, IV, Moskva 2017, pp. 390-398.
- B'jadžini 2019: F. B'jadžini, *Sopostavitel'nyj analiz prekoncessivnyh konstrukcij v ruskom i ital'janskom jazykach na materiale rusko-ital'janskogo parallel'nogo korpusa NKRJA*, in: A. Nikunlassi, E. Protassova (eds.), *Russian Language in the Multilingual World*, Helsinki 2019 (= *Slavica Helsingiensia*, 52), pp. 342-348.
- Bonola 2016: A. Bonola, *Semantiko-pragmatičeskij potencial slova vdrug v ruskom jazyke i ego perevod na ital'janskij: kontrastivnyj analiz na osnove parallel'nogo rusko-ital'janskogo korpusa*, "Vestnik Novosibirskogo Gosudarstvennogo Pedagogičeskogo Universiteta", 2016, 2, pp. 24-37.
- Bonola 2017: A. Bonola, *Il corpus parallelo italiano-russo-italiano come strumento per l'indagine contrastiva delle marche discorsive e testuali: il caso di comunque e i suoi traduttivi russi*, "Supostavitelno ezikoznanie", XLII, 2017, 4, pp. 89-104.
- Bonola 2020: A. Bonola, *Allora e i suoi equivalenti in russo: analisi contrastiva in un corpus parallelo italiano-russo*, in: O. Inkova, M. Nowakowska, S. Scarpel (éds.), *Systèmes linguistiques et textes en contraste. Études de linguistique slavo-romane*, Kraków 2020, pp. 251-269.
- Bonola 2022: A. Bonola, *Russkoe uže i ital'janskoe già: sopostavitel'nyj analiz na osnove parallel'nogo korpusa*, "Vestnik moskovskogo universiteta", 2022, 6, pp. 46-57.
- Bonola, Nosedà 2020: A. Bonola, V. Nosedà, *The use of parallel Corpora for a contrastive (Russian-Italian) description of discourse markers: new instruments compared to traditional lexicography*, in: C. Marras, M. Passarotti, G. Franzini, E. Litta (a cura di), *Atti del IX Convegno Annuale AIUCD. La svolta inevitabile: sfide e prospettive per l'Informatica Umanistica*, Milano, pp. 39-46.
- Borin 2002: L. Borin, *...And never the twain shall meet?*, in: L. Borin (ed.), *Parallel Corpora, parallel worlds. Selected papers from a symposium on parallel and comparable corpora at Uppsala University, Sweden, 22-23 April 1999*, Amsterdam, New York 2002, pp. 1-43.
- Di Filippo 2020: M. di Filippo, *Cromonimi e indeterminatezza. Uno studio contrastivo russo-italiano*, in: O. Inkova, M. Nowakowska, S. Scarpel (éds.) *Systèmes linguistiques et textes en contraste. Études de linguistique slavo-romane*, Krakow 2020, pp. 177-192.

- Di Filippo 2023: M. di Filippo, – “Mi dia due zucchine” / – “Vanno bene 800 grammi?” *Quando i numeri non contano*, “MediAzioni”, xxxvi, 2023, 1, pp. A132-A153.
- Fici 2020: F. Fici, *Quel che resta del Piuccheperfecto nelle lingue slave moderne (un confronto con l’italiano)*, in: P. Macurová, M. di Filippo (a cura di), *Tre lingue, tre cuori. Studi in onore di François Esvan*, Napoli 2020, pp. 29-44.
- Fiči 2020: F. Fiči, *Pljuskvamperfekt v ital’janskom jazyke i ego ekvivalenty v russkom*, in: A.A. Kibrik, K.P. Semenova, D.V. Sičnava, S.G. Tatevosov, A.Ju. Urmančeva (pod red.), *VAProsy jazykoznanija. Megabornik nanostatej*, Moskva 2020, pp. 509-514.
- Gasanova-Mijat 2019: M. Gasanova-Mijat, *Formuly rečevogočiketa*, in: I. Krapova, S. Nistratova, L. Ruvoletto (a cura di), *Studi di linguistica slava. Nuove prospettive e metodologie di ricerca*, Venezia 2019, pp. 191-207.
- Gilquin 2010: G. Gilquin, *Corpus, cognition and causative constructions*, Amsterdam 2010.
- Granger 2003: S. Granger, *The corpus approach: a common way forward for Contrastive Linguistics and Translation Studies?*, in: S. Granger, J. Lerot, S. Petch-Tyson (eds.), *Corpus-based Approaches to Contrastive Linguistics and Translation Studies*, Amsterdam-New York 2003, pp. 17-29.
- In’kova 2018: O. In’kova, *Nadkorpurnaja basa dannyh kak instrument izučenija formal’noj variativnosti konnektorov*, “Komp’juternaja Lingvistika i Intellektual’nye Technologii”, xvii, 2018 (= Trudy Meždunarodnoj Konferencii “Dialog 2018”, Moskva), pp. 240-253.
- In’kova 2020: O. In’kova, *Pokazateli otnošenija soputstvovanija v russkom jazyke i ich ital’janskiekvivalenty*, in: O. In’kova, M. Nowakowska, S. Scarpel (eds.), *Systèmes linguistiques et textes en contraste. Études de linguistique slavo-romane*, Kraków 2020, pp. 46-65.
- In’kova 2023: O. In’kova, *Semantika meždju tem v svete korpurnych dannyh*, “MediAzioni”, xxxvi, 2023, 1, pp. A233-A254.
- Johansson 2003: S. Johansson, *Contrastive linguistics and corpora*, in: S. Granger, J. Lerot, S. Petch-Tyson (eds.), *Corpus-based Approaches to Contrastive Linguistics and Translation Studies*, Amsterdam, New York 2003, pp. 31-44.
- Leone 2019: M. Leone, *Il ‘passato discontinuo’ come categoria semantico-funzionale nella lingua russa contemporanea*, in: I. Krapova, S. Nistratova, L. Ruvoletto (a cura di), *Studi di linguistica slava. Nuove prospettive e metodologie di ricerca*, Venezia 2019, pp. 271-284.

- Levontina, Denisova 2017: I. B. Levontina, G. Denisova, *Ital'janskoe magari i ego russkie perevodnye ekvivalenty: raznye diskursivnye strategii*, in: S. Pavlovič (pod red.), *Komp'juternaja lingvistika i intellektual'nye tehnologii: po materialam meždunarodnoj konferencii "Dialog 2017"*, II, Moskva 2017, pp. 261-271.
- McEnery, Wilson 2001: T. McEnery, A. Wilson, *Corpus Linguistics*, Edinburgh 2001.
- Noël 2003: D. Noël, *Translations as evidence for semantics: an illustration*, "Linguistics", XLI, 2003, 4, pp. 757-785.
- Noseda 2017: V. Noseda, *Corpora paralleli e linguistica contrastiva: ampliamento e applicazioni del corpus italiano-russo nel Nacional'nyj Korpus Russkogo Jazyka*, Tesi di dottorato, Milano 2017, <<http://hdl.handle.net/10280/24613>>.
- Noseda 2018: V. Noseda, *La corpus revolution russa e il corpus parallelo italiano-russo: storia, criteri di compilazione e usi*, "L'Analisi linguistica e letteraria", XXIV, 2018, 2, pp. 115-132.
- Noseda 2019a: V. Noseda, *Il Corpus parallelo italiano-russo per lo studio del causativo in chiave contrastiva*, in: I. Krapova, S. Nistratova, L. Ruvoletto (a cura di), *Studi di linguistica slava. Nuove prospettive e metodologie di ricerca*, Venezia 2019, pp. 357-374.
- Noseda 2019b: V. Noseda, *The use of parallel corpora to investigate causation in Russian*, "Komp'juternaja lingvistika i intellektual'nye tehnologii", XVIII, 2019, Dopolnitel'nyj tom. Po materialam ežegodnoj meždunarodnoj konferencii "Dialog", Moskva, pp. 164-175.
- Noseda 2019c: V. Noseda, *Parallel corpora with multiple translation variants: a compared analysis of fifteen Italian translations of Gogol's short story "The Overcoat"*, in: V. Zacharov (pod red.) *Trudy Meždunarodnoj konferencii Korpusnaja Lingvistika 2019*, Sankt-Peterburg 2019, pp. 405-411.
- Noseda 2020a: V. Noseda, *La concorrenza degli aspetti in russo: particolarità semantiche e possibile resa in italiano*, in: O. Inkova, M. Nowakowska, S. Scarpel (éds.) *Systèmes linguistiques et textes en contraste. Études de linguistique slavo-romane*, Krakow 2020, pp. 85-102.
- Noseda 2020b: V. Noseda, *L'uso dei corpora nell'insegnamento dell'aspetto verbale russo a classi di italofoeni*, "Nuova Secondaria", XXXVII, 2020, 5, pp. 87-91.
- Noseda 2021: V. Noseda, *Le particelle razve e neuželi alla luce del Corpus parallelo russo-italiano*, in: E. Gherbezza, V. Laskova, A. M. Perissutti (a cura di), *Atti dell'VIII incontro di linguistica slava (Udine, 10-12 settembre 2020)*, Roma 2021, pp. 485-509.

- Nozeda 2015: V. Nozeda, *Perevod russkich prefiksial'nych glagolov pamjati v svete dannyh ital'jansko-russkogo parallel'nogo korpusa*, in: V. Zacharov, O. Mitrofanova, M. Chochlova (pod red.), *Trudy meždunarodnoj konferencii "Korpusnaja Lingvistika-2015"*, Sankt-Peterburg 2015, pp. 351-360.
- Pecorari, Pinelli 2023: F. Pecorari, E. Pinelli, *Il connettivo concessivo tuttavia e i suoi corrispettivi russi: un'analisi contrastiva corpus-based*, "MediAzioni", xxxvi, 2023, 1, pp. A154-A176.
- Plungjan 2009: V. Plungjan (pod red.), *Nacional'nyj korpus russkogo jazyka: 2006-2008. Novye rezul'taty i perspektivy*, Sankt-Peterburg 2009.
- Prandi et al. 2005: M. Prandi, G. Gross, C. De Santis, *La finalità. Strutture concettuali e forme d'espressione in italiano*, Firenze 2005.
- Ruvoletto 2020: L. Ruvoletto, *I verbi intensivi russi con circonfisso. Implicazioni pragmatiche e traduzione in italiano*, "Studi Slavistici", xvii, 2020, 2, pp. 117-130.
- Sinclair 2004: J. Sinclair (ed.), *How to use corpora in language teaching*, Amsterdam 2004.
- Slavkova 2017: S. Slavkova, *Ispol'zovanie dannyh nacional'nogo korpusa russkogo jazyka v prepodavanii sovremennogo russkogo jazyka v inostrannoj auditorii*, in: E. Jasin (pod red.), *xvii Aprel'skaja meždunarodnaja naučnaja konferencija po problemam razvitija ekonomiki i obščestva*, iv, Moskva 2017, pp. 427-436.
- Slavkova 2019: S. Slavkova, *Strategii perevoda russkich pristavočnyh glagolov na ital'janskij jazyk (na materiale rusko-ital'janskogo parallel'nogo podkorpusa NKJRJA)*, in: I. Krapova, S. Nistratova, L. Ruvoletto (a cura di), *Studi di linguistica slava. Nuove prospettive e metodologie di ricerca*, Venezia 2019, pp. 511-529.
- Šmelev, Zaliznjak 2017: A. Šmelev, A. Zaliznjak, *Reversivnyj perevod kak instrument lingvističeskogo analiza*, "Komp'juternaja lingvistika i intellektual'nye tehnologii", xvi, 2017, 2 (Po materialam ežegodnoj meždunarodnoj konferencii "Dialog", Moskva), pp. 394-406.
- Zaliznjak 2015: A. Zaliznjak, *Lingvospecifičnye edinicy russkogo jazyka v svete kontrastivnogo korpusnogo analiza "Komp'juternaja lingvistika i intellektual'nye tehnologii"*, xiii, 2015 (Po materialam ežegodnoj meždunarodnoj konferencii "Dialog", Moskva 2015), pp. 683-695.
- Zaliznjak et al. 2018: A. Zaliznjak, G. Denisova, I. Mikaëljan, *Russkoe kak-nibud' po dannym parallel'nyh korpusov*, "Komp'juternaja lingvistika i intellektual'nye tehnologii", xvii, 2018 (Po materialam ežegodnoj meždunarodnoj konferencii "Dialog", Moskva), pp. 803-817.

Abstract

Francesca Biagini, Tatsiana Maiko, Valentina Nosedà

The Italian-Russian Parallel Corpus of the Nacional'nyj Korpus Russkogo Jazyka (NKРJа). Evolution and Applications in Italian Slavistics Research

The aim of this article is to present a comprehensive overview of the studies conducted in Italy using the Italian-Russian parallel corpus of the *Nacional'nyj Korpus russkogo jazyka* (NKРJа), implemented in 2013 and then expanded since 2015. We provide current information on the size of the corpus and a description of the types of research conducted in various fields (contrastive linguistics, translation studies and studies dedicated to the teaching of Russian). The article discusses the practical applications of the corpus and presents the results obtained. On the basis of this overview, the potential and limits of the tool are highlighted, with a view to its continuous and constant improvement.

Keywords

Corpus Linguistics; Parallel Corpora; Russian-Italian Parallel Corpus; Italian Slavistics.