



UNIVERSITÀ CATTOLICA DEL SACRO CUORE
MILANO

Dottorato di ricerca in Scienze Linguistiche e Letterarie
Ciclo XXIX
S.S.D.: L-LIN/21

CORPORA PARALLELI E LINGUISTICA CONTRASTIVA:
AMPLIAMENTO E APPLICAZIONI DEL CORPUS ITALIANO-RUSSO NEL
NACIONAL'NYJ KORPUS RUSSKOGO JAZYKA

Coordinatore: Ch.mo Prof. Dante Liano

Tesi di Dottorato di: Valentina NOSEDA
Matricola: 4211960

Anno Accademico 2015/2016

ABSTRACT

La Linguistica dei corpora - che fa uso di corpora elettronici annotati per lo studio delle lingue - è un approccio ormai diffuso e consolidato. I corpora paralleli, in particolare, in cui i testi in una lingua A sono allineati con la traduzione in lingua B, sono uno strumento molto utile nell'analisi contrastiva. La mancata disponibilità di corpora paralleli di qualità per le lingue di nostro interesse - russo e italiano - ci ha portati a volere ampliare e migliorare il corpus parallelo italiano-russo presente come corpus pilota nel *Nacional'nyj Korpus Russkogo Jazyka* (Corpus Nazionale della Lingua Russa). Il presente lavoro ha avuto pertanto uno scopo applicativo e uno teorico. Da un lato, dopo aver studiato le questioni imprescindibili per la progettazione di un corpus di qualità, sono stati stabiliti i criteri per l'ampliamento e inseriti nuovi testi, consentendo così al corpus parallelo di passare da 700.000 a più di 4 milioni di parole, entità che consente ora di condurre ricerche scientificamente valide. In seguito, sono state proposte tre analisi *corpus-based* così da mettere in luce le potenzialità del corpus ampliato: lo studio dei verbi prefissali di memoria russi e la loro resa in italiano; il confronto tra il causativo analitico italiano "fare + infinito" e il causativo russo; l'analisi comparata di quindici versioni italiane de *Il Cappotto* di N. Gogol'. Le tre analisi hanno consentito di avanzare innanzitutto osservazioni di carattere metodologico in vista di un ulteriore ampliamento e miglioramento del corpus parallelo italiano-russo. In secondo luogo, la prospettiva *corpus-based* si è dimostrata utile per approfondire lo studio di questi temi dal punto di vista teorico.

Parole chiave: linguistica dei corpora, progettazione di un corpus, corpora paralleli, linguistica contrastiva

Corpus Linguistics - which exploits electronic annotated corpora in the study of languages - is a widespread and consolidated approach. In particular, parallel corpora, where texts in a language are aligned with their translation in a second language, are an extremely useful tool in contrastive analysis. The lack of good parallel corpora for the languages of our interest - Russian and Italian - has led us to work for improving the Italian-Russian parallel corpus available as a pilot corpus in the Russian National Corpus. Therefore, this work had a twofold aim: practical and theoretical. On the one hand, after studying the essential issues in order to design a high-quality corpus, all the criteria for expanding the corpus were established and the number of texts was increased, allowing the Italian-Russian parallel corpus, which counted 700.000 words, to reach more than 4 million words. As a result, it is now possible to conduct scientifically valid research based on this corpus. On the other hand, three corpus-based analyses were proposed to highlight the potential of the corpus: the study of prefixed Russian memory verbs and their translation into Italian; the comparison between the Italian analytic causative "fare + infinitive" and Russian causative verbs; The comparative analysis of fifteen Italian versions of *The Overcoat* by N. Gogol'. These analyses first of all allowed to advance some methodological remarks considering a further enlargement and improvement of the Italian-Russian parallel corpus. Secondly, the corpus-based approach has proved to be useful in deepening the study of these subjects from a theoretical point of view.

Key words: corpus linguistics, corpus design, parallel corpora, contrastive linguistics

INDICE

INTRODUZIONE	4
CAPITOLO 1.....	12
1.1. LA LINGUISTICA DEI CORPORA: ASPETTI TEORICI E SPUNTI CRITICI.....	12
1.2. SVILUPPO DELLA LC	18
1.3. DEFINIZIONE E TIPOLOGIE DEI CORPORA.....	23
<i>1.3.1. Le caratteristiche di un corpus</i>	<i>26</i>
1.3.1.1. Il formato elettronico	27
1.3.1.2. L'annotazione	27
1.3.1.3. Rappresentatività e bilanciamento dei testi	29
<i>1.3.2. Tipi di corpora tradizionali</i>	<i>30</i>
1.4. I CORPORA PARALLELI	31
<i>1.4.1. Possibili applicazioni dei CP.....</i>	<i>32</i>
1.5. IL NKRJA E I SUOI CORPORA PARALLELI	38
<i>1.5.1. Nacional'nyj Korpus Russkogo Jazyka: storia e struttura</i>	<i>38</i>
<i>1.5.2. I corpora paralleli del NKRJa</i>	<i>43</i>
<i>1.5.3. Il primo corpus parallelo italiano-russo / russo-italiano del NKRJa</i>	<i>45</i>
1.6. COLLEZIONI DI TESTI PARALLELI ITALIANO-RUSSO AL DI FUORI DEL NKRJA	47
1.8. CRITERI PER LA PROGETTAZIONE DI UN CORPUS	51
<i>1.8.1. Finalità.....</i>	<i>52</i>
<i>1.8.3. Corpus diacronico o sincronico</i>	<i>55</i>
<i>1.8.4. Rappresentatività e bilanciamento</i>	<i>56</i>
<i>1.8.5. Dimensione</i>	<i>58</i>
<i>1.8.6. Corpus di testi completi o di campioni</i>	<i>60</i>
<i>1.8.7. Comparabilità.....</i>	<i>61</i>
<i>1.8.8. Tipologia testuale e classificazione dei testi.....</i>	<i>62</i>
1.9. CRITERI SPECIFICI PER LA PROGETTAZIONE DI CORPORA BILINGUI PARALLELI.....	63
1.10. PROPOSTA PER UN AMPLIAMENTO DEL CORPUS PARALLELO ITALIANO-RUSSO DEL NKRJA.....	65
BIBLIOGRAFIA	71
CAPITOLO 2.....	80
2.1. LE DIMENSIONI	80
2.2. LA SCELTA DEI TESTI	81
2.2.1. <i>Prosa letteraria russa</i>	<i>81</i>
2.2.2. <i>Prosa letteraria italiana</i>	<i>92</i>
2.2.3. SAGGISTICA.....	96
2.3. LE FASI DI COSTRUZIONE DEL CORPUS	102
CAPITOLO 3.....	106

3.1. LO STUDIO DEI PREFISSI VERBALI NELLA RUSSISTICA: DAGLI ANNI '90 AD OGGI..	108
3.2. LO STUDIO DEI PREFISSI VERBALI IN CHIAVE CORPUS-BASED.....	113
3.3. PREFISSI CHE CARATTERIZZANO I VERBI DI MEMORIA NELLA LINGUA RUSSA	118
3.4. ANALISI DELLE TRADUZIONI IN ITALIANO.....	120
3.4.1. Osservazioni metodologiche preliminari.....	120
3.4.2. Le strategie traduttive della resa in italiano	122
3.5. OSSERVAZIONI CONCLUSIVE.....	144
BIBLIOGRAFIA	148
CAPITOLO 4.....	152
4.1. CLASSIFICAZIONE FUNZIONALE DEL VERBO FARE.....	154
4.2. SITUAZIONI CAUSATIVE E VERBI CAUSATIVI: INTRODUZIONE	160
4.3. GLI STUDI LINGUISTICI RUSSI E ITALIANI SUL CAUSATIVO	168
4.3.1. Il causativo italiano	168
4.3.2. Il causativo russo	170
4.4. ANALISI DEGLI ESEMPI	175
4.4.1. Metodologia.....	176
4.4.2. Ruolo e status degli attanti nella situazione causativa.....	177
4.4.3. F causativo: strategie traduttive dall'italiano al russo.....	180
4.5. RISULTATI	184
4.5.1. Riduzione degli attanti ed enunciati a soggetto unico della lingua russa... ..	185
4.5.2. Le sotto-funzioni di F causativo.....	187
4.6. OSSERVAZIONI CONCLUSIVE.....	193
BIBLIOGRAFIA	194
CAPITOLO 5.....	199
5.1. LE TRADUZIONI ITALIANE DEL CAPPOTTO	201
5.2. I DIALOGHI DI AKAKIJ AKAKEVIČ: ANALISI LINGUISTICA	208
5.2.1. Aspetti linguistici, metodologia e limiti	210
5.2.2. I segnali discorsivi.....	214
5.3. L'ANALISI.....	226
5.3.1. La particella <i>mozo</i>	227
5.3.2. La resa dell'ellissi verbale in italiano	239
BIBLIOGRAFIA	247
CONCLUSIONI	251

INTRODUZIONE

Il nostro lavoro si inserisce negli studi relativi alla Linguistica dei Corpora (LC) e più precisamente propone un'applicazione nell'ambito dei corpora paralleli¹. Da più di cinquant'anni, infatti, la LC trova spazio in quanto metodologia diffusa e consolidata e il suo successo è dovuto proprio all'utilizzo dei corpora, strumenti che consentono una ricerca su quantità di dati considerevoli grazie al loro formato elettronico, all'annotazione (cfr. cap. 1, par. 1.3.1.2) e al motore di ricerca che ne permette l'interrogazione: più è sofisticato, maggiori saranno le potenzialità e i risultati dell'analisi condotta.

L'uso dei corpora è espressione di una concezione diffusasi soprattutto in area anglosassone (M. Halliday, J. Sinclair), secondo cui l'indagine linguistica deve considerare esempi di lingua autentica, prodotta dai parlanti in contesti reali. Questo approccio predilige lo studio della lingua "dell'uso" e mette in secondo piano i tradizionali concetti di "corretto" e "non corretto" a favore di "ricorrente" e "non ricorrente"². Per individuare i fenomeni ricorrenti e le strutture preferenziali di una lingua è dunque necessario confrontare in poco tempo un gran numero di dati mediante consistenti corpora elettronici e software di interrogazione. Ma la LC non si ferma allo studio statistico dei dati: le indagini linguistiche basate sui corpora hanno consentito di condurre ricerche in precedenza irrealizzabili, rivelando aspetti della lingua prima ignoti come, per esempio, i fenomeni della collocazione³ e della prosodia semantica⁴, esito dello

¹ Per corpus parallelo si intende una raccolta di testi originali in una lingua (definita anche *Source Language*), allineati con le rispettive traduzioni in una o più lingue diverse (*Target Language*) (Gandin 2009: 134) [Gandin S. (2009). *Linguistica dei corpora e traduzione: definizioni, criteri di compilazione e implicazioni di ricerca dei corpora paralleli*. Annali della Facoltà, 5, Università degli Studi di Sassari, pp. 133-152].

² La distinzione tra correttezza grammaticale e ricorrenza si collega al dibattito tra Linguistica dei corpora e generativismo, di cui parleremo nel primo capitolo, e alla distinzione proposta da N. Chomsky tra *competence* (la competenza del parlante nativo) e *performance* (la lingua effettivamente prodotta). Riguardo a quest'ultima Chomsky afferma: "Everything is possible in language. The question is: how does it relate to the structure of English?" [Aarts (2006). *Corpus Linguistics, Chomsky and Fuzzy Trees Fragments*. In Mair C., Hundt M. *Corpus Linguistics and Linguistics Theory. Papers from the Twentieth International Conference on English Language Research on Computerized Corpora, (ICAME 20)*, Freiburg, Breisgau 1999. Amsterdam, Atlanta: Rodopi, pp. 5-13]. Dall'altro lato J. Sinclair, invita i linguisti a "fidarsi del testo", aprendosi all'eventualità che la nostra intuizione di parlanti nativi (e quindi la nostra competenza) possa essere smentita: "I am advocating that we should trust the text. We should be open to what it may tell us. We should not impose our ideas on it, except perhaps just to get started" [Sinclair J. (2004b). *Trust the text. Language, corpus and discourse*. London, New York: Routledge].

³ Concetto introdotto da J. R. Firth, indicante la co-occorrenza di unità lessicali nel discorso.

⁴ Il concetto di prosodia semantica (ingl: *semantic prosody*) descrive il modo in cui una parola è usata prevalentemente in contesti negativi o positivi. Cfr: Sinclair J., Renouf A. (1991). *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.

stretto rapporto tra lessico e sintassi, tra forma e significato, rilevato in modo puntuale dalla LC.

Come avremo modo di approfondire nel primo capitolo di questo lavoro, la LC ha suscitato numerose critiche, soprattutto in ambito generativista; tuttavia la sua crescita è stata continua e attualmente i corpora elettronici accessibili nel web sono innumerevoli, si adattano a svariate esigenze, poiché presentano caratteristiche di diverso tipo, e esistono per la maggior parte delle lingue studiate al mondo.

Nonostante ciò, nell'ambito dei corpora elettronici si riscontra una lacuna significativa proprio nell'area dei corpora paralleli, soprattutto per alcune coppie di lingue. Per la coppia russo-italiano, ad esempio, risorse di questo tipo sono assai scarse e, ove presenti, di poca utilità per la ricerca scientifica (cfr. par.1.6.). Nascono da qui, dunque, il tema e il fine del nostro lavoro di dottorato, che è consistito nel potenziare il corpus parallelo italiano-russo presente in forma molto limitata all'interno del Corpus Nazionale della Lingua Russa (*Nacional'nyj Korpus Russkogo Jazyka* - NKRJa)⁵. Il nostro intento è contribuire a realizzare uno strumento che permetta ricerche linguistiche affidabili, ossia un CP bidirezionale russo-italiano di entità adeguata a tale scopo⁶, e il lavoro che presentiamo rientra effettivamente in un progetto più ampio, che vede la collaborazione di tre istituzioni: l'Università Cattolica di Milano, l'Istituto di Lingua Russa a Mosca e la Scuola Superiore di Lingue Moderne per Interpreti e Traduttori di Forlì (Università di Bologna).

Il NKRJa, all'interno del quale abbiamo realizzato il nostro progetto, nasce nel 2004 come primo corpus nazionale rappresentativo della lingua russa. È formato da un corpus principale, comprendente testi di diversi generi letterari e risalenti a diversi periodi, e da vari corpora specialistici, tra cui i corpora paralleli. Oggi la sezione dei corpora paralleli include quindici coppie di lingue, tra cui russo e italiano.

Il corpus parallelo italiano-russo, inizialmente creato dai collaboratori del NKRJa come corpus pilota, prima del nostro intervento conteneva cinque opere, tre italiane con la traduzione in russo e due opere russe affiancate dalla traduzione italiana⁷. Fin da subito

⁵ www.ruscorpora.ru

⁶ Il nostro obiettivo iniziale era raggiungere una dimensione analoga all'ENPC (corpus parallelo di testi in inglese e norvegese, in parte modello del nostro progetto), ovvero circa 3 milioni di parole. Tuttavia sarebbe auspicabile a lungo termine raggiungere un volume di 20 milioni di parole, entità attuale del corpus parallelo russo-inglese del NKRJa.

⁷ Bulgakov M. A. *Master i Margarita* (prima e seconda parte) (1929-1940), Collodi C. *Pinocchio* (1883), Dostoevskij F. M. *Belye Noči* (1848), Eco U. *Il nome della rosa* (1980), Pirandello L. *La signora Frola e il signor Ponza, suo genero* (1917).

era stato dunque progettato come corpus bi-direzionale, comprendente testi in lingua A tradotti in lingua B e viceversa. All'inizio del nostro lavoro, esso contava all'incirca 700 mila parole⁸, un'entità troppo limitata per condurre indagini scientificamente valide. Il nostro primo obiettivo è stato pertanto ampliare questo corpus in modo da superare lo stato di corpus-pilota e renderlo fruibile attraverso il portale del NKRJa, creando e testando uno strumento che, pur necessitando di uno sviluppo continuo, è ora accessibile e sufficientemente ampio per essere utilizzato a scopi di ricerca⁹.

Per realizzare questo obiettivo, eminentemente applicativo, è stata necessaria una ricerca preliminare, di cui riferiremo nella nostra dissertazione, e a cui ha fatto seguito l'ampliamento vero e proprio. Il lavoro si è dunque articolato nelle seguenti fasi:

- 1) considerazione e discussione degli aspetti teorici implicati dall'ampliamento e perfezionamento di un corpus parallelo;
- 2) ampliamento del corpus bilaterale italiano-russo. In particolare:
 - selezione dei testi e loro conversione in formato digitale;
 - ottenimento dei permessi da parte degli editori e/o autori per pubblicare in rete (seppur parzialmente) gli originali e le traduzioni in base alle leggi vigenti sui diritti d'autore;
 - allineamento dei testi con un programma reso disponibile dal NKRJa.

Una volta realizzato l'ampliamento del corpus ne abbiamo infine testata l'utilità e l'efficacia svolgendo tre ricerche nell'ambito della linguistica contrastiva.

La nostra dissertazione risulta pertanto divisa in due parti: la prima dedicata alla LC, ai criteri per la progettazione di un corpus e alle fasi dell'ampliamento del nostro corpus parallelo (capitoli 1 e 2), la seconda incentrata sulle tre analisi linguistiche contrastive condotte nel corpus parallelo bilaterale italiano-russo da noi creato, ciascuna esposta in un capitolo a sé (capitoli 3, 4 e 5).

Nel primo capitolo introdurremo e definiremo la linguistica dei corpora e parleremo dei dibattiti che si sono sviluppati attorno alla nascita di questa metodologia e al suo status. Chiariremo quindi cos'è un corpus e descriveremo il NKRJa. Ci soffermeremo poi sul concetto di corpus parallelo in generale e sull'importanza di disporre di un corpus di testi paralleli in russo e in italiano il più completo e sistematico possibile. In seguito ci concentreremo sui corpora paralleli presenti all'interno del

⁸ Quando in riferimento a un corpus si parla di "numero di parole" si intendono i "token", ossia tutte le singole forme di parola presenti nel corpus e non i lessemi (definiti *types*).

⁹ <http://ruscorpora.ru/search-para-it.html>.

NKRJa, con particolare attenzione a quello italiano-russo e, successivamente, descriveremo altri corpora disponibili in rete per queste due lingue. Seguirà un'analisi dei criteri di compilazione di un corpus, per la quale si è scelto di prendere in esame alcuni dei maggiori corpora ad oggi esistenti, non solo paralleli. La scelta di considerare anche altri tipi di corpora – oltre a quelli paralleli – ha diverse ragioni: innanzitutto molti dei criteri di compilazione sono comuni a tipi di corpora differenti; in secondo luogo, non esiste ancora un numero sufficiente di corpora paralleli così sistematici e ben strutturati da costituire un modello di riferimento unanimemente accettato. Al termine della descrizione dei criteri di compilazione esporremo alcune questioni più specifiche relative alla progettazione di un corpus parallelo e presenteremo la nostra proposta di ampliamento.

Nel secondo capitolo si descriverà nel dettaglio il lavoro effettivamente svolto per l'ampliamento del corpus parallelo russo-italiano. Esporremo le nostre considerazioni relative ai criteri di selezione dei testi da noi adottati, alle dimensioni, ai generi testuali scelti e infine presenteremo l'elenco dei testi inseriti con le relative informazioni su autori, traduttori, numero di parole, anno di pubblicazione e genere testuale.

Dopo aver descritto il nostro corpus e il progetto per un suo ulteriore sviluppo nel futuro, presenteremo, nella seconda parte del lavoro, alcune ricerche da noi svolte sulla base dei dati disponibili online dopo il primo ampliamento del CP italiano-russo del NKRJa, risalente a dicembre 2015. Per quanto riguarda la scelta degli aspetti da analizzare, in primo luogo abbiamo voluto differenziare i tipi di analisi per mostrare le potenzialità del corpus parallelo italiano-russo, in secondo luogo, volevamo esplorare aspetti particolarmente interessanti per lo studio contrastivo della lingua russa e italiana. La riflessione su queste due esigenze ci ha portato a formulare i seguenti temi di ricerca: i prefissi verbali della lingua russa e la loro resa in italiano; la costruzione causativa italiana *fare + infinito* e il confronto con il causativo russo; l'analisi comparata delle traduzioni italiane del racconto di N. Gogol' *Il Cappotto*.

I primi due studi, uno sulla prefissazione verbale russa (cap. 3) e l'altro sul valore causativo del verbo *fare* in italiano (cap. 4) sono stati scelti perché riguardano due

strumenti preferenziali¹⁰, del russo il primo e dell'italiano il secondo¹¹, relativamente alla stessa parte del discorso, ossia il verbo. Abbiamo inoltre deciso di condurre la prima analisi partendo dal russo, ossia selezionando un sotto-corpus di originali russi tradotti in italiano, e la seconda dall'italiano, lavorando quindi con un sotto-corpus di originali italiani e traduzioni russe, in modo da vedere in che modo una categoria preferenziale viene resa in una lingua straniera che manca di tale preferenzialità. Sarà così possibile evidenziare eventuali strategie compensative o perdite di significato.

Il terzo capitolo è dunque dedicato all'analisi di un gruppo di verbi prefissali russi: i verbi di memoria con base *pomnit'*, e in particolare *vspomnit'-vspominat'*, *zapomnit'-zapominat'*, *pripomnit'-pripominat'*. Per via delle differenze tipologiche tra russo e italiano, quest'ultimo non dispone di equivalenti traduttivi che possano esprimere con un'unica unità lessicale tutti i tratti semantici di ciascuno dei verbi russi considerati. Inoltre, spesso i dizionari monolingui non chiariscono a fondo le sottili differenze semantiche tra questi verbi, mentre i dizionari bilingui russo-italiano propongono semplicemente *ricordare* come primo traduttivo di tutte le varianti russe, sebbene questi verbi siano dotati di prefissi che conferiscono loro tratti semantici specifici, rendendoli quindi non interscambiabili in contesti diversi. Mediante l'analisi degli esempi estratti dal nostro corpus verranno pertanto identificate le strategie impiegate dai traduttori dal russo in italiano per renderne la semantica, partendo dall'ipotesi che, usando *ricordare* come unico equivalente traduttivo, si possa verificare una perdita di informazione nel testo d'arrivo. Verificheremo questa ipotesi presentando i casi di effettiva perdita di informazione. L'analisi permetterà inoltre di chiarire alcuni aspetti della semantica dei verbi di memoria russi.

Il quarto capitolo tratta invece il verbo *fare* nella costruzione *fare* + infinito: dopo una breve rassegna delle funzioni che questo verbo può svolgere nella lingua italiana, ci

¹⁰ Nel volume *La Comunicazione verbale*, Rigotti e Cigada (2004: 139-140) definiscono una “manifestazione preferenziale quella più tipica, immediata, naturale all'interno di ciascuna lingua storico-naturale, che spesso è anche quella statisticamente più frequente”. Come notano i due autori, la preferenzialità caratterizza quelle che loro chiamano strutture intermedie (lessico, sintassi, ordine delle parole, intonazione, morfologia). Nel lessico, ad esempio, ogni parola polisemica avrà un significato preferenziale. Nella morfologia, invece, alcune desinenze saranno preferenziali nel manifestare una categoria, ad esempio -a per il genere femminile in italiano.

¹¹ In “*Aspekty jazykovej kartiny mira: ital'janskij i russkij jazyki*” (2003: 89-90) [*Aspetti del quadro linguistico del mondo: l'italiano e il russo*] Ju. A. Rylov propone un confronto tra il sistema verbale delle due lingue, evidenziando l'uso italiano del verbo *fare*, la cui dominanza sistemica viene paragonata all'impiego della prefissazione verbale nella lingua russa. Anche a questo studio dobbiamo la nostra decisione riguardo agli aspetti da analizzare.

concentreremo su quella causativa. Inizialmente presenteremo lo stato della ricerca sui verbi causativi e i principali studi ad essi dedicati, in particolare nella russistica e nell'italianistica. In seguito, esamineremo gli esempi di *fare* causativo tratti dal corpus parallelo e le rispettive traduzioni in russo, con lo scopo di identificare alcune strategie di resa dominanti e di chiarire alcuni aspetti semantici e funzionali della costruzione causativa italiana e russa.

Il quinto capitolo, infine, presenta la terza ricerca da noi svolta per testare il nuovo corpus parallelo ampliato ed è dedicato all'analisi di un'unica opera - “*Šinel*” (*Il Cappotto*) di Nikolaj Gogol' - e al confronto tra quindici delle sue traduzioni italiane, disponibili nel corpus. L'analisi si concentrerà sulla resa in italiano del linguaggio del protagonista, Akakij Akakievič, ricco di particelle, preposizioni e intercalari “privi di senso”, come afferma la voce narrante del racconto. Dopo una presentazione delle traduzioni italiane dell'opera, esamineremo alcuni aspetti linguistici della parlata di Akakij Akakievič, in particolare i segnali discorsivi e l'ellissi sintattica. A differenza delle prime due analisi, incentrate su un tema linguistico specifico, qui si intende innanzitutto evidenziare come la possibilità di consultare simultaneamente in un corpus parallelo più traduzioni di uno stesso originale, realizzate in diversi periodi, arricchisca ulteriormente le possibilità che questo strumento offre alla ricerca.

Per quanto concerne la metodologia delle ricerche sopra menzionate è necessaria qualche precisazione. Come vedremo nel primo capitolo, le possibilità di svolgere ricerca linguistica sulla base dei corpora sono innumerevoli. Anche il modo di utilizzare i corpora può variare a seconda dello scopo e del tipo di ricerca. In particolare, vige un dibattito, o quantomeno una distinzione, tra due approcci, definiti rispettivamente *corpus-based* e *corpus-driven*. Come infatti precisa Gilquin (2010: 7-9)¹², i dati di un corpus possono essere considerati in due modi: dal basso verso l'alto, ossia in modalità *corpus-driven* (CD) e dall'alto verso il basso, ossia *corpus-tested* (CT) o *corpus-based* (CB). Altenberg e Granger (2002: 15)¹³ chiariscono un possibile equivoco su quest'ultimo termine (*corpus-based*), che viene impiegato in due modi: uno, più ampio, per indicare qualsiasi indagine che faccia uso di corpora per reperire dati, e uno, più ristretto e opposto a CD,

¹² Gilquin G. (2010). *Corpus, cognition and causative constructions*. Amsterdam, Philadelphia: John Benjamins Publishing.

¹³ Altenberg B., Granger S. (2002). *Recent trends in cross-linguistic lexical studies*. In Altenberg B., Granger S. (a cura di), *Lexis in contrast*. Amsterdam: John Benjamins, pp. 3-48.

comprendente studi che, partendo da un modello, da una teoria o da un'ipotesi, usano il corpus per confutarla, arricchirla o confermarla.

Il metodo CD è di tipo induttivo: i dati di un corpus vengono usati per formulare ipotesi o elaborare una teoria; niente è supposto a priori e i dati sono osservati per ottenere “*new insights into language*”¹⁴ (Granger 2003: 133)¹⁵ o, nel caso dei corpora paralleli, “*to discover types and degrees of cross-linguistic correspondence and to arrive to theoretical statements*”¹⁶ (Altenberg, Granger 2002: 15). Mancando una teoria di riferimento alla base dell'analisi, tale approccio viene spesso accusato di essere superficialmente descrittivo. Tuttavia, i risultati di un approccio CD possono favorire nuovi quesiti e, possibilmente, sono uno stimolo ad elaborare nuove teorie. Al contrario, dal momento che l'approccio CB/CT consiste nel verificare un'ipotesi formulata a priori grazie all'introspezione e alla riflessione linguistica, o nel testare una teoria linguistica preesistente, in questo caso si corre il pericolo di lasciare che i preconcetti influiscano sulla ricerca, in particolare sul modo di interrogare il corpus e sui dati raccolti. Infatti vi è spesso la tentazione di scartare le evidenze che non confermano le nostre ipotesi iniziali.

Secondo McEnery et al. (2006: 8-11)¹⁷ i linguisti CD non prestano particolare attenzione al bilanciamento e alla rappresentatività¹⁸ dei corpora utilizzati, disdegnano l'annotazione, e dunque non considerano categorie linguistiche già codificate. Infine, mentre l'approccio CD, decisamente più radicale, pretende di promuovere un paradigma completamente nuovo entro cui descrivere la lingua, l'approccio *corpus-based* non si pone su questo piano, ma è piuttosto una metodologia usata in tutte le diverse branche della linguistica.

Le nostre ricerche per testare il corpus, espone in questa dissertazione, sono di tipo CB, e dunque hanno come punto di partenza un quadro teorico di riferimento sul tema indagato, ma, al contempo, cercano sempre un compromesso tra considerazione delle teorie di partenza e osservazione dei dati. Si devierà infatti parzialmente dall'approccio CB nel momento in cui si passerà alla pura osservazione degli esempi per

¹⁴ “Nuove intuizioni sulla lingua” (dove non indicato diversamente, la traduzione è nostra).

¹⁵ Granger S. (2003). *The corpus approach: a common way forward for contrastive Linguistics and Translation Studies?* In Granger S., Lerot J., Petch-Tyson S. (eds.), *Corpus-based Approaches to Contrastive Linguistics and Translation Studies*. Rodopi: Amsterdam, New York.

¹⁶ “Per scoprire tipi e gradi di corrispondenze interlinguistiche e formulare assunti teorici”.

¹⁷ McEnery T., Tono Y., Xiao R., (2006). *Corpus-based Language Studies: An Advanced Resource Book*. New York: Routledge.

¹⁸ Rappresentatività e bilanciamento sono due caratteristiche imprescindibili di un corpus di qualità. Data la loro importanza, i due concetti saranno più ampiamente discussi nel capitolo 1 (cfr. par. 1.3.1.3. e 1.8.4).

ricavarne, ove possibile, nuove ipotesi interpretative. In questa fase esplorativa (evidente perlopiù nel capitolo 4), si può dire che prevarrà un approccio CD, seppur in forma limitata.

CAPITOLO 1

LA LINGUISTICA DEI CORPORA E LE SUE APPLICAZIONI: PER UN AMPLIAMENTO DEL CORPUS PARALLELO ITALIANO-RUSSO

*“It has always seemed to me,
ever since I first tried
to become a grammarian,
that grammar was a
subject with too much theory
and too little data”.*
(Halliday 1992: 61)

1.1. La linguistica dei corpora: aspetti teorici e spunti critici

La linguistica dei corpora (LC) può essere definita come un approccio allo studio della lingua, grazie al quale fenomeni linguistici – e non¹⁹ – vengono osservati e analizzati attraverso uno o più corpora di testi. Per “corpus” si intende un insieme di testi in formato elettronico elaborabili da un software.

La LC è un ambito eterogeneo, caratterizzato da numerose distinzioni prettamente concettuali e metodologiche, dal momento che non esiste un unico modo per osservare e impiegare i dati raccolti (McEnery, Hardie 2012:1). Gilquin (2010: 5-6) ricorda in maniera esemplare il titolo della ventiseiesima²⁰ conferenza ICAME²¹ “*Corpus Linguistics: methodology or sub-field*” (Linguistica dei corpora: metodo o sotto-disciplina?), dimostrando come a quarant’anni dalla comparsa del primo corpus elettronico non vi fosse ancora chiarezza relativamente allo status della LC. La maggior parte degli studiosi presenti al convegno sembrava propensa a escludere l’idea di un vero e proprio “*sub-field*” e a considerarla dunque un metodo, ma lamentava, allo stesso tempo,

¹⁹ Molti esempi sono riscontrabili nella sociolinguistica (P. Baker 2010), nel campo dell’acquisizione linguistica dei bambini e in campo sociologico (per i riferimenti precisi cfr. McEnery, Wilson 2001: 3-5).

²⁰ Anno 2005.

²¹ Acronimo di “*International Conference on Advances in Mechanical Engineering*”.

una certa insoddisfazione anche nei confronti di quest'ultima definizione. La LC, infatti, è più di un insieme di strumenti. E, come afferma Teubert (2007: 62), non è sufficiente lavorare con un corpus per essere considerato un linguista dei corpora. In questo senso, è fondamentale la distinzione di Mukherjee (riportata da Gilquin in 2010: 6) tra ricerche "*corpus-aware*" e ricerche "*corpus-based*", per cui, nel primo caso, un ricercatore può essere consapevole dell'esistenza dei corpora e farne uso per le sue ricerche, senza per questo abbracciare interamente la filosofia della LC.

Il termine "approccio", impiegato per definire la LC nelle primissime righe di questo lavoro, sembra essere il ragionevole compromesso tra chi la considera un metodo, e chi invece rifiuta questa definizione e conferisce alla LC lo status di disciplina a tutti gli effetti. McEnery e Hardie (2012: 147) attribuiscono quest'ultima visione alla corrente Neo-Firthiana della LC (dal nome del linguista John Rupert Firth), sostenuta in primo luogo da Sinclair, e portata avanti oggi in particolare da Stubbs, Tognini-Bonelli e Teubert²². La volontà dei Neo-firthiani di considerare la Linguistica dei corpora una disciplina a sé stante è in parte legata al fatto che la nascita della LC in generale, risalente agli anni Sessanta, si distingue fin da subito per una base teorica ben precisa. (McEnery, Wilson 2001) mettono inoltre in luce che il suo emergere è legato alla grammatica generativa di Noam Chomsky, non solo perché le critiche dei generativisti hanno avuto una parte considerevole nel definire la LC, ma soprattutto perché, come ricorda Stubbs (1996), la LC ha fin da subito dato voce alla necessità di usare per la ricerca dati linguistici tratti dall'uso, superando la riflessione teorica di Chomsky, fondata su esempi inventati o sulla mera intuizione dei parlanti nativi. D'altronde, anche prima dell'avvento delle teorie chomskiane gran parte della linguistica si basava sullo studio di testi autentici. McEnery e Wilson, ad esempio, parlano di una "*early corpus linguistics*" ("linguistica dei corpora primitiva") (McEnery, Wilson 2001: 2); come ricorda invece Meyer nella prefazione al volume *English Corpus Linguistics* (2002), tra il 1909 e il 1949 Otto Jespersen aveva usato, per redigere la sua grammatica²³, brani di autori tra cui Chaucer, Shakespeare, Swift e Austin, ossia esempi tratti dalla lingua letteraria di quell'epoca. Il riscontro delle

²² Cfr. Tognini-Bonelli (2001: 1), dove si afferma che "la linguistica dei corpora ha uno status teoretico" o Teubert (2005: 2) che in un saggio intitolato "*My version of Corpus Linguistics*" definisce la LC un "approccio teoretico allo studio della lingua". Non vi è infatti nulla di nuovo, afferma il linguista, nel fare uso di dati empirici autentici per confermare le nostre ipotesi. La vera sfida della LC è quella di "descrivere ciò che non può essere spiegato". Come visto nell'introduzione del nostro lavoro, anche il modo di osservare i dati può variare, e i neo-firthiani tendono a prediligere quello che viene definito approccio "*corpus driven*".

²³ "*A Modern English Grammar on Historical Principles*".

teorie linguistiche mediante dati tratti dall'uso, non è dunque una novità degli ultimi sessant'anni (cfr. par. 1.2.), tuttavia, la riflessione metodologica su questo punto sarà posta in modo consapevole dalla LC, attorno agli anni Sessanta, grazie alla comparsa del primo corpus elettronico, e in contrapposizione al generativismo.

La LC ha dimostrato che l'intuizione del parlante, da un lato, non è in grado di individuare la totalità dei fenomeni presenti in una lingua e, dall'altro, può condurre a risultati che non riflettono l'uso effettivo (Sinclair, Renouf 1991: 100-105). Hunston ricorda che l'intuizione del parlante nativo tradisce il ricercatore in almeno quattro ambiti: collocazione, prosodia semantica, frequenza e fraseologia (Hunston 2002: 20-24). La LC ha quindi introdotto quelli che Biber, Conrad e Rippen definiscono "gli studi sull'uso", volti ad investigare gli strumenti adottati dai parlanti nella produzione linguistica e a trarre conclusioni sul funzionamento della lingua sulla base di testi autentici e contestualizzati. Un corpus di testi, se sufficientemente voluminoso²⁴, può fare luce su ciò che è tipico e ricorrente nella lingua, mentre l'intuizione dei parlanti tende spesso a ricordare fenomeni inusuali (Biber et al. 1998: 3). È così che negli anni Sessanta un gruppo di linguisti britannici, primi fra tutti Firth, Halliday e Sinclair, cominciarono a fondare i loro studi su una serie di principi che avrebbero aperto la strada allo sviluppo della LC, all'interno della quale i neo-firthiani si distinguono per la loro posizione radicale, che rifiuta una definizione di LC nei termini di mero metodo di analisi linguistica, arrivando a scardinare la classica distinzione tra dato e teoria e affermando, al contrario, che "il corpus è la teoria stessa" (McEnery, Hardie 2012: 147).

Nel nostro lavoro opteremo per una posizione intermedia, che considera la LC come un metodo di ricerca fondato sul dato linguistico e non una teoria, ma che, tuttavia, si rifà ad alcuni principi chiave. Questi ultimi, trattati nel dettaglio in (Stubbs 1996: 23-28), possono essere così riassunti:

- la linguistica è una scienza sociale e applicata;
- lo studio della lingua dovrebbe essere supportato da esempi autentici che riflettano l'uso effettivo;
- il linguista dovrebbe considerare testi integrali e non proposizioni, sintagmi o parole isolate²⁵;

²⁴ Successivamente, in particolare nel paragrafo 1.8.5., si rifletterà sulle dimensioni di un corpus.

²⁵ Ciò ha delle implicazioni dirette sul dibattito relativo all'inclusione in un corpus di estratti o testi interi, discusso nel dettaglio nel paragrafo 1.8.6.

- i testi dovrebbero essere comparati l'uno con l'altro per mezzo di corpora appositamente compilati;
- La concezione dualistica del rapporto tra forma e significato, tipica dello strutturalismo saussuriano e del generativismo chomskiano²⁶, è bandita a favore di una concezione monistica della lingua, per cui forma e significato sono inseparabili.

Si tratta di linee guida divenute poi cruciali per la maggior parte della linguistica britannica e per gli studi sulla lingua inglese, che a partire da quegli anni avrebbero influenzato lo sviluppo di tutta la LC: ricordiamo infatti che, sebbene altre lingue fossero state analizzate fin da subito attraverso l'uso di corpora (cfr. i lavori di padre R. Busa e di A. Juillard, di cui parleremo nel par. 1.2.), è nell'ambito della linguistica inglese – inclusa quella studiata in Paesi non anglofoni – che hanno visto la luce sia i principi appena descritti, sia molti dei concetti chiave della LC²⁷ (McEnery, Hardie 2012: 72)²⁸. Più tardi Stig Johansson (1998: 20-21), impegnato principalmente in studi contrastivi tra inglese e norvegese avrebbe dichiarato:

“Our examples²⁹ have illustrated the close connection between correspondences and context, and between lexical and grammatical choices. Corpora may be significant not only in helping us reveal such complex correspondences; they may even transform our ideas of what a linguistic description should be like.”³⁰

Johansson afferma dunque l'unità fra forma e significato, già nominata da Stubbs fra i principi della LC, e ipotizza come dall'aspetto strumentale della LC, che consente di

²⁶ Ronald Carter, nell'introduzione al volume di Sinclair (2004b: 2) scrive: *“Indeed, Chomsky has asserted on several occasions that ‘grammar is autonomous and independent of meaning’* (Infatti, Chomsky ha dichiarato in diverse occasioni che “la grammatica è autonoma e indipendente dal significato”).

²⁷ Ne è un esempio il concetto di “collocazione”, proposto da Firth e ormai comunemente accettato.

²⁸ I due autori elencano i maggiori centri di sviluppo della LC inglese, ossia la University College London, la Lancaster University, l'Università di Birmingham, l'Université Catholique de Louvain, l'Università di Nottingham e la Northern Arizona University (McEnery, Hardie 2012: 73-91).

²⁹ Nel suo lavoro Johansson presenta alcune analisi contrastive sulla base dell'ENPC (*English-Norwegian Parallel Corpus*).

³⁰ “I nostri esempi hanno illustrato lo stretto legame tra connessioni e contesto, così come tra scelte lessicali e scelte grammaticali. I corpora possono essere significativi non solo nel rivelarci queste corrispondenze così complesse, ma possono addirittura trasformare la nostra idea di descrizione linguistica”.

rivelare le connessioni fra lessico e grammatica, si possa giungere a quello teorico, ossia a una radicale ridefinizione della descrizione linguistica.

A questa concezione della lingua e della sua descrizione, fondata esclusivamente sull'uso, si oppongono i generativisti, primo fra tutti Chomsky, i quali sostengono che la competenza linguistica e l'intuizione del parlante prevalgono sulla produzione linguistica effettiva³¹. Non per nulla la teoria della struttura linguistica a cui Chomsky accenna nella sua prima edizione di *Syntactic Structures* non dovrebbe riferirsi ad alcuna lingua in particolare (Chomsky 1957: 11) e in *Aspects of the theory of syntax* aggiunge che le grammatiche particolari di ogni lingua dovrebbero essere completate da una grammatica universale “*to achieve descriptive adequacy*” (Chomsky 1965: 6).

È innegabile che in situazioni di ambiguità sia il contesto a consentire l'adeguata comprensione di un enunciato (Graffi 2008: 19); tuttavia, se il linguista dei corpora parte da questi “inconvenienti” del processo comunicativo per prediligere un approccio basato sullo studio del testo autentico, prodotto in un determinato contesto storico e sociale, al contrario Chomsky parte dal dato di fatto dell'ambiguità per concentrarsi esclusivamente sulla competenza dei parlanti nativi, piuttosto che sulla loro produzione linguistica. Chomsky è perfettamente consapevole che, da un lato, un enunciato grammaticalmente corretto, ma allo stesso tempo molto complesso, può rientrare nella competenza dei parlanti nativi senza essere riscontrabile nell'uso, e che, dall'altro lato, un enunciato agrammaticale, e dunque apparentemente non riconducibile alla sfera di competenza dei parlanti nativi di una lingua, può essere non solo riscontrato nell'uso, ma anche perfettamente comprensibile e accettabile per il destinatario, soddisfacendo dunque lo scopo comunicativo, ossia la corretta trasmissione e ricezione del senso di un messaggio (Rigotti, Cigada 2004)³². Eppure l'accettabilità non ha nulla a che vedere con la competenza su cui dovrebbe fondarsi la grammatica generativa proposta da Chomsky (Chomsky 1965: 11-12).

Simili dibattiti, molto accesi negli anni Sessanta, non si sono affatto esauriti con il passare degli anni. Aaarts, in un lavoro presentato a ICAME nel 1999, riporta parte di

³¹ Chomsky introduce i concetti di “*competence*” – “*the speaker-hearer's knowledge of his language*” - vs “*performance*” – “*the actual use of language in concrete situations*”, in (1965: 4). I due termini vengono tradotti in italiano con “competenza” e “esecuzione”.

³² I due autori propongono una teoria della comunicazione che vede il senso strettamente legato alla pertinenza di un messaggio, “alla significatività di ciò che viene detto, rispetto alle soggettività coinvolte” (Rigotti, Cigada 2004: 56). Basandosi sulle teorie di Pierce, si afferma che la comunicazione efficace “è quella che cambia (...) che tocca e cambia lo *habit* del destinatario” (Rigotti, Cigada 2004: 55). Tale concezione non coinvolge quindi i concetti di grammaticalità o agrammaticalità.

un'intervista rivolta a Noam Chomsky pochi anni prima, dove alla domanda: *what is your view of modern corpus linguistics?*³³ lo studioso risponde: *"it doesn't exist"(...)*³⁴ (Aarts 2000: 5). E aggiunge: (...) *if you take a look at a corpus things appear possible (...). Everything is possible in language. The question is 'how does it relate to the structure of English?' (...)*³⁵ (Aarts 2000: 6).

La LC, in conclusione, si è fatta strada parallelamente allo sviluppo delle teorie generativiste, rafforzandosi dallo scontro con esse (McEnery, Wilson 2001: 4-10).

Le critiche alla LC, tuttavia, non sono circoscrivibili alla polemica con il generativismo; tra i linguisti dei corpora stessi non mancano quanti, pur prediligendo approcci "*corpus-based*", sono consapevoli dei limiti di questi ultimi. Secondo Susan Hunston (2002: 23), uno degli svantaggi della ricerca linguistica basata sui corpora è che un corpus non fornisce informazioni, bensì esempi. Eventuali conclusioni potranno derivare solamente da un'accurata analisi del ricercatore che farà uso, tra l'altro, della sua intuizione di esperto di una data lingua. Vediamo dunque che anche nella LC l'intuizione del parlante, sottolineata da Chomsky, non è da escludersi completamente nel corso dell'indagine linguistica. In secondo luogo, un corpus non informa circa la correttezza o meno di un enunciato, ma presenta quanto ricorre nell'uso, indipendentemente dalla sua grammaticalità (o sensatezza). Ancora, un corpus non corrisponde alla totalità della lingua. Infine, afferma l'autrice, un corpus offre esempi al di fuori del loro contesto. Quest'ultima affermazione è in realtà opinabile e in contrapposizione a quanto affermato in precedenza, dal momento che, come vedremo, l'interfaccia di molti corpora attuali permette di risalire sia alle informazioni extralinguistiche relative alla produzione di un testo (i cosiddetti metadati, relativi ad anno di pubblicazione, autore, età e via dicendo), sia al frammento di testo allargato in cui è inserita l'unità linguistica che si vuole analizzare. Non sarà mai possibile risalire all'intero testo direttamente dal corpus, ma se un corpus è ben compilato si dovrebbe disporre dei dati necessari per procurare altrove tutte le informazioni richieste dalla ricerca. Infine, chiamiamo in causa ancora una volta Aarts, il quale, probabilmente toccato dal confronto con il padre dei generativisti, nel sopracitato contributo del 1999 invita i linguisti dei corpora a riflettere sulla natura dei

³³ "Come vede la linguista dei corpora contemporanea?".

³⁴ "Non esiste".

³⁵ "Se si osserva un corpus, le cose sembreranno possibili. (...) Tutto è possibile nella lingua. Ma la domanda è: che relazione ha questo con la struttura dell'inglese?".

dati ottenuti da un corpus e su come utilizzarli, aspetto fondamentale nell'ambito delle discussioni sullo status e sulla validità della LC.

Chiarite dunque le motivazioni che hanno portato alla nascita della LC, come pure i principi che la governano, passiamo ora in rassegna i principali autori e i contributi prodotti dalla LC, nonché le fasi del suo sviluppo.

1.2. Sviluppo della LC

In questo paragrafo descriveremo le tappe fondamentali della LC³⁶ e l'avvento di questo approccio in Italia e in Russia, i paesi su cui si concentra il nostro lavoro. Nonostante le critiche, infatti, la LC si è fatta strada nel corso dei decenni diffondendosi in maniera esponenziale.

Fra i primi contributi alla LC, tra gli anni Cinquanta e Settanta (sebbene allora non fosse ancora denominata in questo modo), citiamo i lavori di padre Roberto Busa, il primo a creare un corpus di testi elaborabili da un computer, e di Alphonse Juilland, che definiva il proprio metodo “meccano-linguistica”. I due studiosi contribuirono a gettare le basi per quella che oggi conosciamo con il nome di “Linguistica dei corpora”. Busa fu il primo ad introdurre metodi di analisi computazionale nell’ambito degli studi umanistici grazie al suo *Index Thomisticus*, corpus di 11 milioni di parole in cui raccolse le opere di Tomaso D’Aquino (9 milioni di parole) e alcuni scritti di altri autori (quasi 2 milioni di parole) e di cui si servì per studiare il lessico del santo aquinate (McEnery, Hardie 2012: 37) (Busa 2009: 19). I maggiori contributi di Juilland riguardano invece statistiche lessicali e analisi contrastive, attraverso la creazione dei primissimi corpora multilingui comparabili (McHenery, Wilson 2001: 20-22).

Come già accennato, importantissimo per lo sviluppo della LC fu anche l’influsso di Firth e del suo approccio alla lingua nel contesto innovativo di quel periodo, fondato principalmente sull’idea che l’oggetto dell’analisi linguistica dovesse essere la lingua attestata nell’uso, e che il senso di una data parola dipendesse dal contesto in cui è inserita; da qui la sua famosa affermazione: “*you shall know a word by the company it keeps*”³⁷ (Firth 1957: 11), con la quale viene introdotto il concetto di “collocazione”.

³⁶ Alcuni dei corpora qui citati saranno descritti più dettagliatamente nei successivi paragrafi.

³⁷ “Comprenderai una parola considerando cosa accompagna”.

L'inizio degli anni Sessanta segnò la data di creazione dei primi corpora elettronici per lo studio della lingua inglese: il *Brown Corpus*, progettato e compilato da Francis e Kučera alla Brown University (USA), e quello che poi sarà meglio conosciuto come il *London-Lund Corpus*, progettato da Svartik sulla base di un corpus già esistente presso il centro di ricerca SEU (*Survey of English Usage*), fondato da Quirk nel 1959 alla University College di Londra (McHenry, Wilson 2001: 22-23). Successivamente, nel 1970, l'Università di Lancaster e l'Università di Oslo collaborarono per compilare l'equivalente britannico del *Brown Corpus*, ad oggi conosciuto come *LOB Corpus*.

Altro progetto di notevole importanza nella storia della LC è il *COBUILD*, diretto da John Sinclair presso l'Università di Birmingham. Si tratta del primo dizionario compilato sulla base di un corpus, a cui seguirono alcuni strumenti di consultazione redatti sulla base del medesimo principio: ne sono un esempio la *Cobuild English Grammar*³⁸ e il *Cobuild Dictionary of Idioms*³⁹. L'esperienza del *COBUILD*, la cui prima pubblicazione risale al 1987, segnò una svolta nella lessicografia successiva, dimostrando come i lessicografi non potessero più fondare la compilazione dei loro dizionari sulla propria intuizione di parlanti nativi. *COBUILD* nacque come dizionario per apprendenti di livello avanzato della lingua inglese, ma tutt'ora costituisce un ottimo esempio per la compilazione di materiale lessicografico di qualsiasi tipo. In questo caso, solo i dati empirici di un corpus permisero al team di Sinclair di decidere innanzitutto quali voci includere e quali invece eliminare rispetto ai dizionari precedenti. Inoltre, gli esempi del corpus cominciarono non solo ad essere citati direttamente come esempi illustrativi di una voce, ma anche a chiarire molti aspetti legati ai sensi dei lessemi polisemici, suggerendo quali escludere, perché infrequenti o obsoleti, e a quali dare priorità (Sinclair 2004a: 1, Hanks 2009: 215-218).

Oggi, sull'esempio del *COBUILD*, vengono compilati numerosi dizionari con un approccio interamente *corpus-based*, mentre le nuove edizioni di dizionari precedenti sono riviste alla luce dei maggiori corpora, come esaustivamente descritto in (Hanks 2009), a cui si rimanda per un approfondimento del tema⁴⁰.

³⁸ Sinclair J.M. et al. (1990) *Collins Cobuild English Grammar*. London: HarperCollins.

³⁹ Sinclair, J.M. et al. (1995) *Collins Cobuild Dictionary of Idioms*. London: HarperCollins.

⁴⁰ Interessante anche il contributo di Luisa Giacoma (2012) relativo al recente impegno dei lessicografi nella creazione di dizionari sintagmatici che tengano conto soprattutto della collocazione delle parole, piuttosto che dei singoli lessemi. La disponibilità dei corpora, sottolinea l'autrice, è uno dei fattori che ha reso possibile l'avvento di simili dizionari.

A partire dagli anni Novanta, in particolare con il *British National Corpus*, cominciarono ad essere disponibili i primi grandi corpora, composti da almeno 100 milioni di parole. Al giorno d'oggi se ne contano innumerevoli e gli ingenti volumi di materiale reperibile sul web hanno spesso spinto i linguisti a impiegare il web stesso come corpus. Un esempio è *WebCorp*⁴¹, attivo per la lingua inglese e curato dalla *School of English* e la *City University of Birmingham*. Il suo scopo è acquisire testi direttamente dal web tramite *web crawling* – ossia l'estrazione automatizzata di testi dalla rete - e renderli disponibili attraverso un motore di ricerca appositamente studiato per analisi linguistiche. La differenza tra il motore di ricerca di *WebCorp* e un motore di ricerca sul web ordinario (ad esempio Google) sta nel fatto che quest'ultimo consente l'acquisizione di informazioni generali, il primo invece è focalizzato su dati di tipo linguistico (Kehoe, Renouf 2002, Morley 2006)⁴². Web corpora di questo tipo stanno riscuotendo un consenso e una diffusione sempre maggiori, non solo per la lingua inglese. Il progetto *WaCky*⁴³, ad esempio, comprende tre grandi corpora (formati da più di un bilione di parole ciascuno) sia per la lingua inglese, sia per italiano e tedesco: *ukWaC*, *deWaC* e *itWaC*. Per l'italiano, sembra essere il maggiore corpus esistente al momento. Tra i web corpora per la lingua russa, citiamo invece *RuTenTen*⁴⁴, il *General'nyj Internet-Korpus Russkogo Jazyka*⁴⁵, *Araneum Russicum Maius & Minus*⁴⁶ (contenuto nel progetto *Aranea family*) e *The Russian Web Corpus*⁴⁷ (Khokhlova 2016: 238-239).

Tutti i web corpora citati dispongono di un'annotazione di base. *WaCky*, ad esempio, offre la lemmatizzazione e indica l'appartenenza alle varie parti del discorso. In aggiunta, a differenza di altri web corpora, *WaCky* è stato sottoposto ad un processo di "pulizia" e rifinitura a scopi linguistici (Baroni et al. 2009)⁴⁸.

Infine, nota Borin in un lavoro del 2002, negli anni Novanta cominciò ad emergere un campo quasi a sé stante della LC, ossia la "*parallel corpus linguistics*", fondata

⁴¹ <http://www.webcorp.org.uk/> (gennaio 2017).

⁴² Ulteriori contributi su questo tema sono disponibili alla pagina <http://www.webcorp.org.uk:8080/publications.html> (gennaio 2017).

⁴³ Disponibile al sito <http://wacky.sslmit.unibo.it> (gennaio 2017).

⁴⁴ Accessibile dal portale *Sketch Engine* (<https://the.sketchengine.co.uk/auth/corpora/>) (gennaio 2017).

⁴⁵ <http://www.webcorpora.ru/> (gennaio 2017).

⁴⁶ Per le informazioni sui portali d'accesso e sulle caratteristiche del corpus cfr. http://sketch.juls.savba.sk/aranea_about/ (gennaio 2017).

⁴⁷ Accessibile dal portale *Sketch Engine* (<https://the.sketchengine.co.uk/auth/corpora/>) (gennaio 2017).

⁴⁸ Per un approfondimento cfr. Baroni M, Bernardini S., Ferraresi A., Zanchetta E. (2009). *The WaCky Wide Web: a collection of very large linguistically processed web-crawled corpora. Language Resources and Evaluation* 43 (3): 209-226. Si può accedere al testo dalla pagina: <http://wacky.sslmit.unibo.it/doku.php> (gennaio 2017).

sull'uso di corpora multilingui, definiti in generale “paralleli” (Borin 2002: 1). Questo tema verrà trattato in maniera più esaustiva nelle prossime pagine, in particolare ai paragrafi 1.3.1. e 1.4.

Per quanto riguarda lo sviluppo della LC in Italia, le opinioni divergono: da un lato Prat Zagrebelsky, in un saggio relativo all'introduzione della LC nelle università italiane⁴⁹, parla di un processo lento e soprattutto tardivo rispetto al mondo anglosassone. La studiosa ricorda, sulla base della propria esperienza di docente di lingua inglese, come i linguisti e i docenti italiani avessero assunto un atteggiamento piuttosto distaccato di fronte agli sforzi e ai progressi dei colleghi britannici in questo ambito. Solo alla fine degli anni Ottanta, sotto la supervisione di Guy Aston, alcuni linguisti italiani cominciarono a compilare il *PIXI Corpus*, un corpus di lingua parlata in inglese e in italiano (Prat Zagrebelsky 2005: 204). Prat Zagrebelsky segna dunque gli anni Novanta come momento in cui la LC cominciò a diffondersi ufficialmente anche in Italia, inizialmente in collaborazione con università e centri di ricerca esteri, in seguito in modo indipendente.

Dall'altro lato, Barbera (2013: 10-14) non manca di puntualizzare che non solo uno dei pionieri in assoluto della LC fu un italiano, il già citato padre Roberto Busa, ma anche che l'Accademia della Crusca, in particolare l'omonimo Dizionario, fin da subito utilizzò un approccio *corpus-based*, fondato sull'uso di esempi autentici e sulla convinzione che solo questi ultimi permettessero uno studio della lingua completo e approfondito. Barbera ricorda come Francesco Sabatini stesso, oggi presidente onorario dell'Accademia, abbia più volte sostenuto l'opinione secondo cui la linguistica italiana sia stata da sempre legata alla LC.

Oggi i corpora di lingua italiana parlata e scritta sono numerosi e in continua espansione. L'Accademia della Crusca⁵⁰ e la Società di linguistica italiana⁵¹ offrono un elenco dei maggiori corpora e delle maggiori banche dati per l'italiano accessibili online. Tuttavia, ciò che emerge nel consultare queste risorse è la quasi totale assenza di corpora paralleli comprendenti l'italiano compilati secondo metodi tradizionali. La differenza tra corpora tradizionali e web-corpora (quali il sopracitato *WaCky*), come pure le riflessioni sui vantaggi e gli svantaggi di queste risorse, sono ancor oggi materia di discussione.

⁴⁹ Quaderni del CIRSIL – 4 (2005) – www.lingue.unibo.it/cirsil pp. 203-213 (settembre 2016).

⁵⁰ <http://www.accademiadellacrusca.it/it/link-utili/banche-dati-dellitaliano-scritto-parlato> (gennaio 2017).

⁵¹ http://www.societadilinguisticaitaliana.net/index.php?option=com_weblinks&view=category&id=21%3Acorpora-italiani&Itemid=21 (gennaio 2017).

Pertanto, nel successivo paragrafo riporteremo i punti principali di questo dibattito, chiarendo i motivi della nostra preferenza per i corpora tradizionali.

In Russia la diffusione della LC può dirsi parallela alla nascita del *Corpus Nazionale della Lingua Russa* (NKRJa), accessibile sul web dal 2004⁵². I primi passi verso strumenti linguistici *corpus-based* risalgono tuttavia agli anni Settanta, con la compilazione del *Častotnyj slovar' russkogo jazyka* (Dizionario di frequenza della lingua russa) redatto da L.N. Zazorina, la quale si era basata su di un corpus di un milione di parole, discutendone la rappresentatività, la lemmatizzazione, la normalizzazione e facendone un'analisi grafematica.

Oltre al dizionario di Zazorina, la compilazione del NKRJa è stata preceduta da almeno tre progetti:

- nel 1985, la creazione del *Mašinnyj fond russkogo jazyka* (Fondo computerizzato della lingua russa), mai portato a termine;
- alla fine degli anni Ottanta, l'*Uppsal'skij korpus russkich tekstov* (Corpus di testi russi di Uppsala), un corpus di 1 milione di parole composto da testi scritti appartenenti al genere *fiction* e *non-fiction*;
- infine, il *Komp'juternyj korpus tekstov russkich gazet konca XX veka* (Corpus computerizzato di giornali russi della fine del XX secolo), creato tra il 2000 e il 2002 presso la facoltà di lingue dell'Università Statale di Mosca (MGU) e comprendente testi tratti dalla stampa russa (Zacharov 2013: 1-4).

Dal 2004, con l'avvio del NKRJa, anche nella Federazione Russa la LC comincia a svolgere una parte importante all'interno della russistica. In generale, la diffusione di corpora nella slavistica ha avuto luogo proprio in questi anni: tra la fine degli anni Novanta e il 2007 hanno fatto la loro comparsa il vastissimo *Corpus Nazionale della Lingua Ceca* (CNC)⁵³, il *Corpus Nazionale della Lingua Bulgara*⁵⁴ e il *Corpus Nazionale della Lingua Polacca*⁵⁵. Si tratta, in tutti i casi, di corpora monolingui, ad eccezione del CNC, che raggruppa una serie di risorse differenti, tra cui corpora in altre lingue e corpora

⁵² www.ruscorpora.it (gennaio 2017).

⁵³ <http://www.korpus.cz/> (gennaio 2017).

⁵⁴ <http://search.dcl.bas.bg/> (gennaio 2017).

⁵⁵ <http://nkjp.pl/> (gennaio 2017).

paralleli (questi ultimi rientrano nel progetto *InterCorp*, a cui ritorneremo brevemente nel paragrafo 1.6.)⁵⁶.

1.3. Definizione e tipologie dei corpora

È ora utile considerare più nel dettaglio la definizione di “corpus” presentata brevemente all’inizio di questo lavoro, dal momento che il concetto non è omogeneo e pertanto può essere soggetto a diverse interpretazioni.

In LC per “corpus” si intende un insieme di testi in formato elettronico raccolti a scopo di ricerca e secondo precisi criteri. Nei documenti di EAGLES (*Expert Advisory Group on Language Engineering Standards*)⁵⁷, un’iniziativa della Commissione Europea avviata nel 1993 con lo scopo di definire alcuni standard nel settore della linguistica computazionale e del trattamento automatico del linguaggio (settori complessivamente raggruppabili sotto la comune denominazione di *Natural Language Processing*, NLP), si legge che un corpus è:

*“a collection of pieces of language that are selected and ordered according to explicit linguistic criteria in order to be used as a sample of the language”*⁵⁸.

E si aggiunge:

*“Words such as ‘collection’ and ‘archive’ refer to sets of texts that do not need to be selected, or do not need to be ordered, or the selection and/or ordering do not need to be on linguistic criteria. They are therefore quite unlike corpora”*⁵⁹.

⁵⁶ Per una rassegna dettagliata dei corpora delle lingue slave cfr. Reznikova (2009).

⁵⁷ Le informazioni sul progetto sono al sito: <http://www.ilc.cnr.it/EAGLES96/intro.html> (gennaio 2017).

⁵⁸ “Una collezione di estratti di lingua selezionati ed ordinati secondo criteri linguistici precisi, al fine di fungere da esempio di quella lingua”.

⁵⁹ “Termini come ‘collezione’ o ‘archivio’ si riferiscono a testi che non devono essere selezionati o ordinati, o per cui la selezione e il riordino non risponde a criteri linguistici precisi. Pertanto si distinguono dai corpora”.

I corpora compilati tradizionalmente e le collezioni di testi si distinguono dunque dai corpora così come li definisce la LC, in quanto sono raccolte di testi assemblate o senza specifici criteri o secondo criteri non necessariamente linguistici. Come ricorda anche Meyer (2002), facendo appello alla necessità per un linguista di definire un corpus in maniera adeguata, non tutte le raccolte di testi costituiscono un corpus. Zanettin (2012: 153), discutendo i vari tipi di corpora paralleli, distingue tra corpora creati manualmente con l'unico scopo di condurre ricerche linguistiche e corpora assemblati sulla base di fonti multilingue specifiche (come ad esempio *Europarl*, a cui accenneremo in seguito). Sembra dunque opportuno aggiungere anche questi ultimi ai tipi di corpora già presi da noi in considerazione, distinguendo così tre macro gruppi:

- (1) Corpora tradizionali
- (2) Web corpora
- (3) Collezioni di testi prodotti in un contesto specifico e raccolti su una piattaforma dotata di un motore di ricerca.

1) Per corpus tradizionale si intende “una collezione di brani in una data lingua, selezionati e ordinati secondo criteri linguistici precisi, al fine di fungere da esempio di quella lingua” (cfr. definizione di EAGLES). La compilazione di corpora tradizionali, come vedremo, implica una serie di operazioni che rendono il procedimento lungo e complesso: selezione dei testi (in base allo scopo del corpus e a criteri precedentemente stabiliti), acquisizione dei testi (talvolta già in formato elettronico, talvolta in formato cartaceo e quindi da scansionare), conversione nel formato digitale ottimale e ripulitura da eventuali errori (specie in presenza di simboli diacritici), allineamento di testi fonte e testi di arrivo nel caso di corpora paralleli e, infine, annotazione. La chiave di questo processo si trova soprattutto nei criteri di selezione dei testi, grazie ai quali i fruitori del corpus saranno sempre in grado di conoscere l'origine dei brani sottoposti ad analisi e tutte le informazioni extralinguistiche ad essi riferite.

2) La compilazione di Web corpora si basa invece, come già detto, sull'estrazione automatizzata di testi disponibili in rete e presenta senza dubbio numerosi vantaggi: per prima cosa i testi sono facili da ottenere e il processo della loro acquisizione è molto più veloce ed economico che nel caso dei corpora tradizionali; in secondo luogo riflettono in

maniera più accurata i cambiamenti di una lingua storico-naturale (Křen 2009); infine, consentono di trovare e analizzare parole o locuzioni più rare (ad esempio proverbi e modi di dire), poco ricorrenti in corpora tradizionali di dimensioni ridotte (Benko, Zacharov 2016). Zanettin, d'altro canto, prediligendo marcatamente la compilazione di corpora tradizionali, riduce l'unico vantaggio dei Web corpora alla loro ingente dimensione (Zanettin 2012: 56). Vero è che, a seconda delle finalità di una ricerca, è possibile prediligere l'uso di un Web corpus non solo per questioni di volume, ma anche perché il fenomeno ricercato potrebbe non ricorrere tra i generi letterari compresi in un corpus tradizionale (Lüdeling et al. 2005).

Per quanto riguarda gli svantaggi, diversi autori, tra cui Křen (2009), Benko e Zacharov (2016), ne evidenziano diversi: la mancata accuratezza dei metadati (se non, in alcuni casi, la totale assenza); una maggiore possibilità di errori, dal momento che il processo di ripulitura "*post-crawling*" è eseguito automaticamente; e infine, proprio in contrapposizione a Lüdeling, un certo sbilanciamento per quanto riguarda i generi testuali: i testi tratti dal web tendono a non rappresentare affatto alcuni generi, ad esempio quello letterario. Per la precisione, quest'ultimo fattore avrebbe la particolarità di essere un vantaggio e uno svantaggio allo stesso tempo, a seconda dello scopo della ricerca condotta sul corpus. Un ulteriore svantaggio dei web corpora, particolarmente pertinente al nostro lavoro, è l'impossibilità di compilare corpora paralleli basandosi esclusivamente sul *web crawling*, sia per il numero esiguo - verosimilmente non sufficiente - di testi originali e traduzioni disponibili sul web, specie se le due lingue coinvolte sono diverse dall'inglese, sia per eventuali problemi legati alla qualità delle traduzioni, spesso eseguite da software di traduzione automatica, oppure non da professionisti.

3) Il terzo macro gruppo è composto da collezioni di testi redatti in un ambito specifico e ristretto (la Commissione Europea o il Parlamento Europeo, ad esempio) e rese poi disponibili per gli usi legati alla LC. Normalmente, i testi vengono acquisiti e annotati dai gestori di un portale che raccoglie diverse collezioni di questo tipo (cfr. *Opus*, par. 1.6.), e resi poi accessibili attraverso un motore di ricerca analogo a quello dei corpora tradizionali. Come per i Web corpora, il principale vantaggio di tali risorse sta nella facilità e velocità di acquisizione dei testi. A differenza dei Web corpora, tuttavia, è sempre nota la fonte da cui derivano. Gli svantaggi riguardano invece la mancanza di informazioni relative agli autori - senza contare che all'interno di istituzioni internazionali

i redattori di documenti possono non essere madre lingua - e la restrizione dei generi letterari, quasi esclusivamente di tipo giuridico o amministrativo. Per quanto riguarda i corpora paralleli all'interno di questo gruppo, si aggiunge lo svantaggio di avere spesso testi tradotti esclusivamente dall'inglese e mai da altre lingue.

Chiarite le differenze tra questi diversi tipi di strumenti, si precisa che, d'ora in avanti, tutte le riflessioni relative alle caratteristiche specifiche di un corpus e ai criteri di compilazione riguarderanno corpora di tipo tradizionale.

Come accennato, nel processo di compilazione di un corpus importanza fondamentale è attribuita alla finalità, ma nonostante la sua rilevanza, come vedremo anche in seguito, nessuno creerà un corpus sulla base di finalità ristrette (per es. un corpus di proposizioni relative per studiare le proposizioni relative, o un corpus di proverbi per studiare i proverbi); un linguista compilerà piuttosto un corpus di testi su cui potranno essere condotti diversi tipi di analisi linguistica, tra cui l'analisi delle proposizioni relative e/o dei proverbi (Meyer 2002: xi-xii). La finalità sarà piuttosto una guida essenziale per la scelta dei testi da includere. Tuttavia esistono anche corpora, che definiremo *ad hoc*, ossia creati appositamente per lo studio di un aspetto particolare della lingua. Si tratta, tuttavia, di corpora a sé stanti, solitamente di esigue dimensioni e ad uso personale del ricercatore, difficilmente messi a disposizione del grande pubblico. Il web offre diversi strumenti che rendono possibile la creazione di simili corpora direttamente sul proprio computer e consentono, soprattutto, l'elaborazione dei testi in modo da condurre delle analisi. Tra i software più conosciuti e disponibili gratuitamente citiamo AntConc di Laurence Anthony, accessibile dal suo sito⁶⁰. Altri software sono disponibili in rete, ma richiedono il pagamento di una quota per il loro utilizzo⁶¹.

1.3.1. Le caratteristiche di un corpus

In linea teorica e ideale, le caratteristiche oggi considerate come necessarie per qualsiasi tipo di corpus tradizionale sono tre: formato elettronico, annotazione, bilanciamento/rappresentatività. Ciononostante, se la prima caratteristica è ovvia e universale, le restanti hanno diversi livelli di realizzazione, e nel caso della

⁶⁰ <http://www.laurenceanthony.net/> (gennaio 2017).

⁶¹ Citiamo ad esempio, ParaConc, per l'allineamento di testi in diverse lingue: <http://paraconc.com/> (gennaio 2017).

rappresentatività (non sempre facile da raggiungere) il criterio fissato per ottenerla può risultare tutt'altro che univoco.

1.3.1.1. Il formato elettronico

I testi raccolti in un corpus, prodotti manualmente o stampati, innanzitutto dovranno essere letti da uno scanner a fibra ottica e poi convertiti nel formato elettronico adeguato⁶². Il rapporto tra LC e tecnologie informatiche è dunque strettissimo. Anzi, proprio la tecnologia ne ha reso possibile lo sviluppo dal momento che, come ricordano Lenci, Montemagni e Pirelli, “Il computer è a servizio dell'indagine linguistica” (Lenci et al. 2007: 12). Il confronto e l'analisi di testi cartacei richiede tempo e sforzi considerevoli, senza contare la possibilità di incorrere in errori ed omissioni, influenzando così sui risultati non solo dal punto di vista quantitativo, ma anche qualitativo. Il rapporto qualitativo-quantitativo sta invece alla base della linguistica dei corpora e uno dei suoi risultati più innovativi è che l'analisi linguistica assistita da un computer permette di osservare non solo un ingente volume di dati, ma anche ciò che in precedenza risultava inaccessibile all'occhio umano e all'intuizione linguistica. Molte delle indagini e scoperte in campo linguistico degli ultimi decenni sono dunque una novità della LC (Stubbs 1996).

Citiamo infine Sinclair (2004b: 189) il quale, nel chiarire la differenza tra corpus e testo, solleva un'importante questione legata alla differenza tra LC e un approccio “più tradizionale”, basato sull'analisi manuale di testi:

“The essence of the corpus as against the text is that you do not observe it directly; instead you use tools of indirect observation, like query languages, concordances [...]”⁶³.

1.3.1.2. L'annotazione

L'annotazione è uno degli strumenti che differenzia un corpus da un semplice insieme di testi, rendendolo particolarmente adatto ad analisi linguistiche piuttosto elaborate. Per

⁶² Dal 1993 esistono una serie di raccomandazioni sviluppate dalla *Text Encoding Initiative* (TEI recommendations) e relative a formato e principi di codifica (Johansson 1993: 203-205). Per un approfondimento cfr. la pagina web <http://www.tei-c.org/Guidelines/> (gennaio 2017).

⁶³ “L'essenza di un corpus, a differenza del testo, è che non lo osservi in maniera diretta, ma attraverso una serie di strumenti di osservazione indiretta, come linguaggi di interrogazione, concordanze [...]”.

annotazione si intende l'aggiunta di informazioni di carattere linguistico ed extralinguistico ai testi che compongono un corpus. Più precisamente, le informazioni extralinguistiche riguardano l'intero testo, o estratto, mentre le informazioni linguistiche si legano ad ogni singola forma di parola, indicandone, ad esempio, l'appartenenza a una data classe del lessico, il numero, il genere e così via, a seconda dei casi. Ricordiamo che l'annotazione non risponde a degli standard universali; i compilatori di ciascun corpus decideranno, in base ai loro scopi, quali informazioni aggiungere ai propri testi e in che modo presentarle, e quale margine di errore consentire. Pertanto, le informazioni generali relative alla costituzione di un corpus dovrebbero sempre includere dei dati sul metodo di annotazione adottato (Leech in Garside et al. 1997: 2-6).

L'annotazione non ha caratterizzato fin da subito i corpora elettronici - il primo corpus annotato e con un margine di errori relativamente esiguo (pari circa al 2%) risale infatti alla fine degli anni Settanta⁶⁴ (Leech in Garside et al. 1997: viii, 8) -, ma considerate le risorse informatiche attualmente esistenti e visti gli enormi vantaggi apportati da un corpus annotato all'indagine linguistica, oggi è difficile immaginare la compilazione di grandi corpora che non includano questa caratteristica.

Come si accennava, un corpus può disporre di diversi livelli di annotazione. Senza entrare nel dettaglio⁶⁵, diremo che i principali tipi di annotazione sono: l'annotazione metatestuale (informazioni extralinguistiche relative al testo), ortografica (conosciuta anche come *mark-up*, ossia il riconoscimento di segni quali punteggiatura o lettere maiuscole) morfologica, semantica e sintattica; più rare e in via di sviluppo sono l'annotazione stilistica e pragmatica, come pure quella discorsiva; e, infine, per i corpora di lingua parlata può aversi un'annotazione prosodica (Leech in Garside et al. 1997: 12). Leech (1997: 15) ricorda inoltre l'esistenza di altri tipi di annotazione, legati a corpora di carattere specifico. Ne è un esempio il rilevamento degli errori nei cosiddetti "*learner corpora*" (cfr. par. 1.3.2.).

Resta inteso che più ricca e precisa è l'annotazione, maggiore è il valore di un corpus. Pertanto sono auspicabili miglioramenti e sviluppi anche in questo settore della LC.

⁶⁴ Dopo un lavoro durato anni, venne finalmente completata la correzione degli errori di annotazione del Brown Corpus, aggiunta in via del tutto sperimentale dopo la costituzione del Corpus stesso (Leech in Garside et al. 1997: 8).

⁶⁵ Nel par. 1.5. verranno descritti più dettagliatamente i tipi di annotazione di cui dispone il NKRJa.

1.3.1.3. Rappresentatività e bilanciamento dei testi

I concetti di bilanciamento e rappresentatività spesso sono trattati insieme in quanto strettamente collegati⁶⁶: un corpus generale dovrebbe *rappresentare*, idealmente, la lingua parlata e scritta dai nativi di una data comunità linguistica, e ciò richiede che la quantità di testi inseriti sia notevole e le tipologie testuali varie e *bilanciate*, ossia i diversi generi dovrebbero essere rappresentati da sotto-corpora all'incirca della stessa entità anche se, come vedremo, il *bilanciamento* non si riduce esclusivamente a tale criterio.

Il concetto di rappresentatività può apparire limitato e vago, e nel corso dei decenni non è stato esente da critiche. Noam Chomsky osserva, ad esempio, che un corpus di testi è un oggetto finito, mentre la conoscenza umana è sconfinata e capace, in linea di principio, di produrre frasi o testi grammaticalmente corretti sempre nuovi, e quindi assenti nei corpora. La critica chomskiana è sicuramente valida: il linguista dei corpora analizza i dati raccolti in un corpus, escludendo ciò che non vi è contenuto. Se in un corpus generale, comprendente centinaia di milioni di parole, un termine o un sintagma non compaiono, il ricercatore potrebbe non occuparsene, o classificare il fenomeno come raro o poco ricorrente. Si tratta di una scelta opinabile, ma sicuramente coerente con gli scopi della linguistica dei corpora, primo fra tutti quello di studiare la lingua autentica ed effettivamente in uso. In questo senso, la LC non dovrebbe avanzare la pretesa di definire l'essenza della lingua: un corpus, mostrando i fenomeni più ricorrenti, permette al massimo di fare generalizzazioni su quanto è tipico in una lingua (Stubbs 1996) e, per esclusione, su quanto è atipico. Come ha ricordato Sinclair in più occasioni, un corpus non corrisponde alla totalità della lingua. Sono così numerose le sue varietà che, per quanto un corpus possa essere vasto, non comprenderà mai tutte le possibilità offerte da un sistema linguistico. Senza dimenticare, inoltre, che a favore di una scientificità sempre maggiore è ormai pratica diffusa confrontare i dati di un corpus con altri tipi di risultati (Gilquin 2010).

Rappresentatività e bilanciamento dipendono infine da due fattori citati da Lenci, ossia l'effettiva disponibilità dei testi e le capacità intuitive del compilatore (Lenci et al. 2007: 41). La cosa fondamentale è essere consapevoli dei limiti del proprio corpus e renderne coscienti i futuri utenti. Come ricorda Nesselhauf (2011), un corpus è – o dovrebbe essere – una collezione di testi “sistematica”, dove l'aggettivo "sistematico" si

⁶⁶ Entrambi verranno approfonditi nel paragrafo 1.8.4.

riferisce, da un lato, ai criteri extralinguistici stabiliti per la sua progettazione, dall'altro, al fatto che le informazioni relative a questi criteri vengano rese note.

1.3.2. Tipi di corpora tradizionali

A seconda degli scopi a cui devono servire, i corpora possono essere raggruppati in diversi tipi di cui indichiamo di seguito i principali:

- Corpora generali: raccolgono testi di lingua scritta (e in molti casi anche trascrizioni di lingua orale) di diverso genere, prodotti in tutte le varietà di una lingua storico-naturale. Si tratta di corpora di enorme portata costituiti da milioni di parole. Alcuni di essi vengono costantemente aggiornati⁶⁷, come il corpus di lingua inglese *The Bank of English* o il NKRJa. Altri sono invece sincronici e statici: contengono testi prodotti in un dato periodo e non vengono ampliati, come il sopracitato *British National Corpus* (cfr. par. 1.8.).
- Corpora specialistici: comprendono testi di un unico genere o linguaggio settoriale. Un esempio è dato dall'*Air Traffic Control Corpus* (ATCC), una raccolta di trascrizioni riportanti le comunicazioni tra piloti e torre di controllo⁶⁸.
- Corpora paralleli⁶⁹ e comparabili: contengono testi in due o più lingue naturali. In un corpus parallelo i testi in una data lingua A sono tradotti in una lingua B e sottoposti ad un processo di allineamento, per cui testo in lingua A e traduzione in lingua B compaiono affiancati a livello di frase. Invece, in un corpus comparabile i testi in lingua A e i testi in lingua B svolgono la stessa funzione, ossia appartengono alla medesima tipologia testuale, ma, a differenza dei corpora paralleli, non sono uno la traduzione dell'altro. Si nota, tuttavia, che la terminologia in quest'area non è univoca. Alcuni studiosi definiscono "comparabili" anche corpora contenenti testi originali e traduzioni nella stessa lingua, in modo da confrontare le caratteristiche proprie della lingua originale e di quella tradotta (Borin 2002: 3).

⁶⁷ In base a tale caratteristica si distingue tra corpora di riferimento e non (cfr. par. 1.8.2).

⁶⁸ Cfr. Šmídl, Ircing (2014).

⁶⁹ Si è scelto di adottare questa terminologia ormai largamente accettata, sebbene in alcuni casi ci si possa riferire ai corpora paralleli con la denominazione "*translation corpora*" (Granger 2003: 19, Zanettin 2012: 152-153, Aijmer, Altenberg 2000: 15, Johansson 1998: 5). Sui tipi di corpora paralleli cfr. par. 1.4.

- Corpora storici o diacronici: contengono testi di una lingua appartenenti a diversi periodi.
- Learner corpora: comprendono testi prodotti da apprendenti. Possono essere, per citare un esempio, apprendenti di una seconda lingua che producono testi in quella lingua. Questi tipi di corpora sono solitamente molto utili per scopi didattici, per l'analisi degli errori.
- Corpora multimediali: contengono video e materiale multimediale, spesso affiancato da una trascrizione, come nel caso del NKRJa.

1.4. I corpora paralleli

Come accennato nel paragrafo precedente, la terminologia relativa ai corpora paralleli (CP) è talvolta ambigua. Sarà pertanto nostra cura cercare di chiarire le differenze terminologiche e proporre la nostra definizione di corpus parallelo - ossia una raccolta di testi che, seppur redatti in lingue diverse “intendono esprimere lo stesso senso” (Johansson 1998: 5) – dal momento che proprio a questo tipo di corpus è dedicato il nostro lavoro.

Borin (2002: 3-4) riporta la classificazione dei corpora paralleli proposta nel 1999 da Merkel e li divide in diversi tipi, a seconda della relazione che intercorre tra i testi che li compongono. I più diffusi, nonché i più rilevanti per il nostro lavoro, sono quelli che Merkel definisce a) “*translation corpora*”, b) “*target variant corpora*”, c) “*multi-target corpora*” e d) “*mixed-source corpora*”. Di seguito diamo una breve descrizione:

- a) presentano testi in lingua A e le rispettive traduzioni in lingua B.
- b) presentano testi in lingua A e più di una traduzione in lingua B (questo tipo utilizza quasi esclusivamente testi di genere letterario dato che pressoché solo in quest'ambito accade che un'opera venga tradotta più volte da traduttori diversi).
- c) un testo in lingua A è allineato con traduzioni in diverse lingue (B, C, D...).
- d) testi in varie lingue, esprimenti lo stesso senso, sono allineati tra loro, ma non è noto quale sia l'originale.

Secondo la terminologia di Merkel, il corpus italiano-russo oggetto di questo lavoro sarebbe principalmente di tipo a), e in parte di tipo b). Per questo nel definirlo abbiamo

scelto la denominazione generale di “corpus parallelo” che, come si diceva, è comunque largamente accettata e comprensibile. Non verranno invece considerati corpora multilingui, dal momento che le lingue di nostro interesse rimangono due (russo e italiano), come pure corpora di tipo d), poiché si ritiene essenziale, nell’analisi linguistica, conoscere in primo luogo la lingua del testo fonte, ma anche autore e contesto storico di produzione.

Un ulteriore criterio di classificazione dei CP si basa sul numero di lingue in essi contenute e sulla direzione delle traduzioni. Gandin (2009: 134), ad esempio, propone la seguente suddivisione:

- corpora paralleli uni-direzionali, composti da testi originali in una lingua A e dalle traduzioni in lingua B (modello $A \rightarrow B$);
- corpora paralleli bi-direzionali, con testi originali in lingua A e tradotti in lingua B, e testi originali in lingua B, tradotti in lingua A; (modello $A \rightarrow B / B \rightarrow A$);
- corpora a stella, che presentano testi originali in lingua A, tradotti in più di una lingua (ossia i “*multi-target corpora*” di Merkel);
- corpora a diamante, molto complessi, costituiti da testi originali in più di una lingua e da una serie di traduzioni combinate in diverse lingue.

Nel CP italiano-russo del NKRJa, oggetto di questo lavoro, i testi originali in lingua A vengono affiancati dalla rispettiva traduzione in lingua B e viceversa: il rapporto tra le due lingue è quindi bi-direzionale.

Un aspetto fondamentale di qualsivoglia CP, oltre alle caratteristiche già citate nel paragrafo 1.3. (formato elettronico, annotazione, bilanciamento), è l’allineamento fra testi fonte e testi d’arrivo. L’allineamento avviene generalmente a livello di frase ed è realizzato con appositi programmi che operano in modo automatico, ma impongono un lavoro di revisione manuale (cfr. par 1.5.2.).

1.4.1. Possibili applicazioni dei CP

Come tutti i tipi di corpora, anche i CP consentono diverse applicazioni. È difficile classificarle in maniera sistematica e completa, se non altro perché, come afferma giustamente Johansson:

“Once compiled, [a corpus] is transportable and can be used for purposes which may not have been foreseen at the stage of compilation”⁷⁰
(Johansson, 1998: 3).

Proveremo tuttavia a presentarne un elenco il più completo possibile.

Nonostante lo scetticismo iniziale tendente a considerare le traduzioni come testi non rappresentativi di una lingua (Olohan 2004: 13) o “versioni distorte e di seconda mano rispetto ai testi *reali*” (corsivo nostro), scetticismo per cui i testi tradotti venivano deliberatamente esclusi dai corpora (Baker 1993: 233)⁷¹, oggi è largamente riconosciuto che i CP, composti al 50% da traduzioni, sono uno strumento estremamente utile sia in numerosi ambiti della ricerca linguistica, sia nella traduzione umana o automatica. Come ricorda Laviosa (2002: 18), fu Mona Baker (1993), da sempre impegnata negli studi di traduttologia, la prima a dichiarare la validità e l’importanza dei corpora paralleli per studiare la traduzione. Anzi, proprio la sua diversità rende la traduzione degna di essere studiata e analizzata (Baker 1993: 234). Sempre nell’ambito della traduzione, Granger riporta una considerazione di Zanettin, secondo il quale l’uso di CP sarebbe un valido aiuto anche per la formazione di traduttori principianti (Granger 2003: 25).

Come già accennato, oggi le tecnologie permettono di allineare ad un testo fonte più di una traduzione. Ciò consente dunque di condurre ricerche che mettano a confronto le diverse versioni in lingua B di un testo in lingua A. Salkie, in un contributo del 2002, mostra come i CP permettano, tra le altre cose, di cogliere un fenomeno da lui definito “*inventiveness*”, ossia l’uso creativo della traduzione. Il linguista può così tentare una classificazione dei vari tipi di “*inventiveness*” e verificarne la regolarità. Ciò, secondo l’autore, segna anche un possibile punto di incontro tra studi di traduttologia e linguistica (Salkie 2002: 96-98). Confrontando le diverse versioni di un unico testo in lingua A, è possibile dunque stabilire quale tra queste è da considerarsi maggiormente inventiva.

McEnery e Xiao citano invece Mauranen, il quale afferma che i CP sono “indispensabili” per la linguistica contrastiva (McEnery, Xiao 2008: 25). Già alla fine

⁷⁰ “Una volta compilato, [un corpus] è trasportabile e potrebbe essere usato anche per scopi che possono non essere stati previsti in fase di progettazione”

⁷¹ Per indicare le differenze tra testi originali e testi tradotti è stato coniato il termine “*translationese*”, talvolta usato per sottolineare il fatto che i testi d’arrivo spesso sono influenzati dal testo fonte. Il termine ha un valore dispregiativo e intende distinguere la lingua della traduzione dal resto della produzione linguistica, mettendone in evidenza l’inadeguatezza (Zanettin 2012: 12).

degli anni Novanta, Johansson sottolineava come le ricerche di tipo cross-linguistico fossero sempre più numerose nell'ambito della LC (Johansson 1998: 3). Questo perché i corpora in generale, e nello specifico quelli paralleli, garantiscono alla linguistica contrastiva una base empirica piuttosto solida rispetto alla mera intuizione linguistica (Granger 2003: 18), fornendo una piattaforma di dati entro la quale le ipotesi e le intuizioni relative al confronto tra due lingue possono essere effettivamente verificate (Johansson 2003: 35-36). Si noterà, inoltre, che il problema della validità o meno dei corpora nell'analisi contrastiva non è recente; già alla fine degli anni Settanta, quando la LC cominciava appena ad emergere e non esistevano corpora bilingui validi, Filipovič presentò un articolo al XIII Congresso della Società di Linguistica Italiana, dedicato alla linguistica contrastiva, in cui si interrogava, tra le altre cose, sui metodi più appropriati per condurre analisi di questo tipo, promuovendo già allora l'uso dei corpora in questo settore (Filipovič 1982: 10-13).

Tra le applicazioni dei CP Johansson cita anche la possibilità di creare dizionari bilingui basati su corpora, grammatiche contrastive e materiali didattici (Johansson 2003: 39), utili per l'insegnamento di una seconda lingua. A proposito dello studio di una lingua, in particolare di una L2, si è insistito in numerose occasioni sulla validità del "*data-driven learning*", un metodo che, grazie all'uso diretto dei corpora da parte dello studente, potenzia la sua capacità di prendere coscienza delle strutture e degli usi della L2. In particolare, i corpora bilingui permettono di cogliere differenze ed equivalenze tra L1 e L2, momento importante nel processo di apprendimento di una lingua (Granger 2003: 24). Ancora, grazie a CP bilingui, è possibile guidare lo studente, anche principiante, verso l'apprendimento autonomo di vocaboli sconosciuti, come dimostra St. John in uno studio sull'uso di un corpus parallelo per l'apprendimento del tedesco (St. John 2002). Lo stesso è confermato da Bernardini (2004: 2) nel suo contributo al volume di Sinclair "*How to use corpora in language teaching*"⁷².

Interessante, poi, l'idea di commissionare allo studente stesso la compilazione di un corpus, così da potenziare la sua sensibilità e il suo senso critico in riferimento ai corpora già esistenti e disponibili online (Aston 2002). Per una più ampia visione sulle possibilità offerte dalla LC nella didattica si rimanda anche agli altri contributi raccolti nel volume "*Teaching by doing corpus analysis*" (Kettemann, Marko 2002) e in generale

⁷²L'articolo di Bernardini è stato in seguito tradotto in russo e reso disponibile sul sito del NKRJa. È questa la fonte a cui facciamo riferimento.

ai contributi emersi di volta in volta nelle conferenze dedicate a "*Teaching and Language Corpora*" (TALC)⁷³.

Un tipo di applicazione meno diffuso è messo in risalto da Borin (2002: 15-16), e consiste nell'uso dei CP per studiare e conservare lingue minoritarie e in via d'estinzione.

Infine, la comparsa di CP bilingui o multilingui ha permesso alla linguistica computazionale di compiere passi avanti nella programmazione di software di traduzione e, più in generale, ha favorito lo sviluppo del Trattamento Automatico della Lingua (TAL) (Calzolari, Lenci 2004). Borin (2002: 7) denomina quest'ultimo ambito di applicazione "linguistica dei corpora computazionale" e sottolinea come le due tradizioni, quella "computazionale" e quella "generale", si siano quasi totalmente distaccate.

Anche nella russistica è stato svolto fino ad oggi un certo numero di ricerche, basate sui corpora paralleli del NKRJa⁷⁴, seppur quasi esclusivamente per l'inglese e il tedesco, vista la maggiore entità dei corpora paralleli in queste due lingue. Interessante, ad esempio, lo studio di (Gurina, Dobrovol'skij, Kretov 2005) su uno dei maggiori problemi di compilazione dei corpora paralleli, l'allineamento dei testi, dove di frequente il livello di segmentazione dell'originale non è mantenuto dai traduttori. Gli autori si interrogano non solo sulle difficoltà di allineamento derivanti da questa non corrispondenza, ma soprattutto sulle possibili motivazioni che conducono i traduttori a simili scelte. Dello stesso Dobrovol'skij ricordiamo, inoltre, lo studio sulle funzioni espletate rispettivamente dal russo *net* e dal tedesco *nein* (Dobrovol'skij, Levontina 2009) o, ancora, l'analisi della traduzione in tedesco di parole linguospecifiche russe (Dobrovol'skij 2009); infine, di più recente pubblicazione lo studio di alcune particelle russe che segnalano il processo di riattivazione della memoria: attraverso l'analisi di traduzioni di testi russi in lingua inglese e tedesca tratte dal corpus parallelo si analizzano le strategie di resa di tali particelle in queste due lingue (Dobrovolskij, Levontina 2015). Un lavoro di E. Andreeva, e V. Kasevič (2005) riguarda invece il confronto fra tre diverse traduzioni russe dell'opera "*Alice in Wonderland*" di L. Carroll in base ad una serie di parametri, quali numero di parole o di frasi, e la corrispondenza di strutture grammaticali (ed esempio l'uso del passivo) tra originale e traduzioni.

⁷³ La prima conferenza TALC si svolse nel 1994 a Lancaster. Da allora ha scadenza biennale. (<http://www.uni-giessen.de/faculties/f05/engl/ling/talc>) (gennaio 2017).

⁷⁴ Le ricerche di Dobrovol'skij, ad esempio, si sono spesso basate sugli esempi tratti da due corpora paralleli: quello del NKRJa e quello dell'*Austrian Academy Corpus* (AAC).

Naturalmente, lavorare con i CP non sempre è possibile: come ricorda Zanettin (2012: 33), il problema principale riguarda proprio la loro limitata diffusione, in quanto il lavoro di progettazione e creazione di un CP è particolarmente complesso (cfr. anche Borin 2002: 5). Johansson (1998: 6) e Granger (2003), ad esempio, citano la difficoltà di reperire le traduzioni di alcuni testi: per molte tipologie testuali semplicemente non ne esiste alcuna e, in generale, il numero di traduzioni in una lingua sarà sempre minore in confronto alle produzioni originali. A ciò si aggiunge il fatto che le traduzioni dall'inglese in una qualsiasi lingua B tendono ad essere sempre in surplus rispetto alle traduzioni da altre lingue di partenza, rendendo dunque impossibile un adeguato bilanciamento dei CP. Negli ultimi anni molti corpora paralleli sono stati creati appositamente per far fronte al deficit di coppie linguistiche diverse dall'inglese: è il caso del corpus parallelo spagnolo-norvegese descritto da Hareide e Hofland (2012). A queste pecche Johansson – sostenuto da Beeching, interessato alla resa di marche pragmatiche meno frequenti nello scritto – aggiunge la mancanza di materiale per il linguaggio parlato, come pure i problemi legati al “*translationese*” (cfr. nota 71) – citati anche da Borin (2002: 4) – e alla soggettività dei traduttori, le cui scelte di resa del testo fonte in lingua B possono essere oggetto di errate interpretazioni. A questo proposito Beeching (2013: 122-123), utilizzando un corpus parallelo francese-inglese per esplorare la variazione semantica di *quand même*, nonostante promuova i CP per la comprensione di marche pragmatiche, i cui svariati sensi sono poco contestualizzati nei dizionari, se non totalmente assenti, non manca di notare che la cattiva resa o omissione di tali lessemi nelle traduzioni può essere legata semplicemente ad una negligenza del traduttore. Da qui lo scetticismo a cui si accennava all'inizio del paragrafo, e la riluttanza ad usare le traduzioni come base empirica per studi di tipo contrastivo. Per ovviare a questo problema, Johansson propone di approfondire le analisi contrastive basate su di un CP, mettendo a confronto non solo originali in lingua A e traduzioni in lingua B, ma anche originali in lingua A e originali in lingua B, nonché traduzioni in lingua A e traduzioni in lingua B, oppure ancora originali in lingua A e traduzioni in lingua A, nonché originali in lingua B e traduzioni in lingua B (Johansson 1998: 6). Tale procedimento “incrociato” si può realizzare facilmente nel caso di CP bidirezionali, che sono da considerarsi anche comparabili poiché consentono di confrontare testi dello stesso genere che non abbiano tra loro un rapporto di equivalenza traduttiva.

Un ulteriore problema dei CP è l'esistenza di più traduzioni per una singola opera: se quest'ultima è stata tradotta molte volte in una lingua B, sorge la difficoltà di

determinare quali versioni scegliere e quali invece scartare, basandosi sull'abilità dei traduttori, sull'anno di pubblicazione e la reperibilità; senza contare il rischio di incorrere in traduttori fittizi o traduzioni obsolete, eseguite sulla base di una traduzione intermediaria in una terza lingua: in tutti i casi si tratta di informazioni difficilmente reperibili, la cui ricerca prolunga i tempi di compilazione dei corpora paralleli.

Infine, ricordiamo la questione dei diritti d'autore detenuti molto spesso dalle case editrici; puntualizziamo che nel caso dei CP lo sforzo per ottenere i permessi si moltiplica coinvolgendo non un solo testo, ma due o più: originale e traduzione/i.

In linea di principio la legge sul *copyright* imporrebbe ai compilatori di un corpus l'ottenimento di permessi per l'utilizzo dei testi, salvo nei casi in cui l'autore dell'opera sia deceduto da più di 70 anni e non abbia lasciato eredi ancora in vita. La richiesta dei permessi avviene solitamente in forma scritta tramite fax, lettera o e-mail alle case editrici interessate. Il primo passo non è ottenere un consenso, ma quanto meno una risposta: si presume che il mancato riscontro da parte degli editori scaturisca da un'incomprensione della richiesta. Pertanto è indispensabile essere il più chiari ed esauritivi possibile, sottolineando sempre gli intenti puramente scientifici e non commerciali del proprio lavoro. Spesso, suggerisce Zanettin, può essere utile approcciare i detentori di *copyright* in persona, oppure dichiarare tramite comunicazione scritta, meglio se certificata, che il consenso sarà dato per scontato se non altrimenti contestato entro un certo periodo di tempo (Zanettin 2012: 52-55). Una possibile soluzione al problema è proposta da Davies (2009) per la compilazione di COCA e viene illustrata in Zanettin (2012: 54): Davies, previa consultazione di un legale, e viste le modalità di accesso ai materiali del corpus (ossia tramite un motore di ricerca che non consente di visualizzare larghe porzioni di testo, tanto meno opere intere), ha deciso di considerare legittima l'inclusione dei testi, sebbene protetti da *copyright*, appellandosi alla pratica del *fair use* (uso legittimo) in vigore negli Stati Uniti. Con questa presa di posizione i compilatori di COCA hanno evitato di rivolgersi a chi di dovere per ottenere il permesso di pubblicazione.

In Italia, una simile disposizione è contenuta nell'articolo 70 della legge n. 633 del 22 aprile 1941, il cui comma 1 recita:

“Il riassunto, la citazione o la riproduzione di brani o di parti di opera e la loro comunicazione al pubblico sono liberi se effettuati per uso di critica o di discussione, nei limiti giustificati da tali fini e purché non costituiscano

concorrenza all'utilizzazione economica dell'opera; se effettuati a fini di insegnamento o di ricerca scientifica l'utilizzo deve inoltre avvenire per finalità illustrative e per fini non commerciali.”

Tuttavia manca ancora chiarezza su cosa possa essere considerato effettivamente legittimo.

Nonostante gli ostacoli alla progettazione, gli studi condotti finora e la vasta gamma di possibili applicazioni illustrate dimostrano che i CP rimangono senza dubbio uno strumento valido e degno di essere, per quanto possibile, perfezionato. Inoltre, come ricordato da Johansson, se si dispone anche di un corpus comparabile è possibile arricchire e valorizzare l'analisi condotta a partire da un CP.

1.5. Il NKRJa e i suoi corpora paralleli

La linguistica dei corpora iniziò a muovere i primi passi in Unione Sovietica già negli anni Ottanta: da una parte Eršov e Andrjuščenko dedicarono parte dei loro studi alla creazione di un *mašinnyj fond (fondo computerizzato)* della lingua russa. Dall'altra, presso l'Istituto di Lingua Russa Vinogradov venne a costituirsi un dipartimento destinato al citato *fond* che si occupava in particolare della conversione di testi letterari russi in forma elettronica. Ciononostante, in quegli anni le possibilità tecniche e organizzative per la creazione di un corpus completo non erano sufficienti. È con l'inizio degli anni 2000 che la russistica cominciò a disporre di una squadra piuttosto solida di linguisti, pronti a dare vita a quello che oggi conosciamo come il *Nacional'nyj Korpus Russkogo Jazyka* (Corpus Nazionale della Lingua Russa).

1.5.1. Nacional'nyj Korpus Russkogo Jazyka: storia e struttura

Durante un seminario nell'ottobre del 2000, S. Šarov presentò una relazione sulle problematiche della linguistica dei corpora per quanto riguarda i formati di conservazione dei testi, in particolare di quelli letterari. Inoltre, nel 2001 V. Plungjan cominciò a formare una squadra di linguisti che si occupasse attivamente della creazione di un corpus

nazionale e, contemporaneamente, Yandex⁷⁵ fornì il suo supporto tecnico per la programmazione.

Vista la difficoltà di definire e selezionare i testi che avrebbero costituito il corpus in questione, iniziò un vasto dibattito in merito. L'idea iniziale era di costituire un corpus formato dalle opere dei migliori scrittori nazionali, ma fu proprio Šarov a proporre di creare un corpus nazionale che non contenesse esclusivamente testi letterari, ma che rappresentasse la lingua russa in tutte le sue varietà e in tutti i suoi usi. Da qui la denominazione attuale di Corpus Nazionale della Lingua Russa⁷⁶.

Fin da subito si decise che il corpus sarebbe stato arricchito da un buon livello di annotazione, per la quale si decise di riferirsi innanzitutto al *Dizionario Grammaticale della Lingua Russa* di Andrej Zaliznjak⁷⁷. Inoltre si cercò di eliminare almeno in parte i casi di omonimia grammaticale disambiguando parte del corpus⁷⁸.

Per quanto riguarda la classificazione dei testi, Šarov ideò un sistema di annotazione metatestuale (*metarazmetka*) dettagliato, che consente ai fruitori del NKRJa di cercare testi in base al genere, a un particolare periodo, o riguardanti un tema specifico. Il sistema permette di creare in questo modo dei sotto-corpora *ad hoc* per l'interrogazione ragionata del corpus.

Il 29 aprile 2004 il NKRJa venne ufficialmente aperto alla consultazione pubblica e da quella data è in continua evoluzione (Sičinava 2005).

Il NKRJa comprende al momento (gennaio 2017) circa 600 milioni di parole, intese come occorrenze effettive e, come già detto, ha due caratteristiche fondamentali: innanzitutto vuole essere rappresentativo della lingua russa in tutti i suoi generi e pertanto presenta diverse sezioni, a seconda del tipo di testo; in secondo luogo, è annotato a vari livelli.

Nello specifico l'annotazione del NKRJa è di quattro tipi, sebbene non tutti abbraccino l'intero corpus:

⁷⁵ Una delle maggiori società IT russe che da allora garantisce al Corpus un supporto informatico.

⁷⁶ Šarov inizialmente, nel 2002, lavorò alla compilazione del *Большой Корпус Русского Языка* (БОКР), in inglese *Russian Reference Corpus*, che avrebbe dovuto fungere da corpus di riferimento per la lingua russa, avente come modello il BNC. Il corpus avrebbe compreso 100 milioni di parole e testi appartenenti a diversi generi letterari, come descritto in un articolo dello stesso Šarov, pubblicato per la prima volta nel 2003 (Sharoff 2003: 169). La risorsa venne effettivamente creata (anche grazie alla collaborazione dei compilatori del Corpus Nazionale della Lingua Russa) ed è ancora accessibile al sito <http://bokrcorpora.narod.ru/index.html> (gennaio 2017). Il БОКР può essere considerato a tutti gli effetti una versione pilota di quello che poi sarebbe diventato il NKRJa.

⁷⁷ A. A. Zaliznjak (2003). *Grammatičeskij slovar' russkogo jazyka*. Moskva: Russkie slovari (1^a ed. 1977).

⁷⁸ Una parte del corpus – circa 6 milioni di parole – è stata sottoposta a un processo di disambiguazione.

- Metatestuale: applicata a tutte le sezioni e relativa al testo in generale: autore, genere, registro, anno di produzione.
- Morfologica⁷⁹: applicata a tutte le sezioni, comprende le informazioni linguistiche relative ad ogni singolo lessema (appartenenza alla classe, genere, transitività o intransitività per i verbi, aspetto ecc.) e ad ogni forma di parola specifica (caso in cui occorre, numero, tempo, persona ecc.). I simboli usati per la descrizione grammaticale di una parola sono in caratteri latini (V per verbo, S per sostantivi ecc.)⁸⁰.
- Semantica⁸¹: applicata a tutte le sezioni, aggiunge informazioni semantiche ai lemmi del corpus. Ogni parola è abbinata a uno o più tratti semantici, quali, ad esempio, “moto”, “cosa”, “nome proprio” ecc. L’annotazione è realizzata automaticamente da un programma, “*Semmarkup*”, creato da A. E. Poljakov e in questo caso non è stata sottoposta nemmeno parzialmente ad un processo di disambiguazione; pertanto si potrebbero riscontrare degli errori nel caso di parole polisemiche, omonimi od omografi.
- Sintattica⁸²: applicata solo ad una parte del corpus, accessibile dalla sezione “*sintaksičeskij korpus*”⁸³ (corpus sintattico), oltre alle informazioni sopraindicate, aggiunge la struttura sintattica della proposizione in cui è inserito il lemma o il sintagma ricercato.

In tutti i casi descritti, l’annotazione, poiché eseguita automaticamente, può presentare degli errori (ad eccezione dei testi già sottoposti a correzione). Il sistema permette ai fruitori del Corpus di segnalare gli errori in cui si incorre cliccando sulla voce “*soobščit’ ob ošibke*” (“segnalare un errore”) situata accanto alla parola annotata. Gli utenti possono dunque interagire con la piattaforma per il miglioramento delle funzionalità e dell’efficienza di questo strumento.

Nel NKRJa l’annotazione è molto utile non solo per ottenere informazioni aggiuntive sull’oggetto di ricerca, ma anche per: i) creare sotto-corpora *ad hoc*

⁷⁹ Per un approfondimento rimandiamo alla pagina: <http://ruscorpora.ru/corpora-morph.html> (gennaio 2017).

⁸⁰ Una legenda è fornita alla pagina: <http://ruscorpora.ru/corpora-morph.html> (gennaio 2017).

⁸¹ Cfr. <http://ruscorpora.ru/corpora-sem.html> (gennaio 2017).

⁸² Cfr. <http://ruscorpora.ru/instruction-syntax.html> (gennaio 2017).

⁸³ Motore di ricerca del corpus sintattico: <http://ruscorpora.ru/search-syntax.html> (gennaio 2017).

(selezionando autori di interesse, periodo di pubblicazione, genere, argomento e molto altro; ii) affinare la ricerca: grazie alla “*leksiko-grammatičeskij poisk*” (ricerca lessico-grammaticale), è infatti possibile interrogare il corpus non solo ricercando un dato lessema o sintagma, ma anche per categorie grammaticali. Ad esempio, se l’utente deve cercare tutte le forme del sostantivo “*strana*” (in russo “paese”, “nazione”), ad eccezione del genitivo plurale, il motore di ricerca lo consente. In alternativa, è possibile non ricercare affatto una parola o un sintagma specifico, ma tutte le forme che presentano una o più categorie grammaticali, ad esempio “futuro perfettivo”: il sistema in questo caso selezionerà tutte le occorrenze verbali al futuro perfettivo.

Sulla pagina dei risultati il sistema offre inoltre una serie di opzioni aggiuntive: la visualizzazione in formato KWIC (*key words in context*)⁸⁴, la possibilità di visualizzare i risultati senza accento, nel caso si tratti di parti del corpus accentate, la possibilità di ordinare i risultati in base a vari parametri (autore, data di nascita dell’autore, anno di pubblicazione ecc.). Infine, in alcune sezioni del corpus è possibile creare liste di sintagmi (di due, tre, quattro o cinque unità) estrapolati dal contesto.

Ma vediamo ora come si suddivide il Corpus⁸⁵.

I testi inclusi nel NKRJa, appartenenti a diversi generi testuali, ricoprono un periodo che va dall’inizio del XIII all’inizio del XXI secolo. Il corpus è suddiviso come segue:

- 1) Corpus principale (*osnovnoj korpus*): è la sezione più ampia del corpus, che comprende per il 40% testi letterari (opere drammaturgiche e biografie incluse) e per il 60% testi appartenenti ad altri generi (giornalistici, pubblicitari, di critica letteraria, saggistica di vario genere, giuridico-amministrativi e testi non destinati alla pubblicazione, come diari e corrispondenze private). Il corpus principale abbraccia un periodo che va dalla metà del XVIII all’inizio del XXI secolo.
- 2) Corpus sintattico (*sintaksičeskij korpus*): come già accennato, presenta informazioni sulle relazioni sintattiche instaurate dalla parola ricercata con le altre unità linguistiche della frase.

⁸⁴ Un esempio sarà fornito nel prossimo paragrafo.

⁸⁵ <http://ruscorpora.ru/corpora-structure.html> (gennaio 2017).

- 3) Corpus giornalistico (*gazetnyj korpus*): contiene documenti tratti dai mass media contemporanei. Comprende articoli di giornale redatti nel XXI secolo e pubblicati sia da quotidiani cartacei sia da agenzie d'informazione online. Nel 2015 è stato ampliato con l'introduzione della sezione "Stampa regionale ed estera". Quest'ultima contiene articoli che vanno dal 1990 ai giorni nostri, redatti nelle regioni russe e nei paesi della Comunità degli Stati Indipendenti.
- 4) Corpus di testi paralleli (*parallel'nyj korpus*): a sua volta suddiviso a seconda delle lingue allineate con il russo (cfr. par. 1.5.2.).
- 5) Corpus dialettale (*dialektnyj korpus*): comprende testi redatti nei dialetti parlati in varie regioni della Federazione Russa.
- 6) Corpus didattico (*obučajuščij korpus*): esiguo corpus di testi orientato alla didattica della lingua russa nelle scuole (i testi inseriti sono infatti attinenti ai programmi di insegnamento).
- 7) Corpus poetico (*poetičeskij korpus*): comprende testi poetici a partire dal XVIII secolo ed è annotato anche dal punto di vista metrico.
- 8) Corpus della lingua orale (*ustnyj korpus*): comprende le trascrizioni di conversazioni private, discorsi pubblici e film cinematografici.
- 9) Corpus della storia dell'accento russo (*akcentologičeskij korpus*): comprende testi graficamente accentati appartenenti a tre gruppi: tutti i testi del corpus poetico, alcuni testi del corpus orale e un certo numero di testi poetici composti da autori mai pubblicati.
- 10) Corpus multimediale (*mul'timedijnyj korpus*): contiene frammenti di film cinematografici o registrazioni pubbliche (discorsi politici, conferenze) che vanno dagli anni Trenta ad oggi. Ogni video è accompagnato dalla trascrizione.
- 11) Corpus storico (*istoričeskij korpus*): comprende testi redatti a partire dal XIII secolo in slavo ecclesiastico e antico russo (*drevnerusskij* e *staroruskij*). Inoltre, vi sono documenti scritti su corteccia di betulla tra l'XI e il XV secolo.

Chiarite alcune nozioni di base sulla struttura del NKRJa e sulle possibilità offerte dal suo motore di ricerca, ci si concentrerà ora sui corpora paralleli in esso contenuti.

1.5.2. I corpora paralleli del NKRJa

La sezione dedicata ai CP fu creata nel 2005 con il nome di KoParT e con l'intento iniziale di rendere accessibili testi paralleli in russo e in inglese. Al momento della sua compilazione si trattava esclusivamente di opere letterarie, sebbene fin da subito vi fosse l'idea di ampliarla per quanto riguarda la quantità di testi, le tipologie testuali e le lingue. Il CP russo-inglese era dunque un corpus pilota, puntualizzano Dobrovolskij, Kretov e Šarov (2005).

I criteri che a suo tempo guidarono la selezione dei testi erano essenzialmente due: la qualità dal punto di vista letterario (furono scelti innanzitutto i grandi classici della letteratura russa e anglo-americana) e la loro effettiva disponibilità in formato elettronico.

Come per la maggior parte dei CP, il programma di allineamento utilizzato, originariamente denominato ParTeks (oggi "Euclid"), allineava i testi frase per frase, così che ogni frase del testo originale fosse seguita dalla sua traduzione. Precisiamo che per "frase" si intende in questo caso un segmento di testo compreso tra due punti fermi e, purtroppo, a questo proposito, vi sono alcuni problemi:

- La struttura del testo e la divisione in paragrafi, volute dall'autore nell'originale, molto spesso si perdono nella traduzione allineata.

- lo stile grafico delle diverse lingue, o delle diverse case editrici – ad esempio la punteggiatura impiegata per circoscrivere il discorso diretto – fa sì che talvolta il programma non riconosca i segni che indicano il confine di frase. Due frasi separate nel testo fonte da un punto fermo in traduzione potrebbero essere legate da una virgola o da una congiunzione. Viceversa, una frase troppo lunga del testo fonte potrebbe essere spezzata mediante punto fermo dal traduttore. Di conseguenza, i due testi non risultano perfettamente allineati. Ciò dipende dalle lingue considerate, che possono essere più o meno paratattiche, ma anche dallo stile di chi scrive. In entrambi i casi, è dunque necessario un intervento di correzione manuale che, tuttavia, è facilitato dal programma, il quale, pur non essendo in grado di correggere autonomamente questi errori, permette di riconoscerli segnalando le asimmetrie fra il testo originale e la sua traduzione⁸⁶. Dopo l'allineamento, i testi vengono annotati grazie al programma messo a disposizione da Yandex e impiegato per l'annotazione dell'intero NKRJa (Dobrovolskij, Kretov, Šarov

⁸⁶ Il programma inizia col prendere in considerazione i paragrafi. Una volta che questi sono allineati in maniera simmetrica, si passa all'allineamento delle singole frasi e si confronta il numero di frasi che compongono un paragrafo dell'originale con il numero di frasi presenti nella traduzione corrispondente.

2005). La sezione dei CP prevede l'annotazione metatestuale, morfologica e semantica. La ricerca può essere effettuata in una delle due lingue e può interessare una singola parola, più parole o un sintagma.

Oggi (gennaio 2017), tredici anni dopo il primo corpus pilota, i CP del NKRJa sono quindici e comprendono le seguenti coppie di lingue: russo-armeno, russo-bielorusso, russo-bulgaro, russo-buriato, russo-cinese, russo-francese, russo-estone, russo-inglese, russo-italiano, russo-lettone, russo-polacco, russo-spagnolo, russo-svedese, russo-tedesco (a cui si aggiunge un sotto-corpus composto esclusivamente da grandi classici russi tradotti in tedesco, chiamato “*ruskaja klassika v nemeckich perevodach*”), russo-ucraino. In tutti i casi si tratta di corpora paralleli $A \rightarrow B / B \rightarrow A$, vale a dire bi-direzionali, secondo la classificazione di Gandin citata nel paragrafo 1.4. Nel complesso la sezione dei CP include 1.815 documenti, 5.491.636 frasi e 72.533.975 parole.

In aggiunta, vi è un piccolo corpus multilingue composto da 12 testi in inglese tradotti in undici idiomi (russo, olandese, ucraino, polacco, ceco, slovacco, sloveno, croato, serbo, macedone e bulgaro) e contenente 219.642 frasi e 5.025.714 parole.

Nel corso di questi anni vi sono state delle modifiche anche nel programma di allineamento che, in parte, pongono rimedio ai problemi precedentemente esposti, nonostante rimanga sempre necessario un intervento di verifica ed eventualmente correzione manuale, anche se minimo.

Riportiamo nella tabella 1 i dati numerici relativi ai corpora paralleli del NKRJa. Per quanto riguarda il CP italiano-russo, riportiamo le dimensioni del corpus prima dell'ampliamento da noi realizzato:

Tab. 1. Dimensioni dei corpora paralleli del NKRJa

CORPUS	NUMERO DI FRASI	NUMERO DI PAROLE (token)
russo-armeno	175.158	2.136.462
russo-bielorusso	829.829	8.181.435
russo-bulgaro	290.744	3.580.446
russo-buriato	6.674	75.340
russo-cinese	3.309	55.424

russo-estone	46.933	408.326
russo-francese	206.983	3.000.071
russo-inglese	1.606.212	24.677.638
russo-italiano	<u>33.334</u>	<u>706.246</u>
russo-lettone	73.856	777.688
russo-polacco	501.806	6.319.783
russo-spagnolo	25.658	320.917
russo-svedese	43.614	409.485
russo-tedesco	515.403	9.130.803
russo-ucraino	919.437	9.383.924

1.5.3. Il primo corpus parallelo italiano-russo / russo-italiano del NKRJa

Prima di essere ampliato, il corpus di testi paralleli in russo e italiano (CP ita-ru)⁸⁷ era uno dei più piccoli per numero di frasi e di parole, come mostra la tabella 1. All'interno degli 11⁸⁸ corpora di testi paralleli disponibili prima di dicembre 2015, si trovava al terzultimo posto per numero di parole, seguito solamente dal corpus russo-lettone e russo-spagnolo. Le ridotte dimensioni del corpus erano da attribuire principalmente all'assenza di un team di lavoro che potesse dedicarsi con costanza al suo ampliamento, nonostante fosse chiaro fin da subito che un corpus parallelo italiano-russo avrebbe apportato notevoli benefici agli studi legati alla traduzione e alla linguistica contrastiva, trovando numerose applicazioni anche nel campo dell'insegnamento, dell'apprendimento linguistico e della ricerca in generale. Ecco perché ci si è impegnati a lavorare per un suo miglioramento quantitativo e qualitativo.

Fino a dicembre 2015, i testi compresi nel CP ita-ru appartenevano esclusivamente alla prosa letteraria. Si trattava di due romanzi russi tradotti in italiano e tre romanzi

⁸⁷ Precisiamo che in generale, parlando del CP da noi ampliato, si adotterà l'abbreviazione "CP ita-ru", nonostante si tratti di un CP bi-direzionale. Si troverà l'abbreviazione "CP ru-ita" nei capitoli 3 e 5, nel descrivere l'analisi condotta con il sotto-corpus russo-italiano.

⁸⁸ I corpora paralleli russo-buriato, russo-cinese, russo-estone e russo-svedese sono stati aggiunti dopo dicembre 2015.

italiani con rispettiva traduzione in lingua russa. Di seguito riportiamo autori e titoli in ordine alfabetico:

- Bulgakov M. A. *Master i Margarita* (prima e seconda parte) (1929-1940)
- Collodi C. *Pinocchio* (1883)
- Dostoevskij F. M. *Belye Noči* (1848)
- Eco U. *Il nome della rosa* (1980)
- Pirandello L. *La signora Frola e il signor Ponza, suo genero* (1917)

Si tratta senza dubbio di autori illustri della prosa italiana e russa, ma questo corpus, oltre ad essere molto esiguo e, di conseguenza, scarsamente rappresentativo, non era sistematico, ossia al momento della compilazione non erano stati formulati (e probabilmente non esistevano affatto) i criteri che hanno guidato la scelta dei testi i quali, per altro, appartengono a periodi abbastanza diversi. Infine, non è mai stato chiaro se si trattasse di testi integrali o di frammenti, dal momento che manca una descrizione dettagliata di questo micro corpus.

Dal punto di vista del suo funzionamento – per la ricerca di parole, sintagmi o frasi - il CP ita-ru è strutturato analogamente al resto del NKRJa. È possibile cercare una forma di parola precisa, oppure un lessema realizzato in tutte le sue forme. Una volta inoltrata la richiesta, comparirà una schermata con frasi contenenti l'elemento ricercato, seguite dalla rispettiva traduzione. Come già detto, un'opzione permette di ordinare i risultati sulla base di diverse variabili, quali la data di pubblicazione dell'opera, o l'età dell'autore. Infine è possibile ottenere i risultati in formato KWIC, anche se in questo caso i risultati compariranno in una sola lingua e sarà necessario cliccare su un'apposita freccia per ottenere anche la traduzione del paragrafo (vedi Fig.1 e 2). Segnaliamo che, per qualsiasi chiarimento relativo all'utilizzo del corpus, sul sito del NKRJa è possibile accedere a un manuale d'istruzioni abbastanza completo⁸⁹.

Riportiamo ora, come esempio, i risultati della ricerca del lessema "madre":

⁸⁹ Alla voce “*instrukcija*”, in alto a destra sulla pagina del motore di ricerca.

Fig. 1. Esempio di visualizzazione in formato standard

Ora, via, si capisce che una figliuola, maritandosi, lasci la casa della **madre** per andare a convivere col marito, anche in un'altra città [...] ⁹⁰. [Luigi Pirandello. La signora Frola e il signor Ponza, suo genero (1917)]

Что ж, вполне понятно, когда, выйдя замуж, девушка покидает родительский дом и следует за своим супругом, [...] [Луиджи Пиранделло. Госпожа Фрола и ее зять господин Понца (Марина Гасанова-Мийат, 2006)]

Fig. 2. Esempio di visualizzazione in formato KWIC

alle folle, prima che santa **madre** chiesa si muovesse. E tu ←...→
del Signore e l'aiuto della **Madre** di Dio. ←...→
interno alle leggi di santa **madre** chiesa. Essi vollero sempre migliorare ←...→
quanto quella della nostra santa **madre** chiesa e non pensavano affatto ←...→
e accettar l'opinione di santa **madre** chiesa. Ma Michele, durissimo: "Io ←...→
animale "agnoscit", riconosce la propria **madre** , e ne riconosce la voce ←...→
con le questioni di santa **madre** chiesa. ←...→
abbazia dalla stessa regina Elena, **madre** dell'imperatore Costantino, dopo che ←...→
racchiusa ancora nel ventre della **madre** . E in quella nebbia dell' ←...→
tutta prima di distinguere mia **madre** , poi mi resi conto dell' ←...→
contro il padre e la **madre** che avevano messo al mondo ←...→
nuda come l'aveva fatta sua **madre** , batté le mani e gridò ←...→

Dopo aver illustrato le caratteristiche del CP italiano-russo del NKRJa passiamo ora in rassegna altri corpora, paralleli e non, i cui pregi e difetti sono serviti per l'elaborazione del nostro progetto di ampliamento del CP appena descritto.

1.6. Collezioni di testi paralleli italiano-russo al di fuori del NKRJa

Sia per la lingua russa, sia per l'italiano, considerati separatamente, esistono diversi CP disponibili gratuitamente in rete. Un'enorme quantità di corpora paralleli, o meglio collezioni di testi – attenendosi alla distinzione proposta nel par. 3 – è raccolta, ad

⁹⁰ Taglio nostro.

esempio, sui già citati *InterCorp*⁹¹ e *Opus*⁹², la maggiore risorsa online ad essi dedicata. Quest'ultimo raccoglie dati provenienti da diversi settori e ricopre più di 90 lingue. Si tratta perlopiù di testi legislativi o amministrativi reperiti dai database delle Nazioni Unite o di altre istituzioni, quali l'Unione Europea, ma vi sono anche sottotitoli di film, articoli di giornale e altri testi tratti da varie risorse online a cui *Opus* rimanda direttamente (Tiedermann 2012). *Opus* permette di avviare una ricerca nelle coppie di lingue desiderate e fornisce all'utente la lista di corpora comprendenti le lingue richieste. È necessaria un'attenta analisi delle fonti, dal momento che a volte si tratta di file tradotti sia in russo che in italiano, ma con originali in un'altra lingua e il portale in questo caso non fa distinzione. Le collezioni di testi paralleli per il russo e l'italiano che abbiano come originale una delle due lingue sono le seguenti: *OpenSubtitles2013*, *OpenSubtitles2012*, *OpenSubtitles2011*, *Books*, *EUbookshop*, *Tatoeba*.

Opus consente di scaricare i file allineati, oppure, in alternativa, presenta un motore di ricerca che attinge ai dati di ciascuna collezione singolarmente; la ricerca non può essere effettuata su tutte le collezioni simultaneamente. I risultati ottenuti vengono incolonnati, con frase in L1 sulla sinistra e la rispettiva traduzione in L2 sulla destra; in alternativa possono essere visualizzati in formato KWIC.

Considerata la mancanza di chiarezza relativa a molte di queste risorse (autori dei testi e delle traduzioni, anno di pubblicazione, lingua fonte e lingua d'arrivo ecc.), non ci è consentita una descrizione precisa. Ci limiteremo ad inserire dunque in una tabella riassuntiva (tab. 2) solo i dati effettivamente disponibili.

⁹¹ Il progetto *InterCorp* è nato con l'intento di creare un ampio corpus parallelo comprendente tutte le lingue studiate alla Facoltà di Arte dell'Università Carolina di Praga. Una parte del corpus – definita *core corpus* – è stata compilata in modo tradizionale, allineando testi letterari manualmente con la rispettiva traduzione in una data lingua, sebbene in alcuni casi le coppie di lingue allineate siano entrambe traduzioni. In aggiunta, *InterCorp* si compone di testi processati automaticamente, quindi collezioni (cfr. definizione di EAGLES) di estratti da diverse risorse, in parte accessibili attraverso altri portali (ad esempio *Opus*). Nella home page di *InerCorp* si avvertono gli utenti della possibilità di incontrare errori di allineamento proprio per l'assenza di qualsiasi intervento manuale. Queste le fonti: *Project Syndacate e PressEurop* (collezione di commenti politici); *Acquis Communautaire corpus* (comprendente testi legali dell'Unione Europea); *Europarl* (accessibile anche da *Opus*); *Open Subtitles* (accessibile anche da *Opus*). Il corpus è di vastissima portata. A maggio 2015 il *core corpus* contava 195 milioni di parole intese come token complessivamente. La parte ceco-italiano ne contava 7,2 milioni. Di entità ancora maggiore è la sezione tratta dalle fonti sopra citate, dove si raggiungono 1,229 milioni di parole (sempre considerando tutte le coppie di lingue). Il corpus è dotato di una parziale annotazione morfologica. Per un approfondimento cfr.: <http://ucnk.ff.cuni.cz/intercorp/?req=page:info&lang=en> (gennaio 2017).

⁹² <http://opus.lingfil.uu.se/> (gennaio 2017).

Tab 2. Collezioni di testi paralleli ru-ita disponibili online

CORPUS	DOCUMENTI	FRASI	TOKEN	DESCRIZIONE
OpenSubtitles 2011-2012-2013	10.853	7,6 milioni	91.6 milioni	Portale che presenta la traduzione di sottotitoli in diverse lingue. Le traduzioni sono realizzate e caricate sul sito dagli utenti, i quali svolgono il lavoro volontariamente e gratuitamente.
Books	2	17.900	600.000	Collezione di libri allineati in varie lingue da <u>Andras Farkas</u> ⁹³ . Per russo e italiano tuttavia gli unici testi disponibili sono i due volumi di <i>Anna Karenina</i> di L. Tolstoj.
EUbookshop	37	6.600	200.000	È una libreria online composta da pubblicazioni, dal 1952 ai giorni nostri, in più di 50 lingue. La maggior parte dei documenti è redatta da enti dell'Unione Europea. La libreria è gestita dall'ufficio pubblicazioni dell'UE in Lussemburgo.
Tatoeba	/	2.700	33.900	Corpus parallelo aperto e "cooperativo" composto esclusivamente da frasi tradotte in 129 lingue. Anche in questo caso si tratta di traduzioni eseguite dagli utenti. Chiunque può offrire il proprio contributo e chiunque può impiegare il materiale tradotto a scopi di ricerca, per manuali, applicazioni e altro ancora.
totale	10892	7,6 milioni	92,5 milioni	

Come avviene tradizionalmente nell'ambito della LC, il confronto con corpora già esistenti fornisce modelli da impiegare nella futura progettazione di un corpus. Come riporta Šarov, la nascita e lo sviluppo degli stessi corpora di lingua russa, tra cui БОКР (cfr. nota 76), hanno seguito le linee guida dettate dai loro predecessori, in particolare dai corpora di lingua inglese (Sharoff 2006: 168). Anche per l'ampliamento del CP ita-ru, dunque, ci atterremo a tale pratica. Tuttavia, le risorse elencate nella tabella 2 mancano della sistematicità posta come caratteristica necessaria da Nesselhauf (2011) nel paragrafo 1.3. e, in generale, non rispondono ai criteri che individueremo per caratterizzare un corpus bilanciato e utile a scopi di ricerca. Per questo motivo, prenderemo in esame sia altri CP creati fino ad ora al di fuori del NKRJa e non necessariamente in lingua russa e

⁹³ Reperibili al sito http://www.farkastranslations.com/bilingual_books.php (gennaio 2017).

italiana (perché assenti), sia i principali corpora generali che si può dire abbiano segnato la storia della LC.

1.7. Corpora a confronto

Per stabilire i criteri utili alla progettazione di un corpus parallelo ci riferiremo, inizialmente, ai contributi di quanti, in passato, hanno affrontato questo problema. In particolare si è deciso di prendere in esame i seguenti corpora:

- tre corpora generali di lingua inglese, che dagli anni Sessanta ad oggi hanno fornito validi modelli nell'ambito della LC, ossia il *Brown Corpus* (BrC), il *British National Corpus* (BNC) e il *Corpus of Contemporary American English* (COCA);
- uno dei principali corpora di riferimento della lingua italiana: CORIS/CODIS;
- un corpus parallelo di testi in inglese e norvegese, che riteniamo particolarmente significativo: *l'English-Norwegian Parallel Corpus* (ENPC);
- CEXI (*the English Italian Translational Corpus*), che, sebbene mai realizzato per mancanza di fondi, intendeva essere il primo corpus sistematico e bilanciato di testi paralleli in inglese e italiano (Zanettin 2012: 49)⁹⁴.

Elenchiamo i corpora di riferimento con alcune informazioni tecniche relative al tipo di corpus e alla sua dimensione, misurata in numero di token:

Tab. 3. Elenco e descrizione dei corpora di riferimento

NOME	TIPO DI CORPUS	TOKEN
BrC	Corpus generale ING americano	1 milione
BNC	Corpus generale ING britannico	100 milioni
COCA	Corpus generale ING americano	450 milioni
CORIS/CODIS	Corpus generale ITA	130 milioni
ENPC	Corpus parallelo ING-NORV	2,6 milioni
CEXI	Corpus parallelo ING-IT	4 milioni

⁹⁴ Sebbene il corpus non sia mai stato realizzato, sembra utile esaminarne la progettazione, che invece è disponibile, proprio perché rappresentava uno dei pochi progetti per la creazione di un corpus parallelo bilanciato per la lingua italiana.

Abbiamo considerato innanzitutto il BrC e il BNC poiché i due corpora hanno contribuito a plasmare in maniera considerevole l'ancor giovane storia della LC, nata proprio in ambito anglosassone. In particolare il BNC deve la sua esistenza a John Sinclair, uno dei fondatori della linguistica dei corpora, che a suo tempo si occupò in prima persona della sua creazione e progettazione. Pertanto, anche se datati, questi corpora vantano una sistematicità talmente elevata da essere ancora oggi un ottimo esempio per la compilazione di qualsivoglia corpus elettronico. Il COCA è invece, attualmente, il corpus generale più vasto per l'inglese americano e può essere considerato come uno dei corpora generali più bilanciati esistenti. La differenza e novità rispetto al BrC e al BNC sta nel suo continuo monitoraggio, che prevede ogni anno un'aggiunta di 20 milioni di parole equamente distribuite tra le varie tipologie testuali in esso rappresentate.

Per quanto riguarda i restanti corpora, si è voluto prendere in esame, innanzitutto, il corpus generale CORIS/CODIS, uno dei più importanti corpora di riferimento per la lingua italiana creato presso l'Università di Bologna. Infine abbiamo considerato due corpora paralleli: CEXI, per quanto riguarda la lingua italiana, e l'ENPC, un corpus parallelo particolarmente strutturato che si presta molto bene come modello per la costituzione di un CP rappresentativo. Quest'ultimo, in particolare, merita maggiore attenzione per due motivi: il primo riguarda il suo prestigio e la popolarità di cui gode nell'ambito della LC, tant'è che, a sua volta, è stato d'esempio per la progettazione sia di CEXI (Zanettin 2002), sia di molti altri CP, quali ad esempio COMPARA, un corpus di testi paralleli in inglese e portoghese (Frankenberg-Garcia, Santos 2003); il secondo motivo riguarda la sua sistematicità e la relativa completezza rispetto ad altri corpora paralleli oggi esistenti.

1.8. Criteri per la progettazione di un corpus

La compilazione di un corpus richiede innanzitutto un attento lavoro di preparazione e progettazione che definisca chiaramente, e soprattutto preventivamente, le sue caratteristiche. A questo proposito Sinclair afferma:

“A corpus is a remarkable thing, not so much because it is a collection of language text, but because of the properties that it acquires if it is well-designed and carefully-constructed”⁹⁵ (Sinclair 2004a: 1).

Le scelte legate ai vari parametri, quali ad esempio dimensioni e tipologie testuali, sono correlate a una serie di problemi teorici che il ricercatore è tenuto ad affrontare. Presentiamo quindi nel modo più esaustivo possibile le questioni teoriche che accompagnano il momento della progettazione di un corpus e le soluzioni proposte attraverso la descrizione dei corpora sopracitati.

1.8.1. Finalità

La prima questione, messa in evidenza sia dagli autori del BNC, sia da Rossini Favretti nella sua guida alla compilazione di CORIS, è legata alla finalità del progetto. Dalla finalità del nostro corpus, infatti, dipenderà la maggior parte delle restanti variabili. Stabilire la finalità del corpus significa anche definirne le caratteristiche tipologiche.

Nel caso di CORIS, l'intenzione consisteva nel creare un corpus generale e, più precisamente, un corpus di testi rappresentativi dell'italiano contemporaneo (Rossini Favretti 2000). Analogamente il BNC doveva essere un corpus rappresentativo della lingua inglese contemporanea (Burnard 2000), ossia dagli anni Sessanta all'inizio degli anni Novanta.

Un metodo valido, per stabilire lo scopo di un corpus in fase di progettazione, consiste nel definire innanzitutto i suoi possibili usi. Gli autori del BNC, ad esempio, stabilirono i seguenti: ricerca linguistica, insegnamento, intelligenza artificiale, acquisizione di informazioni ed elaborazione del linguaggio naturale (Burnard 2000).

Per quanto concerne gli altri corpora, notiamo che il BrC nacque con uno scopo pressoché analogo a quello di BNC e CORIS: compilare un corpus rappresentativo dell'inglese americano di quel periodo, vale a dire i primi anni Sessanta. Mentre più elaborato è il caso di COCA, che intendeva esplicitamente apportare delle innovazioni rispetto ai corpora precedenti. Come spiega Davies (2009: 3), COCA è stato creato innanzitutto con il fine di ottenere un corpus per l'inglese americano che ovviasse alle

⁹⁵ Un corpus è una cosa eccezionale, non tanto perché è una collezione di testi, ma per le caratteristiche che acquisisce se ben progettato e attentamente compilato.

lacune del BNC, da un lato, e dell'*American National Corpus* (ANC)⁹⁶, dall'altro. Il COCA si pone dunque lo scopo di essere vasto, bilanciato, rappresentativo e soprattutto costantemente monitorato, aspetto che invece non ritroviamo né nel BNC, né nel BrC. In sintesi, lo scopo comune ai corpora citati è senza dubbio la rappresentatività, anche se, come già ribadito, non si tratta di una variabile di semplice definizione.

Per quanto riguarda i corpora paralleli, anche in questo caso la struttura e il design dipendono dal futuro impiego dei dati. Se lo scopo è permettere ricerche di traduttologia letteraria, ne consegue che i testi da inserire saranno esclusivamente letterari. Se l'idea è di concentrarsi su di un singolo autore, verrà in questo caso creato un corpus *ad hoc* con i suoi scritti originali e le rispettive traduzioni, creando un corpus A→B. Se, al contrario, lo scopo è predisporre materiale utile per la traduzione specialistica, si tratterà di identificare le macro-aree a cui attingere per il recupero dei testi: potrebbe trattarsi di articoli di giornale, articoli accademici, testi legali, manuali tecnici e via dicendo (Gandin 2009).

L'ENPC nacque come strumento per la traduzione e gli studi di linguistica contrastiva. L'intento iniziale era creare un corpus bi-direzionale di testi paralleli in inglese e norvegese, prevedendo poi l'inserimento di testi non paralleli, ma comparabili, appartenenti cioè allo stesso genere, ma senza essere l'uno la traduzione dell'altro. Tuttavia, quest'ultimo sotto-corpus non fu mai realizzato (Johansson, Ebeling, Oksefjell 2002: 2). Johansson, Ebeling e Oksefjell mettono in luce come l'ENPC consenta di avanzare negli studi contrastivi grazie all'analisi e al confronto non solo tra testi originali e le loro traduzioni, ma anche tra testi originali in inglese e testi originali in norvegese, nonché tra le traduzioni nelle due lingue. Pertanto il corpus è da considerarsi anche comparabile (per ulteriori riflessioni relative a questo aspetto cfr. par. 4.1.).

CEXI, infine, fu pensato come progetto dell'Università di Interpreti e Traduttori di Forlì (Università di Bologna) con l'intento di favorire gli studi connessi alla traduzione e all'apprendimento della lingua. Zanettin, in fase di progettazione, sottolineava come il corpus avesse non solo finalità descrittive, ma anche applicative: non si trattava solo di apprendere qualcosa sulla lingua, ma anche di imparare a usarla parlando, scrivendo o traducendo (Zanettin 2002: 330).

⁹⁶ Corpus generale e rappresentativo dell'inglese americano, rilasciato per la prima volta nel 2003 e contenente ad oggi circa 15 milioni di parole (Ide, Suderman 2004). Il corpus è accessibile dal sito: <http://www.anc.org/> gennaio 2017).

1.8.2. Corpus statico o di monitoraggio

Dopo aver stabilito le finalità del corpus da creare, come ricorda Kennedy (1998:60), è necessario scegliere se realizzare un corpus statico o un corpus dinamico, meglio conosciuto come “corpus di monitoraggio”. Un corpus statico è anche definito “di riferimento” (*reference corpus*), poiché, rimanendo invariato, presenta sempre le medesime caratteristiche, ma soprattutto gli stessi risultati, e può essere dunque impiegato come punto di riferimento per le ricerche che richiedono dati invariabili.

I corpora di prima generazione nacquero quasi tutti come statici, con lo scopo di considerare la lingua in un intervallo di tempo ben definito. Si pensi al Brown Corpus, con i suoi testi pubblicati esclusivamente nell’anno 1961. Analogamente, il BNC non pianificò mai alcun tipo di monitoraggio successivo al momento della sua creazione.

I corpora più recenti tendono invece ad essere dinamici e aperti. Tale scelta è determinata in parte dalla convinzione che una lingua storico-naturale sia un sistema in continua evoluzione e che proprio questa evoluzione possa essere in molti casi di grande interesse; dall’altro lato lo sviluppo delle tecnologie informatiche permette il continuo monitoraggio di un corpus; non era così nei primi anni Sessanta, quando le possibilità tecnologiche di allora non consentivano di elaborare contemporaneamente un gran numero di dati. L’esempio più palese è il caso di COCA a cui ogni anno vengono aggiunte 20 milioni di parole. Tuttavia, anche il continuo monitoraggio di un corpus non risulta privo di difficoltà, poiché la quantità rischia di prevalere a discapito della qualità. È questo un problema di cui era ben cosciente Sinclair e ce lo ricorda anche Rossini Favretti (2000) quando tratta il problema della dimensione durante la progettazione di CORIS/CODIS. In (Barbera 2013) questo tema prende il nome di “finitzza” di un corpus. Secondo l’autore si tratta di una caratteristica imprescindibile per almeno due motivi: in primo luogo per poter rendere ripetibili le analisi condotte, conferendo così maggiore scientificità ai risultati ottenuti; la seconda motivazione riguarda la possibilità di condurre, grazie ai corpora, analisi statistiche, che, come Barbera puntualizza, esigono dati stabili.

Dal nostro punto di vista, sebbene sia difficile condurre analisi di tipo statistico su dati in continua variazione, senza dubbio maggiore è il loro numero, maggiori sono le possibilità di ottenere una quantità di risultati utili per avanzare ipotesi valide. Il corpus di monitoraggio rimane quindi l’alternativa più desiderabile, purché le informazioni

riguardanti il numero di testi e di parole che compongono il corpus vengano costantemente aggiornate e rese note ai fruitori nel corso dell'ampliamento.

Il continuo monitoraggio di un corpus richiede, naturalmente, risorse e sforzi costanti che spesso scoraggiano la sua realizzazione. L'ENPC, ad esempio, prevedeva l'aggiunta di testi in inglese e norvegese comparabili. In seguito, tuttavia, quest'ultimo progetto è stato abbandonato lasciando spazio alla creazione di un sotto-corpus di testi paralleli multilingui.

CEXI, infine, nasceva con un piano di monitoraggio ben preciso: da un lato si pensava di ampliare gli estratti già esistenti, dall'altro si prevedeva di introdurre nuovi testi. Per permettere la prima opzione, venne da subito espressa la volontà di ottenere dalle case editrici i permessi per testi interi, così da poter aggiungere nuovi campioni degli stessi in una fase di espansione successiva (Zanettin 2002: 340).

1.8.3. Corpus diacronico o sincronico

Oltre a fissare le finalità e scegliere fra corpus statico e di monitoraggio, è necessario stabilire se si vuole creare un corpus diacronico o sincronico. I testi prescelti, infatti, possono risalire a un periodo di tempo circoscritto o possono appartenere a periodi storici differenti. In quest'ultimo caso prevale l'intento di fare luce sull'evoluzione storico naturale della lingua.

I corpora che abbiamo preso in esame sono di tipo sincronico, ossia composti da testi risalenti ad un numero di anni predefinito e limitato. Una volta stabilito tale parametro, è necessario fissare il periodo di pubblicazione dei testi (Zanettin 2002).

Il BrC, ad esempio, contiene testi pubblicati nel 1961, mentre il BNC comprende testi di fiction a partire dal 1960 e testi informativi dal 1975 fino ad arrivare agli anni Novanta (Burnard 2000). In entrambi i casi la scelta è dettata dalla volontà di dare una visione della lingua contemporanea tralasciando gli aspetti evolutivi. CORIS include testi risalenti approssimativamente agli anni Ottanta e Novanta. La scelta rispecchia la volontà di "consentire, tramite generalizzazione, una descrizione dell'italiano corrente nell'uso comune" (Rossini Favretti 2000: 43). Infine, CEXI prevedeva di includere testi pubblicati tra il 1976 e il 2000, salvo rare eccezioni, e preferibilmente ancora in stampa al momento della creazione del corpus (Zanettin 2002: 332). Anche l'ENPC contiene testi pubblicati tra la fine degli anni Settanta e l'inizio degli anni Novanta. Quello di COCA può essere definito, invece, come un caso particolare: esso rappresenta, infatti, un corpus di

monitoraggio, pertanto in continua evoluzione; ciononostante, i testi più datati risalgono agli anni Novanta, vale a dire ad un periodo relativamente recente. Al momento si può dunque affermare che prevale l'aspetto sincronico, ma se il progetto di monitoraggio verrà realmente rispettato a lungo termine, negli anni a venire comincerà a delinearsi anche una componente diacronica.

In questo contesto il NKRJa è innovativo, poiché comprende testi che ricoprono un periodo molto più ampio – più precisamente dagli anni Cinquanta ai nostri giorni – e contiene, inoltre, sia una sezione con testi del XIX secolo e della prima metà del XX, sia una sezione di testi ancora più antichi, risalenti ad un periodo di tempo che va dall'XI al XIV secolo (Rachilina 2009). Questo aspetto, sicuramente di grande interesse, manca in tutti i corpora da noi esaminati.

1.8.4. Rappresentatività e bilanciamento

Il termine "rappresentatività" indica la possibilità di estendere a una lingua nella sua globalità i caratteri riscontrati in un corpus di testi di quella stessa lingua (Leech 1991). Tuttavia, considerando che una lingua storico naturale può presentare al suo interno numerose varietà, la domanda da porsi è: "Cosa vogliamo realmente rappresentare?" Secondo Biber, citato da Hareide e Hofland (2012: 76), il compito principale è definire con chiarezza la popolazione che produce i testi.

Anche in questo caso entrano in gioco fattori concomitanti: da un lato, la finalità del corpus, dall'altro, il concetto di bilanciamento, che, come afferma Sinclair, è di ancor più complessa definizione. Se un corpus ha la pretesa di essere "generale", la rappresentatività sarà data dal bilanciamento tra testi di diverso tipo. Allo stesso modo si cercherà di comprendere testi di vari periodi, a seconda del lasso di tempo che si intende coprire. È però necessaria una precisazione: non è auspicabile inserire testi di generi diversi in quantità identiche. Bisognerà reperire testi appartenenti ad un dato genere in numero proporzionale alla sua effettiva diffusione. Un corpus specialistico, a sua volta, soddisferà altri criteri di bilanciamento. In un corpus creato per analizzare gli errori degli apprendenti di una L2 e formato quindi da testi redatti da studenti, il criterio del bilanciamento imporrà infine che tutti gli studenti coinvolti nel progetto compongano, approssimativamente, lo stesso numero di testi.

Tale descrizione rappresenta un ideale, purtroppo non sempre realizzabile e, talvolta, non rispondente agli obiettivi. Il *Penn Treebank*⁹⁷, ad esempio, non risponde al criterio di bilanciamento, poiché il suo unico fine è rendere disponibile ed elaborabile da una macchina un numero cospicuo di testi per lavori di annotazione e *parsing*⁹⁸. Si tratta dunque di un corpus abbastanza voluminoso (quasi 5 milioni di parole), ma eterogeneo e senza alcuna pretesa di bilanciamento (Meyer 2002: xii).

In ogni caso, indispensabile è la trasparenza, ovvero informare gli utenti sulle caratteristiche del corpus di lavoro. Ritorna dunque il concetto di sistematicità. Anche Sinclair ne sottolinea l'importanza, affermando che, dal momento che un corpus è anche soggettivo, in quanto molto dipende dal ricercatore che lo progetta, all'utente non dovrebbero bastare i risultati della sua ricerca: è necessario risalire anche alle ragioni che hanno condotto alla compilazione di un dato corpus. La documentazione di qualsiasi scelta compiuta nel processo di progettazione e compilazione di un corpus diviene dunque un momento imprescindibile (Sinclair 2004a: par. 4).

In generale il concetto di rappresentatività, come ricorda Rossini Favretti, è uno dei criteri più dibattuti nell'ambito della LC e di difficile risoluzione, anche perché, secondo l'autrice, si lega ad altri criteri oltre alla finalità: ad esempio la dimensione e la disponibilità dei testi. Alla base della progettazione di CORIS vi era l'interrogativo su quanto dovesse essere vasto un corpus per essere rappresentativo. Rossini Favretti fa notare che la variabile della dimensione si differenzia in base al periodo storico e all'avanzamento tecnologico. I primi corpora erano in grado di elaborare un numero limitato di dati, e molti testi non erano affatto disponibili come avviene invece oggi. Attualmente la questione della dimensione implica spesso il problema opposto: date le enormi potenzialità informatiche e la maggiore reperibilità di testi, può sembrare spesso difficile decidere cosa escludere.

La rappresentatività, tuttavia, non si riduce esclusivamente alle dimensioni di un corpus o alla varietà delle tipologie testuali in esso incluse. Nella guida al BNC si distingue tra testo prodotto e testo recepito (Burnard 2000: 7). Non tutti i testi redatti in una lingua sono necessariamente recepiti dall'intera comunità linguistica, ossia da quanti condividono una determinata lingua storico-naturale. Se un corpus di riferimento vuole essere “una raccolta di testi autentici e ricorrenti nell'uso, in formato elettronico e

⁹⁷ <https://www.cis.upenn.edu/~treebank/> (settembre 2015).

⁹⁸ Il *parsing* è un processo di analisi linguistica impiegato per ricostruire la struttura sintattica di una frase, ossia i costituenti e le relazioni di dipendenza grammaticale (Calzolari, Lenci, 2004: 58).

rappresentativi di uno Stato o di una varietà della lingua” (Rossini Favretti 2000: 41), sarà d’obbligo l’inclusione di testi largamente recepiti dal pubblico. Abbiamo già affermato, all’inizio di questo paragrafo, che i testi appartenenti ad un dato genere dovrebbero essere in numero proporzionale alla sua effettiva diffusione. Ma non solo. La questione qui esposta concerne anche testi riconducibili ad un unico tipo testuale. Ad esempio, per quanto riguarda i testi letterari, un’opera, o le opere di un autore particolarmente noto, talvolta plasmano il bagaglio culturale di un popolo e servono da modello, anche dal punto di vista linguistico, per diverse generazioni. Tali opere saranno dunque inserite senza dubbio in un corpus rappresentativo di una lingua. Nel caso di testi contemporanei, uno dei criteri di selezione può essere il numero di vendite, preferendo un best-seller ad un’opera scarsamente acquistata. Lo stesso approccio, come suggerisce Burnard, può essere applicato a giornali e riviste.

Anche per la progettazione di CEXI ci si è avvalsi di questo criterio per la selezione dei testi, scartando opere particolarmente rare o eccessivamente costose e pertanto verosimilmente meno vendute (Zanettin 2002: 332).

D’altro canto, il criterio della ricezione perde senso per alcune tipologie testuali, ad esempio nel caso di documenti specialistici, destinati alla stretta cerchia di coloro che operano nel settore e non al grande pubblico.

1.8.5. Dimensione

La dimensione di un corpus è un aspetto di primaria importanza e coinvolge non solo il numero di parole (*tokens*), ma anche il numero di categorie testuali da includere e la quantità di testi per ogni categoria.

Sembra facile pensare che un corpus, per essere completo, debba includere il maggior numero di testi possibile e, in particolare, appartenenti a più tipologie testuali. Ciò consentirebbe, infatti, di dare una più ampia visione delle strutture linguistiche proprie di un dato genere testuale dal punto di vista lessicale, sintattico e stilistico. In realtà la dimensione dipende essenzialmente, ancora una volta, dallo scopo che guida la progettazione del nostro corpus (Gandin 2009). Per la compilazione di un dizionario “*corpus-based*”, ad esempio, il numero di lemmi dovrà essere certamente elevato. Inoltre, nei casi di polisemia, è auspicabile disporre di almeno un esempio per ognuno dei significati di un lessema (Kennedy 1998). Al contrario, per la costituzione di un corpus *ad hoc*, finalizzato alla ricerca di una struttura linguistica ben precisa, saremo meno

esigenti in termini quantitativi. Si ribadisce tuttavia che in ognuno di questi casi è desiderabile, così come per i sopracitati parametri, stabilire fin dall'inizio una dimensione, quanto meno approssimativa. Se poi si prevede un continuo monitoraggio, è necessario riconsiderare la dimensione di un corpus, che non potrà essere stabilita a priori, ma subirà, piuttosto, variazioni costanti. In alternativa, è possibile la sostituzione invece che l'inserimento di nuovi testi, come era previsto inizialmente per la sezione principale del NKRJa (*osnovnoj korpus*), la cui dimensione target stabilita in origine era di 100 milioni di parole. Rachilina (2009: 11) spiega che il monitoraggio del corpus non avrebbe dovuto comportare una variazione di questa cifra. Nel caso di aggiunta di nuovi testi, risalenti ad esempio al periodo 2005-2010, si pensava di rimuovere testi più datati o meno significativi includendoli in una banca dati indipendente, ugualmente accessibile, ma non più parte del corpus stesso. Questa decisione, tuttavia, sembra essere stata modificata nel corso degli anni, dal momento che ad oggi l'*osnovnoj korpus* conta più di 260 milioni di parole e l'intero NKRJa, compresi i sotto-corpora, ammonta a circa 600 milioni di parole.

Al criterio della dimensione si lega anche la scelta se includere testi integrali o campioni, aspetto che affronteremo nel prossimo paragrafo. Come spiega Zanettin (2002) in riferimento alla progettazione di CEXI, una volta stabilita la dimensione approssimativa del corpus che si andrà a comporre, sarà necessario scegliere se includere un minor numero di testi in formato integrale, oppure più testi in formato ridotto. A livello di rappresentatività, se il corpus non prevede di raggiungere notevoli dimensioni, la seconda opzione potrebbe essere preferibile. Quest'ultimo è stato il caso di CEXI, che con 4 milioni di parole prevedeva di includere solo campioni.

Infine, per un confronto utile alla progettazione di un CP, segnaliamo che l'ENPC è composto da testi che vanno dalle 10 mila alle 15 mila parole per un totale di 200 estratti, ossia 100 per lingua. Il numero complessivo di parole si aggira attorno ai 2 milioni e mezzo (Johansson, Ebeling, Oksefjell 2002).

L'aspetto puramente quantitativo tuttavia non è sufficiente per fare la differenza. Come già accennato, il BrC, con il suo milione di parole, rappresenta comunque un esempio assai valido di corpus di prima generazione. I corpora cosiddetti di terza generazione, comprendenti centinaia di milioni di parole, possono arrivare ad includere testi raccolti senza alcun criterio di selezione, o meglio sulla base di un unico criterio: reperire il maggior numero di testi possibile in base alla disponibilità. E, per citare direttamente Lecch, ricordiamo che “*a collection of machine-readable text does not make*

*a corpus*⁹⁹ (Leech 1991: 10). La quantità, insomma, non dovrebbe mai oscurare l'aspetto qualitativo. Anche Rossini Favretti riporta l'esempio del BrC, che nonostante le sue modeste dimensioni rispetto ai grandi corpora odierni, non ha mai smesso di fungere da ottimo strumento per le ricerche relative all'inglese americano scritto all'inizio degli anni Sessanta.

1.8.6. Corpus di testi completi o di campioni

Per questioni di spazio o di *copyright*, così come per svariate scelte legate ai criteri sopraelencati, non tutti i corpora sono costituiti da testi in versione integrale. Come ricorda Stubbs (1996: 28), in base al principio per cui lo studio della lingua dovrebbe considerare testi interi e non proposizioni isolate, l'idea di inserire degli estratti è stata spesso criticata. Eppure nella maggior parte dei corpora presi qui in esame sono stati inseriti campioni di dimensione predefinita piuttosto che interi testi. In molti casi, una volta decise le tipologie testuali, il numero di testi per ciascuna categoria e le rispettive dimensioni, i campioni vengono selezionati in maniera casuale.

Il BrC è costituito da estratti di 2000 parole appartenenti a 15 generi testuali differenti. Le dimensioni relativamente ridotte del corpus (ricordiamo, 1 milione di parole) hanno favorito la propensione verso tale modalità, così da garantire l'inserimento di un maggior numero di testi, mantenendo inoltre un bilanciamento a livello di numero di parole tra testi di diverso tipo.

Analogamente, il BNC è un corpus di campioni che non superano le 45.000 parole per testi non letterari e le 40.000 parole per opere letterarie. Tali cifre sono, naturalmente, indicative, dal momento che si è cercato di non interrompere gli estratti nel bel mezzo di una frase o di un pensiero (Burnard 2000).

CEXI avrebbe dovuto contenere 320 campioni da circa 12.500 parole ciascuno. La scelta presenta un'analogia con il caso del BrC ed è infatti dipesa soprattutto dalle dimensioni stabilite per il corpus: l'introduzione di testi interi, per un totale di 4 milioni di parole, avrebbe consentito di caricare solo un massimo di 15 testi, compromettendo così ogni pretesa di rappresentatività (Zanettin 2002). Anche nel caso dell'ENPC, la scelta di includere degli estratti anziché opere complete è stata dettata dalla volontà di bilanciare

⁹⁹ "Una collezione di testi elaborabili da una macchina non fa di per sé un corpus".

il più possibile il corpus, includendo testi scritti e tradotti da un numero significativo di autori e traduttori (Johansson, Ebeling, Oksefjell 2002).

Diversa è la situazione di CORIS, per cui si è scelto di privilegiare l'impiego di testi interi. Come spiega Rossini Favretti, la presenza di campioni garantisce omogeneità e standardizzazione, ma allo stesso tempo compromette sia la corretta contestualizzazione dei dati, influenzati dalla selezione fatta dal compilatore, sia la rappresentatività (Rossini Favretti 2000: 47). Vediamo dunque che, nel caso di CORIS e di CEXI, due scelte opposte hanno in realtà la pretesa di soddisfare il medesimo criterio: la rappresentatività. Si conferma, dunque, la profonda complessità di tale concetto, sempre influenzato da numerosi parametri, come, in questo caso, la dimensione del corpus.

1.8.7. Comparabilità

Un aspetto meno discusso, ma ugualmente rilevante per il nesso diretto con il principio della sistematicità, è quello della comparabilità di un corpus. Per essere tale, un corpus dovrebbe essere confrontabile con altri corpora, ed è questo uno dei segni del suo essere sistematico. Di comparabilità parla anche Rossini Favretti nella guida a CORIS/CODIS, in particolare relativamente alla scelta dei sotto-corpora destinati a comporre un corpus generale. Il BNC, ad esempio, nasce con una serie di caratteristiche stabilite secondo una metodologia chiara e accessibile e non a caso Burnard afferma: *“This approach makes it possible for other researchers and corpus compilers to review, emulate or adapt concrete design”*¹⁰⁰ (Burnard 2000: 6). Ne risulta che ogni nuovo corpus è compilato sulla base dei precedenti; per esempio, il LOB corpus, apertamente costruito sul modello del BrC, è l'equivalente per l'inglese britannico. Nel caso dei CP, come già accennato, molti corpora esistenti, tra cui CEXI, sono stati progettati basandosi su di un modello preciso, ossia l'ENPC. Tale pratica consente inoltre, per quanto possibile, di colmare continuamente le lacune createsi.

Ogni corpus, di per sé, dovrebbe soddisfare il criterio della comparabilità e risultare potenzialmente ripetibile. Anche per questo motivo è fondamentale che la metodologia impiegata e le scelte di progettazione siano chiaramente esplicitate.

¹⁰⁰ “Questo approccio consente ad altri ricercatori o compilatori di corpora di rivedere, imitare o adattare il proprio progetto di compilazione”.

1.8.8. Tipologia testuale e classificazione dei testi

Anche la classificazione dei testi di un corpus dipende dalla finalità. Il BNC, ad esempio, raccoglie i suoi testi in base al dominio (ossia l'argomento, per i testi non letterari. Questi ultimi rientrano in un dominio a parte, sotto la voce "testi immaginari"), alla data e al mezzo (se si tratta di un libro, un periodico, un quotidiano ecc.). Una simile classificazione soddisfa il criterio di comparabilità in quanto presenta una struttura facilmente imitabile.

Se si considera invece l'ENPC, molto più limitato rispetto al BNC, si vedrà che la classificazione dei testi si basa semplicemente sulla distinzione tra fiction e non-fiction. Il metodo può funzionare nel caso di corpora costituiti da pochi milioni di parole, dove i testi non letterari non raggiungono un numero sufficiente per poter essere suddivisi in sotto-corpora ulteriori.

Prima di essere classificati, i testi vengono selezionati seguendo i criteri prefissati. Per prima cosa viene stabilita la tipologia. In seguito si identifica il mezzo, ovvero il tipo di fonti a cui si intende attingere: giornali, riviste, libri (per cui si può ulteriormente distinguere tra libri pubblicati e non), libri in versione elettronica ecc. Una volta prescelto il mezzo, è utile definire nuovi dettagli. Ad esempio, gli autori del BNC stabilirono che i libri dovessero rispondere ad una serie di requisiti: dovevano essere stati pubblicati da una casa editrice britannica, non dovevano costare più di una data somma, dovevano avere un certo numero di pagine, e infine dovevano aver avuto un numero prestabilito di vendite e quindi essere stati largamente recepiti dal pubblico. Per soddisfare quest'ultimo criterio è sembrato utile prediligere *best-seller* e vincitori di premi, oppure ci si è basati sulle statistiche relative ai prestiti bibliotecari (Burnard 2000).

I progettisti di CEXI, invece, prevedevano di attingere solamente a libri per adulti, escludendo dunque pubblicazioni per bambini o semplificate. Si tratta, in questo caso, di una scelta convenzionale. Inoltre, per garantire una certa omogeneità, sarebbero stati inclusi solo testi in prosa (Zanettin 2002: 332). Nel caso di CP, tuttavia, i criteri che determinano la selezione dei testi devono tener conto anche della traduzione (cfr. par. 9).

1.9. Criteri specifici per la progettazione di corpora bilingui paralleli

I criteri esaminati fino a questo punto sono di carattere sufficientemente generale per poter essere applicati alla progettazione di vari tipi di corpora, tanto monolingui, quanto bilingui. La compilazione di un corpus parallelo richiede però di considerare una serie di aspetti che risultano, in altri casi, per nulla o poco pertinenti e che esporremo in questo paragrafo.

Nel caso di corpora bilingui paralleli la selezione dei testi dipende in primo luogo dall'esistenza di una traduzione nella lingua d'arrivo. Questo aspetto, purtroppo, compromette le dimensioni del corpus in questione, che non potrà mai raggiungere i volumi solitamente garantiti da qualsivoglia corpus generale di terza generazione (Frankenberg-Garcia 2009).

Il secondo passo consiste nel definire i dettagli ulteriori, in parte già citati in precedenza, quali l'anno di pubblicazione, il grado di ricezione da parte del pubblico, l'attendibilità della traduzione. L'Index Translationum¹⁰¹, un database dell'UNESCO che registra quasi¹⁰² tutti i libri tradotti dal 1979 al 2012, offre un elenco delle traduzioni esistenti divise per generi. Zanettin spiega che per CEXI 1500 titoli offerti dall'Index Translationum rispondevano ai requisiti predefiniti. Tuttavia, sulla base di alcuni calcoli preliminari, ci si aspettava di ottenere il permesso da case editrici ed autori per circa 200 titoli. Come già spiegato, il problema legato ai diritti d'autore riguarda qualsiasi tipo di corpus, ma nel caso di CP è evidentemente più complesso per il fatto che i permessi da richiedere sono due: testo originale e traduzione in lingua d'arrivo.

Accanto a queste considerazioni introduttive, esaminiamo di seguito due criteri particolarmente rilevanti nella compilazione di un corpus parallelo: la validità della traduzione e il bilanciamento tra le due lingue.

A. Validità della traduzione

Il primo criterio riguarda la qualità della traduzione in nostro possesso. È necessario assicurarsi che il traduttore sia affidabile. Zanettin (2002) osserva che per la compilazione

¹⁰¹ <http://www.unesco.org/xtrans/> (gennaio 2017).

¹⁰² Purtroppo abbiamo notato delle lacune.

di CEXI sono state evitate auto-traduzioni o traduzioni che non fossero state svolte da un parlante nativo della lingua d'arrivo.

Va inoltre sottolineato che i classici della letteratura spesso presentano più di una traduzione e in alcuni casi i numeri sono piuttosto importanti. Un esempio significativo, nel caso di traduzioni dal russo all'italiano, è quello del *Cappotto* di Gogol', che vede in Italia più di 30 traduzioni diverse, sulla base di una recente ricerca di Olga Inkova (2014, 2016). In questo caso risulta difficile selezionare la migliore e sembra più utile proporre la disposizione parallela di diverse traduzioni dell'opera, così da permettere analisi contrastive. Malmkjær afferma addirittura che includendo in un CP una sola traduzione gli studi di traduttologia ne risultano fortemente limitati e suggerisce, pertanto, di includere il maggior numero di traduzioni possibili per una stessa opera (Laviosa 2002: 28). Un'opera come *Il Cappotto*, per tornare all'esempio citato, dovrebbe dunque essere inclusa in un corpus per due motivi: innanzitutto, perché un numero così elevato di traduzioni testimonia l'importanza del racconto di Gogol' nella cultura d'arrivo; in secondo luogo, per rendere possibile agli studiosi di traduttologia il confronto automatico e computerizzato tra diverse traduzioni, consapevoli del fatto che si tratta di un ambito di ricerca finora poco esplorato, ma certamente ricco di potenzialità. Si può concludere, dunque, che uno dei criteri per la selezione delle opere dovrebbe considerare la quantità di traduzioni esistenti.

B. Bilanciamento tra le due lingue

Una questione fondamentale nella creazione di un corpus parallelo bi-direzionale, ma anche di difficile soluzione, è quella del bilanciamento tra il numero di testi nelle due lingue.

In un corpus $A \rightarrow B / B \rightarrow A$ le traduzioni da lingua A a lingua B dovrebbero essere in egual numero rispetto alle traduzioni da lingua B a lingua A. Per alcune coppie di lingue, tuttavia, risulta piuttosto difficile raggiungere un tale bilanciamento. Ad esempio, è estremamente semplice trovare, in qualsivoglia Paese, traduzioni dall'inglese, ma purtroppo non è vero il contrario. È un problema di cui era cosciente anche Zanettin, il quale afferma di aver incontrato non poche difficoltà in questo senso nella compilazione di CEXI (Laviosa 2002: 41). Nelle situazioni di forte sbilanciamento tra lingua A e lingua B, appare plausibile abbandonare l'idea di compilare un corpus bi-direzionale, optando,

invece, per una soluzione uni-direzionale. Fortunatamente, nel caso del russo e dell'italiano esistono molte opere tradotte in entrambe le direzioni, anche se in seguito segnaleremo comunque un certo sbilanciamento.

1.10. Proposta per un ampliamento del corpus parallelo italiano-russo del NKRJa

Considerati i limiti del primo corpus parallelo italiano-russo del NKRJa (d'ora in poi "corpus pilota"), formuliamo alcune proposte per un suo adeguato ampliamento e perfezionamento. In questo paragrafo esporremo il nostro progetto di ampliamento iniziale. Nel secondo capitolo verrà poi illustrato quanto effettivamente realizzato.

È stato indispensabile porsi obiettivi realistici, tenendo presente che, data la sua mole, difficilmente il lavoro, così come descritto nel presente paragrafo, sarebbe stato realizzabile per intero durante i tre anni del Dottorato di Ricerca. Lo scopo è stato pertanto quello di impostare le basi per un progetto di ampliamento a lungo termine.

Inoltre, ribadiamo che il nostro progetto si inserisce in un contesto di più ampio respiro (il NKRJa) realizzato da diversi gruppi di lavoro, sia in Russia sia in Italia (cfr. cap. 2). Un tipo frequente di contributo, diffuso anche per altri corpora, è la messa a disposizione, da parte di collaboratori esterni, di testi già allineati. Sebbene questo coordinamento delle forze sia alla base della costituzione di un corpus ampio, la collaborazione volontaria presenta il pericolo che si aggiungano al corpus testi che non rispondano ai criteri da noi elencati, mancando così alcuni degli obiettivi preposti ¹⁰³.

La nostra proposta ha tenuto in considerazione i seguenti punti:

- a) Finalità e tipologia
- b) Dimensione
- c) Scelta delle tipologie testuali
- d) Corpus sincronico o diacronico
- e) Corpus di estratti o di testi interi
- f) Selezione dei testi e delle traduzioni per ogni tipologia testuale
- g) Rappresentatività e bilanciamento
- h) Sistematicità

¹⁰³ Nel secondo capitolo presenteremo nel dettaglio una situazione di questo tipo.

a. *Finalità e tipologia*

Il Corpus di testi paralleli in russo e italiano del NKRJa, anche nella versione di cui abbiamo proposto l'ampliamento, continuerà ad essere un corpus parallelo bi-direzionale. Presenterà dunque un certo numero di testi originali in italiano e russo, allineati con la rispettiva traduzione; sarà un corpus parallelo generale, e non specialistico, perché la presenza di più generi testuali consente un maggior numero di applicazioni future; e, infine, ha l'obiettivo di diventare un “*target variant corpora*” (cfr. par.1.4.), costituito, cioè, da più traduzioni di una stessa opera dove presenti.

Riguardo ai generi da inserire, ricordiamo che è molto difficile, se non impossibile, trovare un buon numero di testi specialistici tradotti nella coppia di lingue dai noi considerata (russo e italiano); sarebbe più facile con l'inglese, in quanto usato come lingua franca in tutti gli ambiti scientifici o giuridico-amministrativi a livello internazionale. Il nostro corpus conterà, dunque, testi appartenenti a diversi generi dello scritto. L'opzione di includere trascrizioni di lingua parlata è un'idea contemplabile, ma considerata la difficoltà di trovare trascrizioni di lingua parlata tradotte in entrambe le direzioni per la coppia di lingue in oggetto, al momento questo non può essere fissato come obiettivo. Una soluzione potrebbe prevedere l'inserimento di copioni di film con sottotitoli tradotti, poiché il film tende a riprodurre il linguaggio parlato. Come abbiamo visto, *OpenSubtitles* rende disponibili sottotitoli con la rispettiva traduzione, ma tale opzione andrebbe attentamente ponderata dal momento che le traduzioni sono svolte da utenti volontari, anziché da traduttori esperti.

Per quanto concerne la finalità, il corpus, dato il suo carattere generale, potrà fornire materiale utile per la linguistica contrastiva, la traduttologia, le applicazioni glottodidattiche, la linguistica computazionale, la filologia, la linguistica storica e gli studi letterari.

b. *Dimensione*

Considerando gli esempi di CP analizzati e i limiti di tempo che circoscrivono il nostro lavoro, abbiamo fissato un obiettivo preliminare volto al raggiungimento di 3 milioni di nuove parole in tre anni. Non è stato possibile stabilire con esattezza il numero di documenti necessari a raggiungere questo scopo poiché ciò dipende dalla loro lunghezza.

Stabilendo una lunghezza fissa per gli estratti che vada dalle 10.000 alle 15.000 parole, si otterrebbe una cifra che va dai 200 ai 300 testi. Considerato però che alcuni di essi, soprattutto specialistici, potrebbero presentare un numero di parole inferiore, e che per alcune opere letterarie l'esigenza di non interrompere un capitolo a metà potrebbe portare ad avere estratti che superino le 15.000 parole, una previsione al dettaglio era prematura all'inizio del lavoro di ampliamento.

c. Scelta delle tipologie testuali

Come già menzionato, il corpus intende essere generale ed abbracciare diverse tipologie testuali, la scelta delle quali dipende dall'effettivo scambio tra russo e italiano e quindi dalla presenza di testi tradotti da lingua A a lingua B e viceversa.

Escludendo la lingua parlata, per i motivi sopra esposti, il primo genere di cui abbiamo deciso di occuparci è quello letterario. Ciò è giustificato innanzitutto dal fatto che il primo nucleo del corpus in origine era composto da pochi testi letterari. Il primo obiettivo voleva pertanto essere quello di ampliare questa sezione così che potesse raggiungere dimensioni accettabili per la ricerca. Inoltre, i testi letterari non solo sono maggiormente reperibili, ma spesso presentano, in particolar modo se datati, più di una traduzione.

Una seconda sezione includerà testi di saggistica, ma essendo più ristretto lo scambio di saggistica tra russo e italiano, la scelta è limitata a quanto effettivamente reperibile. I titoli sono stati scelti tra gli elenchi dell'Index Translationum, che divide i libri a seconda della materia trattata: bibliografie (1 libro tradotto dal russo all'italiano, 3 libri dall'italiano al russo), filosofia-psicologia (47 ru-it, 67 it-ru), religione-teologia (42 ru-it, 29 it-ru), giurisprudenza-scienze sociali-istruzione (360 ru-it, 110 it-ru), scienze naturali-scienze esatte (88 ru-it, 32 it-ru), scienze applicate (19 ru-it, 72 it-ru), arte-giochi-sport (101 ru-it, 150 it-ru), storia-geografia-biografie (159 ru-it, 66 it-ru).

A questo elenco si aggiungono poi le opere letterarie. L'Index Translationum consente di rilevare subito una discrepanza tra lingua fonte e lingua d'arrivo: i libri in lingua russa tradotti in italiano sono in numero maggiore rispetto ai libri in italiano tradotti in russo. Ma la disuguaglianza è accettabile, e in ogni caso non ha costituito un problema, poiché consapevoli che la quantità di testi che sarebbe stata di fatto inserita nel corpus sarebbe stata comunque inferiore al numero complessivo delle traduzioni esistenti

dall'italiano al russo. Escludendo la prima categoria (le bibliografie), per via dei numeri molto esigui e della sua scarsa utilità per gli studi linguistici e traduttologici, includendo 10 estratti di 10-15 mila parole per ogni categoria (5 ru-it e 5 it-ru), contando originali e traduzioni abbiamo calcolato di ottenere un corpus che solo per la saggistica avrebbe raggiunto un milione e mezzo/due milioni di parole. Si tratterebbe di un risultato estremamente significativo, che si spera di poter ottenere negli anni futuri, ma eravamo quasi certi di non poterlo realizzare a breve termine. Pertanto, per questioni di bilanciamento, ci è sembrato ottimale cominciare con l'introdurre due testi per ogni categoria (il primo tradotto dal russo all'italiano, e il secondo dall'italiano al russo).

Inizialmente si pensava di poter inserire testi del settore giuridico-amministrativo, ossia documenti di enti ed organizzazioni internazionali, contratti, legislazioni. Tuttavia, per questo tipo di documenti è difficile ottenere originali e traduzioni nella coppia di lingue da noi considerata.

d. Corpus sincronico o diacronico

Per quanto riguarda la diacronia e la sincronia, abbiamo formulato proposte differenti in base alle diverse tipologie testuali che avrebbero costituito il corpus.

Nella sezione dei testi letterari abbiamo pensato di inserire opere appartenenti a periodi storici differenti: dai classici dell'800 fino a romanzi pubblicati nell'ultimo decennio. Questo sotto-corpus non sarebbe, pertanto, sincronico, ma diacronico. Al contrario, per quanto concerne le altre tipologie testuali, si è pensato di inserire in un primo momento testi più recenti, risalenti ai secoli XX e XXI.

e. Corpus di estratti o di testi interi

La maggior parte dei corpora sono costituiti da campioni di lunghezza predefinita, piuttosto che da testi integrali. La necessità di includere testi interi piuttosto che campioni dipende dall'uso che ci si prefigge. Come già accennato, nel caso di un testo letterario, si può verificare la necessità di analizzare la traduzione di una determinata parola o struttura linguistica nel corso di tutto il romanzo. Nel caso di testi settoriali, invece, potrebbe darsi la necessità di verificare la traduzione di strutture ricorrenti in parti specifiche del testo, quali introduzione o conclusioni. Nel caso di campioni estratti in maniera casuale, un

simile tipo di ricerca è da escludere a priori. Tuttavia, come ricorda Zanettin, la questione della dimensione non è da sottovalutare: dal momento che il nostro corpus non prevedeva in tre anni di raggiungere volumi particolarmente considerevoli, abbiamo deciso di prediligere la presenza di un maggior numero di testi, seppur non interi e quindi più brevi. Nel nostro caso, principalmente per questioni di tempo, si è stabilito che le 10.000-15.000 parole sopracitate potessero essere un traguardo iniziale ragionevole. Il numero esatto dipende poi dall'effettiva lunghezza dei testi caricati. Nel caso di racconti, o di documenti non letterari, l'ammontare complessivo di parole potrebbe rimanere al di sotto della suddetta soglia. In questi casi i testi verrebbero inseriti nella loro interezza.

f. Selezione dei testi e delle traduzioni per ogni tipologia testuale

Si è deciso che la selezione dei testi si sarebbe basata su di un certo numero di parametri, primo fra tutti la reperibilità della traduzione. Per i testi letterari si è stabilito di scegliere tra i romanzi russi e italiani rispettivamente tradotti nell'altra lingua, preferendo i romanzi che hanno avuto una fortuna particolare nel Paese d'arrivo. Quest'ultimo criterio è stato stabilito sulla base di:

- numero di traduzioni presenti: preferenza per romanzi tradotti da almeno cinque traduttori diversi, con il conseguente obiettivo di inserire tutte le traduzioni nel caso in cui i numeri non siano troppo elevati (oltre le cinque traduzioni).
- cataloghi delle case editrici per opere più recenti, così da verificare cosa è stato tradotto e possibilmente più venduto, ove tali informazioni siano disponibili.

In questa fase iniziale dell'ampliamento abbiamo ritenuto opportuno inserire una sola opera per autore, preferibilmente la più tradotta o letta, senza escludere che in una fase successiva si possa procedere con l'aggiunta di ulteriori opere per lo stesso autore.

Nel caso di testi non letterari, si è stabilito come criterio fondamentale la qualità della traduzione, preoccupandosi innanzitutto che il traduttore sia esperto e madrelingua della lingua d'arrivo. Nel caso di articoli di giornale (per l'ampliamento a lungo termine), si selezioneranno i domini in base ai settori che legano maggiormente Russia e Italia, sebbene l'effettiva reperibilità dei testi metta a volte un limite oggettivo a questo criterio.

g. Rappresentatività e bilanciamento

Includendo nel CP testi e traduzioni di tipi diversi, pensiamo di compiere già un discreto passo verso la tanto auspicata rappresentatività dei corpora generali. Tuttavia, affinché un tipo testuale non venga maggiormente rappresentato rispetto a altri, idealmente si dovrebbe pensare di distribuire i testi equamente. In realtà è risaputo che nel caso di un CP, è molto più facile reperire la traduzione di un testo letterario, piuttosto che di altri tipi di testi. Per questo motivo si è deciso di partire dall'inserimento di testi letterari, come già accennato, per poi cercare di non trascurare la ricerca e l'allineamento di altri tipi di testi. Infine, sempre per ragioni di bilanciamento, si è ritenuto di dover scegliere opere letterarie risalenti a diverse epoche: a partire dal XIX secolo fino ai giorni nostri.

Per quanto riguarda il bilanciamento tra testi in lingua A e testi in lingua B, bisogna tener presente che esiste uno sbilanciamento tra russo e italiano legato al numero di traduzioni di una stessa opera: mentre alcune opere russe sono state tradotte in italiano più di venti volte da traduttori diversi, non è accaduto mai il contrario. Esistono opere italiane tradotte varie volte in russo, ma il numero di traduzioni si aggira tra 3 e 5. Solo in un caso esistono per un singolo romanzo italiano 8 traduzioni in lingua russa (*Spartaco* di R. Giovagnoli).

h. Sistematicità

Da ultimo, per rispettare il criterio della sistematicità abbiamo deciso di annotare ogni passo del processo di compilazione, nonché ogni cambiamento o aggiunta rispetto alla presente proposta.

Bibliografia

- AA.VV. *Introduction to the Brown Corpus* (Tratto da: <http://www.uni-giessen.de/cms/faculties/f05/engl/ling/help/materials/restricted/brown>) (settembre 2015).
- Aarts B. (2000). *Corpus Linguistics, Chomsky and Fuzzy Trees Fragments*. In Mair C., Hundt M. *Corpus Linguistics and Linguistics Theory*. Papers from the Twentieth International Conference on English Language Research on Computerized Corpora, (ICAME 20), Freiburg, Breisgau 1999. Amsterdam, Atlanta: Rodopi, pp. 5-13.
- Aijmer K., Altenberg B. (2000). *The English-Swedish Parallel Corpus: A resource for contrastive research and translation studies*. In Mair C., Hundt M. (eds.), *Corpus Linguistics and Linguistics Theory*. Papers from the Twentieth International Conference on English Language Research on Computerized Corpora, (ICAME 20), Freiburg, Breisgau 1999. Amsterdam, Atlanta: Rodopi, pp. 15-33.
- Andreeva E. G., Kasevič V. (2005). *B. Grammatika i leksika (na materiale anglo-russkogo korpusa parallel'nyh tekstov)*. In *Sovremennye informacionnye tehnologii i filologija*. Moskva: IMLI RAN, pp. 297-307.
- Anthony, L. (2013). *A critical look at software tools in corpus linguistics*. *Linguistic Research*, Volume 30, Issue 2, pp 141-161.
- Aston G. (2002). *The learner as corpus designer*. In Kettermann B., Marko G. (eds.), *Teaching and Learning by Doing Corpus Analysis*. Proceeding of the Fourth International Conference on Teaching and Language Corpora, Graz 19-24 July 2000, Amsterdam, New York: Rodopi, pp. 9-25.
- Aston G. (2000). *I corpora come risorse per la traduzione e per l'apprendimento*. In Bernardini S., Zanettin F. (a cura di), *I corpora nella didattica della traduzione*. Forlì: Biblioteca della Scuola Superiore di Lingue Moderne per Interpreti e Traduttori, pp. 21-29.
- Baker M. (1993). *Corpus linguistics and translation studies – Implications and applications*. In Baker M., Francis G., Tognini-Bonelli E. (eds.), *Text and technology. In honour of John Sinclair*. Philadelphia, Amsterdam: John Benjamins Publishing, pp 233-250.

- Baker M. (2000). *Linguistica dei corpora e traduzione. Per un'analisi del comportamento linguistico dei traduttori professionisti*. In Bernardini S., Zanettin F. (a cura di), *I corpora nella didattica della traduzione*. Forlì: Biblioteca della Scuola Superiore di Lingue Moderne per Interpreti e Traduttori, pp 31-44.
- Baker P. (2010). *Sociolinguistics and Corpus Linguistics*. Edinburgh: Edinburgh University Press.
- Barbera M. (2013). *Linguistica dei corpora e linguistica dei corpora italiana. Un'introduzione*. Milano: Qu.A.S.A.R. s.r.l.
- Baroni M, Bernardini S., Ferraresi A., Zanchetta E. (2009). *The WaCky Wide Web: a collection of very large linguistically processed web-crawled corpora*. *Language Resources and Evaluation* 43 (3): 209-226.
- Beeching K. (2013). *A parallel corpus approach to investigating semantic change*. In Aijmer K., Altenberg B. *Advances in corpus-based Contrastive Linguistics*. Amsterdam, Philadelphia: John Benjamins Publishing, pp. 103-125.
- Benko V., Zacharov V. P. (2016). *Very large corpora: new opportunities and new challenges*. In Komp'juternaja lingvistika i intellektual'nye tehnologii: po materialam ežegodnoj meždunarodnoj konferencii "Dialog", 15, Moskva: RGGU, pp. 83-98.
- Bernardini S. (2004). *Corpora in the classroom. An overview and some reflections on future developments*. In Sinclair, J. (ed.), *How to use corpora in language teaching*. - Amsterdam [u.a.]: Benjamins.
- Biber D., Conrad S., Reppen R. (1998). *Corpus Linguistics*. Cambridge: Cambridge University Press.
- Borin L. (2002) ... *And never the twain shall meet?* In Borin L. (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: Rodopi, pp. 1-43.
- Botley S., McEnery T., Wilson A. (eds), (2000). *Multilingual Corpora in Teaching and Research*. Amsterdam, Atlanta: Rodopi.
- Burnard L. (ed.) (2000). *Reference Guide to the British national Corpus (world edition)*. (<http://www.natcorp.ox.ac.uk/archive/worldURG/urg.pdf>) (gennaio 2017).

- Busa R. (2009). *Dall'Index Thomisticus alle Lingue disciplinate: Sessant'anni di studio computazionale delle lingue*. In "Linguistica e modelli tecnologici di ricerca", SLI 51. Roma: Bulzoni, pp. 19-23.
- Calzolari N., Lenci A. (2004). *Linguistica computazionale. Strumenti e risorse per il trattamento automatico della lingua*. In *Mondo Digitale* n.2, pp. 56-69.
- Chomsky N. (1957). *Syntactic structures*. The Hague: Mouton.
- Chomsky N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: The M.I.T. Press.
- Davies M. (2009). The 385 + million-word *Corpus of Contemporary American English (1990-2008+)*. *Design, architecture and linguistic insights*. In *International Journal of Corpus Linguistics* 14 (2). John Benjamins Publishing Company. Pp. 159-190.
- Davies M. (2012). *Expanding horizons in historical linguistics with the 400-million-word Corpus of Historical American English*, *Corpora*, Volume 7, Issue 2, pp 121-157.
- Dobrovol'skij D. O., Kretov A. A., Šarov C. A. (2005). *Korpus parallel'nyh tekstov. Arhitektura vozmožnosti ispol'zovanija*. In *Nacional'nyj korpus ruskogo jazyka: 2003-2005*. Moskva: Indrik. Pp. 263-296.
- Dobrovol'skij D. O. (2009). *Korpus parallel'nych tekstov v issledovanii kul'turno-specifičnoj leksiki*. In *Nacional'nyj korpus ruskogo jazyka: 2006—2008. Novye rezul'taty i perspektivy*. Sankt-Peterburg: Nestor-Istorija, pp 383—401.
- Dobrovolskij D. O., Levontina I. B. (2009). *Russkoe net, nemeckoe nein, angliskoe no: sopostavitel'noe issledovanie semantiki na baze parallel'nych korpusov*. In *Komp'juternaja lingvistika i intellektual'nye tehnologii: po materialam ežegodnoj meždunarodnoj konferencii "Dialog"*, 8, Moskva: RGGU, pp. 97-101.
- Dobrovolskij D. O., Levontina I. B. (2015). *Modal'nye časticy i ideja aktualizacii zabytogo (na materiale parallel'nych korpusov)*. In *Komp'juternaja lingvistika i intellektual'nye tehnologii: po materialam ežegodnoj meždunarodnoj konferencii "Dialog"*, 14, Moskva: RGGU, pp. 104-117.
- Filipovič R. (1982). *Some basic methodological questions of contrastive analysis*. In Calleri D., Marengo C. (a cura di), *Linguistica contrastiva*. Atti del XIII Congresso Internazionale della società di linguistica italiana, Asti, 26-28 maggio 1979. Roma: Bulzoni, pp. 3-16.
- Firth J. R. (1957). *Papers in linguistics 1934-1951*. London: Oxford University Press.

- Fischer K. (2006). *Towards an understanding of the spectrum of approaches to discourse particles: introduction to the volume*. In Fischer K. (ed.), *Approaches to discourse particles*. Oxford, Amsterdam: Elsevier, pp. 1-19.
- Frankenberg-Garcia A., Santos D. (2003). *Introducing COMPARA, the Portuguese-English parallel corpus*. In Zanettin F., Bernardini S., Stewart D. (eds.), *Corpora in Translator Education*. Manchester; St. Jerome, pp 71-87.
- Frankenberg-Garcia, A. (2009). *Compiling and using a parallel corpus for research in translation*. In *International Journal of Translation*, vol. XXI-1, pp 57-71.
- Gandin S. (2009). *Linguistica dei corpora e traduzione: definizioni, criteri di compilazione e implicazioni di ricerca dei corpora paralleli*. Annali della Facoltà, 5, Università degli Studi di Sassari, pp. 133-152.
- Garside, R., Leech, G. and McEnery, T. (1997). *Corpus Annotation. Linguistic information from computer text corpora*. Harlow: Longman.
- Giacoma L. (2012). *Lessicografia bilingue e innovazione: il superamento dei confini della lessicografia tradizionale nella resa dell'intorno linguistico del lemma ne Il nuovo dizionario di Tedesco (Giacoma, Kolb, Zanichelli/Klett: 2009)*. In Ferreri S. (a cura di), *Lessico e lessicologia: atti del 44° congresso internazionale di studi della Società di linguistica italiana (SLI), Viterbo, 27-29 settembre 2010*, pp. 479-486.
- Gilquin G. (2010). *Corpus, cognition and causative constructions*. Amsterdam, Philadelphia: John Benjamins Publishing.
- Graffi G. (2008). *Che cos'è la grammatica generativa*. Roma: Carocci.
- Granger S. (2003). *The corpus approach: a common way forward for contrastive Linguistics and Translation Studies?* In Granger S., Lerot J., Petch-Tyson S. (eds.), *Corpus-based Approaches to Contrastive Linguistics and Translation Studies*. Rodopi: Amsterdam, New York.
- Gurina E. E., Dobrovol'skij D. O., Kretov A. A. (2005). *Asimetrija členenija teksta v parallel'nom korpuse (na materiale romana T. Dradžer "Sestra Kerri")*. Vestnik Voronežnogo Gos. Universiteta. Serija "Lingvistika i mežkul'turnaja komunikacija", № 1, pp. 67-74.
- Halliday M. A. K. (1992). *Language as system and language as instance: the corpus as a theoretical construct*. In Svartvik (ed.), *Directions in Corpus Linguistics*. Berlin: Mouton de Gruyter, pp. 61-78.

- Hanks P. (2009). *The Impact of Corpora on Dictionaries*. In P. Baker (ed.), *Contemporary Corpus Linguistics*, Continuum, London, pp 214-236.
- Hareide L., Hofland K. (2012). *Compiling a Norwegian-Spanish parallel corpus*. In Oakes M. P., Ji M. (eds.), *Quantitative methods in corpus-based Translation Studies*. Amsterdam, Philadelphia: John Benjamins Publishing, pp. 75-113.
- Hunston S. (2002). *Corpora and Applied Linguistics*. Oxford: Oxford University Press.
- Ide N., Suderman K. (2004). *The American National Corpus First Release*. In *Proceedings of the Fourth Language Resources and Evaluation Conference (LREC)*, Lisbon, 1681-84.
- Inkova O. (2014). *Tradurre il titolo: le traduzioni italiane del "Cappotto" di Gogol'*. In *Kwartalnik neofilologiczny*. Université de Varsovie, LXI, 1, pp. 41-56.
- Inkova O. (2016). *Le prime traduzioni romanze del "Cappotto" di Gogol': precisazioni cronologiche e traduttologiche*. In *Diversité et identité culturelle en Europe*, tome 13/2, București: Editura Muzeul Literaturii Române, pp. 93-120.
- Johansson S. (1993). *Some aspects of the recommendations of the Text Encoding Initiative, with special reference to the encoding of language corpora*. In Kytö M., Rissanen M., Wright S. (eds.), *Corpora across the centuries*. Proceedings of the First International Colloquium on English Diachronic Corpora. St. Catherine's College Cambridge, 25-27 March 1993, pp. 203-212.
- Johansson S. (1998). *On the role of corpora in cross-linguistic research*. In Johansson S., Oksefjell S. (eds.), *Corpora and cross-linguistic research: theory, method and case studies*. Amsterdam, New York: Rodopi, pp. 3-24.
- Johansson S., Ebeling J., Oksefjell S. (2002). *English-Norwegian Parallel Corpus: Manual*. Department of British and American Studies: University of Oslo.
- Johansson S. (2003). *Contrastive linguistics and corpora*. In Granger S., Lerot J., Petch-Tyson S., (eds.), *Corpus-based Approaches to Contrastive Linguistics and Translation Studies*. Rodopi: Amsterdam, New York.
- Kennedy G. (1998). *An Introduction to Corpus Linguistics*. Longman: London, New York.
- Kehoe, A., Renouf A. (2002). *WebCorp: Applying the Web to Linguistics and Linguistics to the Web*. WWW2002 Conference, Honolulu, Hawaii (tratto da <http://www.webcorp.org.uk:8080/publications.html>) (dicembre 2015).

- Kettermann B., Marko G. (eds.), *Teaching and Learning by Doing Corpus Analysis*. Proceeding of the Fourth International Conference on Teaching and Language Corpora, Graz 19-24 July 2000, Amsterdam, New York: Rodopi.
- Khokhlova M. V. (2016). *Large Corpora and Frequency Nouns*. In Komp'juternaja lingvistika i intelektual'nye tehnologii: po materialam ežegodnoj meždunarodnoj konferencii "Dialog", 15, Moskva: RGGU, pp. 237-250.
- Křen M. (2009). *The SYN Concept: Towards One-Billion Corpus of Czech*. In Mahlberg M., González-Díaz V., Smith C. (eds) *Proceedings of the Corpus Linguistics Conference*. Liverpool: University of Liverpool (tratto da <http://ucrel.lancs.ac.uk/publications/cl2009/>) (dicembre 2015).
- Laviosa S. (2002). *Corpus-based translation studies*. Rodopi: Amsterdam, New York.
- Leech G. (1991). *The state of the art in corpus linguistics*. In Aijmer K., Altenberg B. (ed.), *English corpus linguistics*. Longman: London, New York.
- Leech G. (1997). *Introducing corpus annotation*. In Garside, R., Leech, G. and McEnery, T. (eds.), *Corpus Annotation. Linguistic information from computer text corpora*. Harlow: Longman, pp. 1-18.
- Lenci A., Montemagni S., Pirelli V. (2007). *Testo e computer: elementi di linguistica computazionale*. Carrocci: Roma.
- Lüdeling A., Evert S., Baroni M. (2005). *Using Web Data for Linguistic Purposes* (tratto da <http://www.webcorp.org.uk:8080/publications.html>) (dicembre 2015).
- McEnery T., Wilson, A. (2001). *Corpus Linguistics*. Edinburgh: Edinburgh University Press.
- McEnery T., Tono Y., Xiao R., (2006). *Corpus-based Language Studies: An Advanced Resource Book*. New York: Routledge.
- McEnery T., Xiao R. (2008). *Parallel and Comparable Corpora: What is happening?* In Anderman G.M., Rogers M. (eds.), *Incorporating Corpora: The Linguist and the Translator*. Clevedon: Multilingual Matters.
- McEnery T., Hardie A. (2012). *Corpus linguistics. Method, Theory and Practice*. New York: Cambridge University Press.
- Meyer C. F. (2002). *English corpus linguistics. An introduction*. Cambridge University Press: Cambridge.

- Morley B. (2006). *WebCorp: a tool for online linguistic information retrieval and analysis*. In Renouf A., Kehoe A. (eds.), *The changing face of Corpus Linguistics*. Rodopi pp. 283-296.
- Nesselhauf N. (2011). *Corpus linguistics: a practical introduction*. (<http://www.as.uniheidelberg.de/personen/Nesselhauf/files/Corpus%20Linguistics%20Practical%20Introduction.pdf>) (febbraio 2016).
- O'Keefe A., McCarthy M., (eds) (2010). *The Routledge handbook of corpus linguistics*. London; New York, NY: Routledge.
- O'Keefe, A. and Walsh, S. (2012). *Applying corpus linguistics and conversation analysis in the investigation of small group teaching in higher education*. *Corpus Linguistics and Linguistic Theory*, Volume 8, Number 1, pp 183-208.
- Olohan M. (2004). *Introducing corpora in translation studies*. London and New York: Routledge.
- Prat Zagrebelsky M. T., (2005). *L'introduzione della corpus linguistics o linguistica dei corpora, nelle università italiane: una ricostruzione 'personale' dagli anni '60 a oggi*. In *Quaderni del CIRSIL*, 4 ([www.lingue.unibo.it /cirsil](http://www.lingue.unibo.it/cirsil)).
- Rachilina E. V. (2009). *Korpus kak tvorčeskij proekt*. In Plungjan V.A. (pod red.), *Nacional'nyj korpus russkogo jazyka: 2006-2008. Novye rezul'taty i perspektivy*. Nestor-Istorija: Sankt-Peterburg, pp. 1-26.
- Reznikova T. I. (2009). *Slavjanskaja korpusnaja lingvistika: sovremennoe sostojanie resursov*. In *Nacional'nyj korpus russkogo jazyka: 2006—2008. Novye rezul'taty i perspektivy*. Sankt-Peterburg: Nestor-Istorija, pp. 402-461.
- Rigotti E., Cigada S. (2004). *La comunicazione verbale*. Milano: Apogeo.
- Rossini Favretti R. (2000). *Progettazione e costruzione di un corpus di italiano scritto: CORIS/CODIS*, in R. Rossini Favretti (a cura di), *Linguistica e informatica. Multimedialità, corpora e percorsi di apprendimento*. Bulzoni: Roma, pp. 39-56.
- Salkie R. (2002). *How can linguists profit from parallel corpora?* In Borin L. (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: Rodopi, pp. 93-109.
- Sharoff S. (2003). *Methods and tools for development of the Russian Reference Corpus*. In *Proc. of Corpus Linguistics Conference*, April, 2003, Lancaster, UK.

- Sičinava D. V. (2005). *Nacional'nyj korpus russkogo jazyka: očerk predystorii*. In *Nacional'nyj korpus russkogo jazyka: 2003-2005*. Moskva: Indrik, pp. 21—30.
- Sinclair J. et al. (eds) (1990) *Collins Cobuild English Grammar*. London: HarperCollins.
- Sinclair J., Renouf A. (1991). *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.
- Sinclair J. et al. (eds.) (1995) *Collins Cobuild Dictionary of Idioms*. London: HarperCollins.
- Sinclair J. (2003). *Reading concordances. An introduction*. London: Pearson.
- Sinclair J. (2004a). *Corpus and Text — Basic Principles*. In Wynne M (ed.), *Developing Linguistic Corpora: a Guide to Good Practice*. Oxford: AHDS Literature Languages and Linguistics. (Tratto da: <http://www.ahds.ac.uk/creating/guides/linguistic-corpora/> (maggio 2014).
- Sinclair J. (2004b). *Trust the text. Language, corpus and discourse*. London, New York: Routledge.
- St. John E. (2001). *A case for using a parallel corpus and concordancer for beginners of a foreign language*. In *Language Learning & Technology*, Vol. 5, 3, pp. 185-203.
- Stubbs M. (1996). *Text and Corpus Analysis*. Oxford: Blackwell.
- Šmídl L., Ircing P. (2014). *Air Traffic Control Communication (ATCC) Speech Corpus*. (Tratto da: <https://www.clarin.eu/content/papers-posters-and-demos-cac2014>) (gennaio 2016).
- Teubert W. (2005). *My version of corpus linguistics*. In *International Journal of Corpus Linguistics*, 10, 1, pp. 1-13.
- Teubert W. (ed) (2007). *Text Corpora and Multilingual Lexicography*. Amsterdam: Benjamins.
- Tiedermann J. (2012). *Parallel Data, Tools and Interfaces in OPUS*. In *Proceedings of the 8th International Conference on Language Resources and Evaluation (LREC'2012)*, pp. 2214-2218 (tratto da http://www.lrec-conf.org/proceedings/lrec2012/pdf/463_Paper.pdf) (dicembre 2016).
- Tognini-Bonelli E. (2001). *Corpus Linguistics at work*. Amsterdam: John Benjamins.
- Tortorelli G. (1996). *«L'Italia che scrive» 1918-1938. L'editoria nell'esperienza di A.F. Formiggini*. Milano: Franco Angeli.

- Zacharov V. P. (2013). *Korpusnaja lingvistika v Rossii*. Retro 2013. Crimea, Ukraine, Peshanoye. (www.retro.crimea.edu/publications/2013_4_zakharov.pdf) (luglio 2015).
- Zaloznjak A. A. (2003). *Grammatičeskij slovar' russkogo jazyka*. Moskva: Russkie slovari.
- Zanettin F. (2000). *Parallel Corpora in Translation Studies: issues in corpus design and analysis*. In Olohan M. (ed.), *Intercultural Faultlines. Research Models in Translation Studies I. Textual and Cognitive Aspects*. Manchester: St. Jerome Publishing, pp 105-118.
- Zanettin F. (2002). *CEXI. Designing an English-Italian Translational Corpus*. Mair C., Meyer C.F., Oostdijk N. (eds.), *Language and Computers* 42 (1). Rodopi: New York. Pp. 329-343.
- Zanettin F. (2012). *Translation-driven corpora*. Routledge: New York.
- Zanettin F. (2013). *Corpus Methods for Descriptive Translation Studies*. In *Procedia - Social and Behavioral Sciences*, Vol. 95. Corpus Resources for Descriptive and Applied Studies. Current Challenges and Future Directions: Selected Papers from the 5th International Conference on Corpus Linguistics (CILC2013), pp. 20-32.
- Zasorina L. N. (1977) *Častotnyj slovar' russkogo jazyka*. Moskva: Russkij jazyk.
- Zemskaja E. A. Kitajgorodskaja M. V., Širjaev E. N. (1981). *Russkaja razgovornaja reč'*. *Obščie voprosy. Slovoobrazovanie. Sintaksis*. Moskva: Nauka.

CAPITOLO 2

IL CORPUS PARALLELO ITALIANO-RUSSO: COMPILAZIONE E RISULTATI

In questo breve capitolo descriveremo nel dettaglio le fasi dell'ampliamento del corpus parallelo e il risultato che abbiamo ottenuto. Sulla base del criterio di sistematicità discusso nel primo capitolo, documenteremo l'intero lavoro, creando così uno strumento di consultazione utile non solo per proseguire il progetto, ma anche per le ricerche future.

Il lavoro di ampliamento è stato svolto in tre diverse sedi: l'Istituto di Lingua Russa a Mosca, il Dipartimento di Slavistica dell'Università Cattolica di Milano e la Scuola Superiore di Lingue Moderne per Interpreti e Traduttori di Forlì (Università di Bologna). Il presente capitolo è dedicato principalmente al lavoro dell'unità milanese, ma descriveremo in breve anche l'operato dei restanti collaboratori i cui frutti sono confluiti nel corpus attualmente accessibile online all'indirizzo <http://www.ruscorpora.ru/search-para-it.html> (gennaio 2017).

I testi sono stati acquisiti ed allineati tra giugno 2014 e marzo 2015 e i risultati di questo primo ampliamento sono comparsi su *ruscorpora.ru* a dicembre 2015. A partire da questa data il corpus pilota ita-ru, composto da 700.000 occorrenze, è arrivato a contarne più di 4 milioni. Vediamo ora nel dettaglio le scelte che hanno guidato questa prima fase del lavoro.

2.1. Le dimensioni

Ad eccezione di alcune opere di dimensioni già di per sé ridotte (*Il Cappotto* di Nikolaj Gogol', *Il grande inquisitore* di Feder Dostoevskij e *Novecento* di Alessandro Baricco), l'unità milanese ha deciso di inserire nel corpus parallelo estratti che contassero tra le 10.000 e le 15.000 parole. La cifra è indicativa e talvolta questi parametri hanno subito lievi modifiche per due motivi: in primo luogo si è deciso di non interrompere l'estratto nel bel mezzo di un capitolo; di conseguenza, si è dato il caso in cui un capitolo contenesse meno di 10.000 parole e che con l'inserimento di due capitoli, al contrario, si superassero le 15.000. In secondo luogo, il volume degli estratti italiani è risultato sempre maggiore

rispetto alla versione russa, che si trattasse di originale o traduzione. Ciò è dovuto chiaramente al carattere più analitico della lingua italiana rispetto al russo. Pertanto, se un originale russo ammontava a 13-14 mila parole, poteva accadere che la traduzione italiana ne contenesse più di 15 mila. Viceversa, ove l'originale italiano contava 11-12 mila parole, la traduzione russa poteva aggirarsi attorno alle 9.000 parole. In ogni caso, il volume di ogni coppia di testi allineati conta un minimo di 20.000 e un massimo di 30.000 parole.

Quanto alla scelta degli estratti, si è deciso di partire dall'inizio di ogni opera. A seconda della lunghezza dei singoli capitoli, ogni estratto può contenerne 1-2 o addirittura 6-7. Sono rari i casi in cui il primo capitolo di un'opera sia già sufficientemente lungo da non richiedere l'aggiunta di ulteriori pagine. Nei casi in cui invece mancasse una vera e propria suddivisione in capitoli, l'estratto è stato interrotto in un punto dove potesse dirsi concluso un pensiero.

2.2. La scelta dei testi

Dopo aver stabilito i criteri generali per la compilazione del corpus, descritti nel primo capitolo (par. 1.10) (tipologia testuale e periodo di appartenenza), siamo passati alla scelta dei testi da inserire. Dal momento che il corpus parallelo pilota conteneva solo testi letterari, si è deciso di ampliare inizialmente questa sezione, così da renderla al più presto adatta alle ricerche scientifiche. Più precisamente, abbiamo iniziato a inserire opere russe tradotte in italiano (e non il contrario) poiché, prevedendo le ricerche che avremmo in seguito condotto (cfr. capitoli 3 e 5), abbiamo giudicato più urgente disporre di testi in questa direzione.

2.2.1. Prosa letteraria russa

Per quanto riguarda la ricezione in Italia della prosa letteraria russa, è stata rilevata un'enorme discrepanza fra i classici dell'Ottocento e le opere della seconda metà del XX secolo e del secolo XXI: per i primi esistono innumerevoli edizioni e traduzioni, mentre le opere più recenti sono generalmente pubblicate da una sola casa editrice e in un'unica versione (fanno eccezione poche opere, quali *Il maestro e Margherita*, che pur essendo più recente – 1966 – è stato tradotto nove volte). Pertanto, abbiamo suddiviso le opere russe tradotte in italiano in due gruppi:

A) opere dei secoli XIX e XX, tradotte più volte in italiano;

B) opere più recenti (XX e XXI secolo), tradotte una sola volta in italiano.

Ricordiamo che i dati relativi al numero di traduzioni sono stati raccolti grazie ai database delle biblioteche nazionali¹⁰⁴ e all'Index Translationum, già citato nel primo capitolo (cfr. par. 1.9.).

Per il gruppo A, inizialmente intendevamo costituire fin da subito un *target-variant corpus*, allineando per ogni opera il maggior numero possibile di versioni italiane, ma trattandosi di una mole di testi consistente, abbiamo cominciato ad acquisire e allineare innanzitutto le opere del gruppo B, così da conferire al corpus entro tempi brevi una certa eterogeneità, disponendo fin da subito di un buon numero di testi di vari autori e traduttori.

I titoli sono stati scelti sulla base di criteri diversi a seconda del gruppo. Per quanto riguarda le opere del gruppo B, fonte principale sono stati i cataloghi delle case editrici italiane. Nella maggior parte dei casi, per ogni autore è stata tradotta più di un'opera. Tuttavia si è deciso inizialmente di selezionarne una per autore. La scelta dell'opera è stata dettata in parte dalla sua importanza nel panorama letterario e dalla sua popolarità nel Paese d'arrivo.

Di seguito presentiamo una tabella con autori e titoli selezionati. Ricordiamo che per ogni opera sono stati inseriti solo alcuni estratti ad eccezione di un'opera acquisita e allineata in versione integrale dall'unità di Forlì (*La filiale* di Dvlatov)¹⁰⁵. Le opere sono disposte in ordine cronologico di pubblicazione.

Tab 4. Opere letterarie del gruppo B disponibili nel NKRJa

AUTORE	ANNO e TITOLO (originale e traduzione)	ANNO TRADUZIONE e TRADUTT.	CASA EDITRICE (originale e traduzione)	CAPITOLI ALLINEATI	N. PAROLE ru	N. PAROLE ita
I. Bunin	1923 <i>Mitina ljubov' / L'amore di Mitja</i>	1934 R. Kufferle	AST / Carabba editore	1-14	8627	11839

¹⁰⁴ <http://www.sbn.it/opacsbn/opac/iccu/free.jsp> (gennaio 2017).

¹⁰⁵ Come vedremo, la presenza di opere in versione integrale sbilancia parzialmente il progetto iniziale. Nelle conclusioni proporrò una soluzione per ovviare a questo problema.

V. Zazubrin	1923 <i>Šepka / La scheggia</i>	1990 S. Vitale	«Sibirskie ogni» / Adelphi	1-4	8188	11439
A. Platonov	1933-1936 <i>Ščastlivaja Moskva / Mosca felice</i>	1996 O. Discacciati S. Vitale	Eksmo / Adelphi	1-6	10712	13732
V. Nabokov	1937 <i>Dar / Il dono</i>	1998 S. Vitale	Slovo / Adelphi	Parte del cap. 1	11195	15077
Č. Ajtmatov	1966 <i>Proščaj, Gul'sary! / Addio Gul'sary</i>	1988 C. Di Paola S. Leone	Eksmo / Mursia	1-5	15565	21743
S. Sokolov	1973 <i>Škola dlja durakov / La scuola degli sciocchi</i>	2007 M. Crepax	Simposium / Salani	1	10449	12802
S. Dovlatov	1987 <i>Filial / La filiale</i>	2010 L. Salmon	Azbuka / Sellerio	Versione integrale	25298	31517
Ju. Druznikov	1988 <i>Angely na končike igly / Angeli sulla punta di uno spillo</i>	2006 F. Aceto	Kul'tura / Feltrinelli	1-4	11234	12854
V. Pelevin	1993 <i>Žizn' nasekomych / La vita degli insetti</i>	2000 V. Piccolo	Eksmo / Minimum fax	1-4	9312	12234
L. Ulickaja	1997 <i>Veselye pochorony / Funeral Party</i>	2004 E. Guercetti	Vagrius / Frassinelli	1-7	10055	12271
B. Akunin	1998 <i>Azazel' / La regina d'inverno</i>	2000 P. Pera	AST / Frassinelli	1-4	11028	14975
V. Erofeev	2004 <i>Chorošij Stalin / Il buon Stalin</i>	2008 L. Montagnani	Zebra / Einaudi	1- e parte del cap. 2	10672	13921
A. Kurkov	2005 <i>Zakon ulitki / I pinguini non vanno in vacanza</i>	2006 B. Osimo	Amfora / Garzanti	1-8	12146	15140

Naturalmente l'elenco presenta lacune, ma si tratta di una lista aperta che nel tempo è destinata ad essere integrata.

Per quanto riguarda il gruppo A, gli autori sono stati selezionati innanzitutto per la loro popolarità nel paese di arrivo, evidente dal gran numero di traduzioni italiane esistenti. Alcuni, pur non presentando una quantità di versioni italiane così elevata, sono stati comunque scelti per l'importanza che hanno rivestito nel panorama letterario internazionale.

Anche in questo caso è stato deciso di inserire inizialmente solo un'opera per ogni autore, scegliendo quella tradotta il maggior numero di volte.

Nella tabella 5 presentiamo i titoli del gruppo A inizialmente previsti per l'inserimento. In questo caso, i dati sono ordinati secondo il numero di traduzioni eseguite in italiano, per dare più visibilità a questo aspetto. Seguiranno alcune precisazioni circa quanto effettivamente disponibile al momento. Precisiamo che questo gruppo comprende anche *Il grande inquisitore*, che in base ai nostri parametri di classificazione è da considerarsi un caso particolare: il testo è tratto infatti dal capitolo quinto de *I fratelli Karamazov* di Dostoevskij (tradotto diverse volte in italiano), ma viene spesso considerato come un'opera nell'opera, e come opera a sé stante è stato tradotto da S. Vitale e nel 2010.

Tab 5. Proposta di inserimento opere letterarie del gruppo A

AUTORE	TITOLO (originale e traduzione) e ANNO	N. TRADUZIONI italiane
N. Gogol'	1842 <i>Šinel' / Il Cappotto</i>	32
F. Dostoevskij	1866 <i>Prestuplenie i nakazanie / Delitto e Castigo</i>	24
I. Turgenev	1862 <i>Otcy i deti / Padri e figli</i>	21
L. Tolstoj	1875-1877 <i>Anna Karenina / Anna Karenina</i>	16
M. Lermontov	1840 <i>Geroj našego vremeni / Un eroe del nostro tempo</i>	14

A. Čechov	1900 <i>Tri sestri / Le tre sorelle</i>	10
M. Bulgakov	1966-1967 <i>Master i Margarita / Il Maestro e Margherita</i>	9
V. Šalamov	1973 <i>Kolymskie rasskazy / I racconti di Kolyma</i>	3
V. Grossman	1980 <i>Žizn' i sud'ba / Vita e Destino</i>	2
B. Pasternak	1957 <i>Doktor Živago / Il dottor Živago</i>	2
F. Dostoevskij	1879 <i>Velikij inkvizitor / Il grande inquisitore</i>	1

Se questo è quanto ci si era prefissi, la considerazione delle ore di lavoro necessarie per l'acquisizione e l'allineamento di ogni testo ci ha portato alla conclusione che sarebbe stato impossibile, entro i tre anni del dottorato, inserire per ciascuna opera sopra elencata tutte le versioni italiane esistenti. A ciò si aggiunge il fatto che le versioni più antiche sono spesso ormai difficili da reperire.

In secondo luogo, durante la prima fase di allineamento, l'unità di Milano è stata messa in contatto con una studentessa universitaria di Mosca, K. D. Balakina, intenta ad allineare indipendentemente, per la sua tesi di laurea, alcune opere degli autori citati. Essendo a conoscenza del progetto del NKRJa, avrebbe poi ceduto i suoi file allineati ai programmatori dell'Istituto di Lingua Russa. Pertanto, si è deciso di sospendere temporaneamente l'acquisizione dei testi del gruppo A per evitare sovrapposizioni con il lavoro della tesista, la quale, per altro, operando secondo i propri criteri, ha imposto delle sostanziali modifiche ai piani originari in quanto:

1) ha selezionato i testi in base alla reperibilità su Internet, per disporre fin da subito del formato elettronico; ciò ha immancabilmente condotto ad una serie di conseguenze: per alcuni autori si dispone fin da subito di più opere (tradotte, peraltro, dallo stesso traduttore); mentre per i romanzi più famosi nella cultura di arrivo, tradotti numerose volte in italiano, non è stata scelta la versione italiana più nota o autorevole, ma quella disponibile online;

2) ha allineato testi interi anziché estratti;

3) l'allineamento è avvenuto in un'unica direzione (italiano-russo), vista la difficoltà di reperire sul web in modalità *open access* opere originali in lingua italiana.

Il bilanciamento auspicato nella fase di progettazione del corpus è dunque venuto parzialmente a mancare, sebbene il contributo della studentessa abbia indubbiamente consentito al corpus di raggiungere in minor tempo una dimensione accettabile ai fini della ricerca. Naturalmente, nel tempo sarà possibile tornare ad un maggiore bilanciamento inserendo un maggior numero di originali italiani.

Di seguito presentiamo l'elenco delle opere allineate da K. D. Balakina (tra cui una versione integrale de *I racconti di Kolyma* di Šalamov) e le relative informazioni disponibili¹⁰⁶.

Tab 6. Opere aggiunte da collaboratrice esterna

AUTORE	ANNO e TITOLO (originale / traduzione)	ANNO TRADUZIONE e TRADUTTORE	N. PAROLE (originale + traduzione)
N. Gogol'	1835-1852 <i>Mertvyje duši / Le anime morte</i>	2009 P. Nori	169132
I. Gončarov	1859 <i>Oblomov / Oblomov</i>	2000 A. Michettoni	339104
I. Turgenev	1862 <i>Otcy i deti / Padri e figli</i>	1908 F. Verdinois	120749
F. Dostoevskij	1866 <i>Prestuplenie i nakazanie / Delitto e castigo</i>	1983 G. Kraiski	385359
L. Tolstoj	1875-1877 <i>Anna Karenina / Anna Karenina</i>	1961 M. B. Luporini	597787
A. Čechov	1885 <i>Rasskazy / Racconti</i>	2004 F. Malcovati	229550
	1892 <i>Palata N. 6 / La corsia N. 6</i>	2004 F. Malcovati	36521

¹⁰⁶ I metadati sono stati tratti in parte dal sito del NKRJa sulla base di quanto compilato dalla collaboratrice stessa (titolo, autore, traduttore e numero di parole) e in parte (anno di traduzione) sono stati da noi integrati. Non disponiamo tuttavia delle informazioni sugli editori. Inoltre, il numero di parole reperibile tramite il corpus non fa distinzione tra originale e traduzione. Pertanto, non avendo allineato questi testi in prima persona, non siamo stati in grado di indicare il numero di parole dei testi presi singolarmente.

M. Gor'kij	1897-1898 <i>Byvšie ljudi / Gli ex-uomini</i>	1905 E. W. Foulques	43660
	1897-1898 <i>Ozornik / Il burlone</i>	1906 E. W. Foulques	12830
	1897-1898 <i>Suprugi Orlovy / I coniugi Orlof</i>	1905 E. W. Foulques	39740
	1897-1898 <i>Toska / L'angoscia</i>	1906 E. W. Foulques	28571
V. Nabokov	1967 ¹⁰⁷ <i>Lolita / Lolita</i>	1993 G. Arborio Mella	214434
V. Šalamov	1973 <i>Kolymskie rasskazy / I racconti di Kolyma</i>	1999 M. Binni	62535

Da parte nostra, le opere del gruppo A al momento inserite sono le seguenti: 15 versioni de *il Cappotto*¹⁰⁸ di Gogol', *I racconti di Kolyma* di Šalamov, per cui è stato possibile disporre fin da subito di due versioni, e *Il grande inquisitore* di Dostoevskij. Si puntualizza che *Il Cappotto* di Gogol' è l'opera russa più tradotta in italiano e quindi degna di una nota di riguardo. Inoltre, trattandosi di un racconto di 10.008 parole, è stato possibile l'allineamento per intero, a differenza di gran parte delle opere selezionate. Per questi motivi, si è deciso di inserire sin dall'inizio il maggior numero di versioni italiane possibili, così da poterle utilizzare per la nostra ricerca presentata nel capitolo 5. Si riportano i dati relativi a queste opere nella tabella 7.

¹⁰⁷ Si intende l'anno di pubblicazione della versione russa. Il romanzo fu redatto prima in inglese dall'autore e pubblicato nel 1955.

¹⁰⁸ Quando ci riferiamo a quest'opera usiamo il titolo maggiormente diffuso in Italia - *Il Cappotto* - sebbene alcuni traduttori abbiano proposto titoli diversi. Questi ultimi sono segnalati nella tab 7., accanto al nome del traduttore (per un approfondimento sulle versioni italiane di Šinel', cfr. par. 5.1).

Tab 7. Opere letterarie del gruppo A disponibili nel NKRJa

AUTORE	TITOLO ¹⁰⁹ (traduzione)	ANNO TRADUZIONE e TRADUTTORE	CASA EDITRICE ¹¹⁰ (originale/ traduzione)	CAPITOLI ALLINEATI	TOKEN ru	TOKEN ita
N. Gogol'	<i>Il cappotto</i>	1922 C. Rebora	Feltrinelli (1992) ¹¹¹	Versione integrale	10008	13253
	<i>Il cappotto</i>	1937 Duchessa D'Andria	Einaudi (1960)			11808
	<i>Il mantello</i>	1941 T. Landolfi	Rizzoli			12555
	<i>Il cappotto</i>	1949 O. Del Buono	Rizzoli			12777
	<i>Il cappotto</i>	1957 L. Pacini-Savoj	Newton Compton (2012)			11682
	<i>Il cappotto</i>	1963 ¹¹² G. Pacini	De Agostini			13113
	<i>Il cappotto</i>	1964 A. Julovic	Sansoni			12084
	<i>Il cappotto</i>	1967 P. Zveteremich	Garzanti			12187
	<i>Il cappotto</i>	1980 E. Bazzarelli	Rizzoli			12220
	<i>Il cappotto</i>	1986 S. Beffa	A. Peruzzo			11988
	<i>Il cappotto</i>	1986 F. Mariano	Mondadori			11887
	<i>Il cappotto</i>	1993 L. De Nardis	Newton Compton			12426
	<i>La mantella</i>	1994 S. Prina	I Meridiani Mondadori			12582

¹⁰⁹ Per l'anno di pubblicazione e il titolo originale cfr. tab 5.

¹¹⁰ La versione originale di *Šinel'* è quella contenuta in Gogol' N.V. (1937-1952) *Polnoe sobranie sočinenii v četyrnadcati tomach*. Leningrad: Izadel'stvo Akademii nauk SSSR, 1938, T. 3, *Povesti*, pp. 139-174.

¹¹¹ Indichiamo tra parentesi l'anno della versione di riferimento, se diverso da quello della prima edizione.

¹¹² Nel NKRJa è riportato un anno di traduzione non corretto (1949).

	<i>Il cappotto</i>	1995 E. Guercetti	Rizzoli			11768
	<i>Il cappotto</i>	2001 F. Legittimo	Marsilio			11069
V. Šalamov	<i>I racconti della Kolyma</i>	1999 M. Binni	Vagrius (1998) /Adelphi	I primi 6 racconti	9690	13133
	<i>I racconti di Kolyma</i>	2005 S. Rapetti	Vagrius (1998) Einaudi			14790
F. Dostoevskij	<i>Il grande inquisitore</i>	2010 S. Vitale	Pečatnyj dvor / Salani	Versione integrale	7744	8910

Riportiamo ora, in una tabella riassuntiva, le opere letterarie, dal russo all'italiano, disponibili nel corpus dal dicembre 2015, comprese le due opere già presenti in origine nel corpus pilota (*Le notti bianche* e *Il Maestro e Margherita*). I titoli sono disposti in ordine cronologico di pubblicazione.

Tab 8. Opere letterarie dal russo all'italiano nel NKRJa (dicembre 2015)

AUTORE	ANNO e TITOLO (originale / traduzione)
N. Gogol'	1842 <i>Šinel' / Il cappotto</i>
N. Gogol'	1835-1852 <i>Mertvye duši / Le anime morte</i>
F. Dostoevskij	1848 <i>Belye noči / Le notti bianche</i>
I. Gončarov	1859 <i>Oblomov / Oblomov</i>
I. Turgenev	1862 <i>Otcy i deti/ Padri e figli</i>

F. Dostoevskij	1866 <i>Prestuplenie i nakazanie / Delitto e castigo</i>
L. Tolstoj	1875-1877 <i>Anna Karenina / Anna Karenina</i>
F. Dostoevskij	1879 <i>Velikij inkvizitor / Il grande inquisitore</i>
A. Čechov	1885 <i>Rasskazy / Racconti</i>
	1892 <i>Palata N. 6 / La corsia N. 6</i>
M. Gor'kij	1897-1898 <i>Byvšie ljudi / Gli ex-uomini</i>
	1897-1898 <i>Ozornik / Il burlone</i>
	1897-1898 <i>Suprugi Orlovy / I coniugi Orlov</i>
	1897-1898 <i>Toska / L'angoscia</i>
I. Bunin	1923 <i>Mitina ljubov' / L'amore di Mitja</i>
V. Zazubrin	1923 <i>Šepka / La scheggia</i>
A. Platonov	1933-1936 <i>Ščastlivaja Moskva / Mosca felice</i>
V. Nabokov	1937 <i>Dar / Il dono</i>

M. Bulgakov	1929-1940 ¹¹³ <i>Margherita / Master i Margarita</i>
Č. Ajtmatov	1966 <i>Proščaj, Gul'sary! / Addio Gul'sary Il Maestro e</i>
V. Nabokov	1967 <i>Lolita / Lolita</i>
V. Šalamov	1973 <i>I racconti di Kolyma / Kolymskie rasskazy</i>
S. Sokolov	1973 <i>La scuola degli sciocchi / Škola dlja durakov</i>
S. Dovlatov	1987 <i>La filiale / Filial</i>
Ju. Druznikov	1988 <i>Angeli sulla punta di uno spillo / Angely na končike igly</i>
V. Pelevin	1993 <i>La vita degli insetti / Žizn nasekomych</i>
L. Ulickaja	1997 <i>Funeral Party / Veselye pochorony</i>
B. Akunin	1998 <i>La regina d'inverno / Azazel'</i>
V. Erofeev	2004 <i>Il buon Stalin / Chorošij Stalin</i>
A. Kurkov	2005 <i>I pinguini non vanno in vacanza / Zakon ulitki</i>

¹¹³ Ci si riferisce all'anno di stesura, così come compare sull'annotazione tesutale del NKRJa.

2.2.2. Prosa letteraria italiana

Per quanto riguarda le opere italiane tradotte in russo, non è stata necessaria la suddivisione in due gruppi, poiché gran parte delle opere inserite è stata pubblicata in Russia in un'unica versione (spesso esistono traduzioni in lingua russa tradotte e pubblicate in paesi ex-sovietici. In questa sede sono state considerate solo opere tradotte e pubblicate in territorio russo). Fanno eccezione *I promessi sposi* di Manzoni (6 versioni pubblicate in URSS)¹¹⁴ e due romanzi che, a differenza dell'Italia, dove sono meno noti, hanno goduto in Unione Sovietica di un'ampissima ricezione: *Lo Spartaco* di Raffaello Giovagnoli (8 versioni in URSS)¹¹⁵ e *Cipollino* di Gianni Rodari (di cui si attestano 3 versioni integrali, più alcuni rifacimenti per bambini)¹¹⁶. In due casi – *La Locandiera* di Goldoni e *Cuore* di De Amicis – le prime traduzioni dall'italiano al russo risalgono al periodo imperiale, prima della riforma dell'alfabeto, e non ci risultano ulteriori versioni in alfabeto russo moderno (ad eccezione di quelle inserite)¹¹⁷, pertanto non sono state prese in considerazione. Come per le opere dal russo all'italiano, anche in questo caso per dare fin da subito un corpus vario, i tre romanzi con più di una versione nella lingua d'arrivo sono stati inizialmente tralasciati. Nella tabella 9 presentiamo (in ordine cronologico) l'elenco dei testi da noi allineati e già disponibili nel NKRJa. A questi si aggiungono le tre opere già comprese nel corpus pilota (*Pinocchio* di C. Collodi, *Il nome della rosa* di U. Eco e la novella *La signora Frola e il signor Ponza suo genero* di L. Pirandello) e sette racconti di Dino Buzzati inseriti dall'unità di Forlì (I racconti sono tratti dalla raccolta *Sessanta Racconti*, 1958, Mondadori, e pubblicati in Russia da AST-Astrel' con il titolo *Šest'desjat rasskazov*). Non sono invece disponibili le informazioni bibliografiche per i testi che formavano il corpus pilota, prima dell'allineamento. Tutti questi testi sono stati inseriti in versione integrale. Per distinguere tra i testi da noi allineati e quelli inseriti da altre unità, riportiamo i dati relativi a questi ultimi in una tabella a parte (tab. 10).

¹¹⁴ Cinque pubblicate in Russia: V. Mežević (1854, Mosca), E. Nekrasova (1899, San Pietroburgo), I. Šitec (1936, Leningrado), A. G. Smirnov (1984, Mosca), Georgievskaja (2011, Mosca); e una traduzione del 1990 di A. Vajšnoras pubblicata a Vilnius.

¹¹⁵ Tre versioni pubblicate in Russia – da noi allineate – e le restanti pubblicate in varie parti dell'ex Unione Sovietica: B. Musaeu (1982, Baku), S. Tiraturjan (1982, Erevan), A. Gudajtis (1983, Vilnius), N. Kipiani (1984, Tbilisi), D. Faur e M. Rošu (1990, Chišinau)

¹¹⁶ Z. Potapova (1955, Mosca), I. Sokolova (2005, Smolensk), F. Dvin (2006, Mosca).

¹¹⁷ I dati relativi alle opere tradotte dall'italiano al russo e al numero di traduzioni esistenti sono stati raccolti grazie al database della biblioteca nazionale "Lenin" di Mosca, oltre che all'index Translationum.

Nella tabella 11 compariranno invece ulteriori testi selezionati, ma non ancora caricati sul corpus. Una parte delle opere selezionate, infatti, non è ancora stata allineata, oppure l'allineamento è avvenuto dopo dicembre 2015. Esse saranno pertanto visibili su *ruscorpora* nel 2017.

Tab 9. Opere letterarie dal russo all'italiano nel NKRJa (dicembre 2015)

AUTORE	ANNO e TITOLO (originale / traduzione)	ANNO TRADUZIONE e TRADUTTORE	CASA EDITRICE (originale e traduzione)	CAPITOLI ALLINEATI	N. PAROLE ita	N. PAROLE ru
C. Goldoni	1753 <i>La locandiera</i> / <i>Traktirshchica</i>	1933 A. Dživelegov	Marsilio (2007) / Iskusstvo	I atto, scena 1-22	7096	10068
E. De Amicis	1886 <i>Cuore / Serdce</i>	1958 V. Davidenkova	Feltrinelli (2008) / Lenizdat	cap 1-2 (ottobre e novembre)	12295	12105
L. Pirandello	1904 <i>Il fu Mattia Pascal / Pokojnyj Mattia Paskal'</i>	1967 G. Rubcova, N. Rykova	Garzanti (1993) / Chudož. lit. Leningr. otd-nie ¹¹⁸	cap 1-5	15017	13253
I. Calvino	1957 <i>Il Barone rampante / Baron na dereve</i>	1965 L. Veršinin	Einaudi / chudož. lit. - Moskva	cap 1-5	14599	12633
G. Bassani	1962 <i>Il Giardino dei Finzi-Contini / Sad Finci- Kontini</i>	2008 I. Soboleva	Einaudi / Tekst	I parte: cap 1-6	11224	10050
U. Eco	1988 <i>Il pendolo di Focault / Majatnik Fuko</i>	2006 E. Kostjukovič	Bompiani / Symposium	cap 1-6	12163	11003
A. Baricco	1994 <i>Novecento / Legenda o pianiste</i>	2005 N. Kolesova	Feltrinelli / Symposium	Versione integrale	11700	10457
A. Camilleri	1996 <i>Il cane di terracotta / Sobaka iz terrakotty</i>	2005 A. Kondjurina	Sellerio / Inostranka	cap 1-5	12360	11042

¹¹⁸ Chudožestvennaja literatura – Leningradskoe otdelenie.

V.M. Manfredi	1998 <i>Aléxandros, III, Il confine del mondo / Aleksandr makedonskij. Predely mira</i>	2003 M. Kononov	Mondadori / EKSMO	cap 1-7	12867	10877
N. Ammaniti	2001 <i>Io non ho paura / Ja ne bojus'</i>	2005 V. Nikolaev	Einaudi / Machaon	cap 1-2	13573	11422
G. Faletti	2002 <i>Io uccido / Ja ubivau</i>	2005 I. Kostantinova	Baldini / Castoldi / Azbuka Klassika	cap 1-4	13248	10000

Tab 10. Altre opere letterarie dall'italiano al russo disponibili sul NKRJa (dicembre 2015)

AUTORE	ANNO E TITOLO (originale / traduzione)	ANNO TRADUZIONE e TRADUTTORE	N. PAROLE (ru+ita) ¹¹⁹
C. Collodi	1883 <i>Pinocchio / Priključenija Pinokkio</i>	1959 E. Kazakevič	73576
L. Pirandello	1917 <i>La signora Frola e il signor Ponza, suo genero / Gospoža Frola i ee zjat' gospodin Ponca</i>	2006 M. Gasanova-Mijat	5999
D. Buzzati	1942 <i>Eppure battono alla porta / I vse že stučat v dver'</i>	2010 I. Smagin	7466
	1942 <i>I sette messaggeri / Sem' goncov</i>	2010 F. Dvin	2807
	1942 <i>Il borghese stregato / Očarovannyj buržua</i>	2010 P. Chlodovskij	4312
	1942 <i>Il mantello / Plašč</i>	2010 G. Kiselev	2857

¹¹⁹ Anche in questo caso abbiamo potuto ricavare solo il numero di parole complessivo.

	1942 <i>L'assalto del Grande Convoglio / Napadenie na bol'shoj konvoj</i>	2010 P. Chlodovskij	8846
	1942 <i>L'uccisione del drago / Kak ubili drakona</i>	2010 F. Dvin	9018
	1942 <i>Sette piani / Sem' etažej</i>	2010 G. Kiselev	9382
U. Eco	1980 <i>Il nome della rosa / Imja rozy</i>	1989 E. Kostjukovič	341954

Tab. 11. Opere letterarie dall'italiano al russo non ancora disponibili nel NKRJa ma previste per il 2017

AUTORE	ANNO e TITOLO (originale / traduzione)	ANNO TRADUZIONE e TRADUTTORE	CAPITOLI ALLINEATI	CASA EDITRICE (originale e traduzione)	N. PAROLE ita	N. PAROLE ru
A. Manzoni	1827 <i>I promessi sposi / Obručennye</i>	2011 N. Georgievskaja	cap 1-3	Gruppo editoriale Fabbri (1980) / Azbuka	16054	14767
R. Giovagnoli	1873-4 <i>Lo Spartaco/ Spartak</i>	1881 S. Stepnjak-Kravčinskij	cap 1-2	Parenti Editore (1955) / Entropoligraf (2014)		9253
		1936 L. Frenkel'		Žurgazob'edinenija		8225
		1954 A. Jasnaja		Kniga		11576
E. Salgari	1898 <i>Il corsaro nero / černyj korsar</i>	1991 G. Smirnov	cap 1-5	Fabbri (1977) / SP "Inart"	13689	11038
I. Svevo	1923 <i>La coscienza di Zeno / Samopoznanie Dzeno</i>	1972 S. Buševa	cap 1-4	Mondadori (2002) / Chudož. lit. Leningr. otdnie	17879	18055

A. Moravia	1929 <i>Gli indifferenti / Ravnodušnye</i>	2010 L. Veršinin	cap 1-5	Bompiani (1992) / AST: Astrel'	16897	13815
G. Guareschi	1948 <i>Don Camillo: Mondo piccolo / Malyj mir</i>	2014 O. Gurevič	cap 1-9	Rizzoli (1974) / Granat	11519	10238
G. Rodari	1951 <i>Le avventure di Cipollino / Priključenija Čipollino</i>	1955 Z. Potapova	cap 1-8	Editori Riuniti / EKSMO-press (2001)	10187	10050
N. Ginzburg	1963 <i>Lessico familiare / Semejnye besedy</i>	1989 G. Smirnov	manca una suddivisione in capitoli	Einaudi / Raduga	11327	9554
L. Sciascia	1976 <i>I pugnatori / Palermskie ubijcy</i>	1982 E. Solonovič, Z. Potapova	manca una suddivisione in capitoli	Adelphi (2003) / Izvestija	11533	10491
B. Pitzorno	1984 <i>La casa sull'albero / Dom na dereve</i>	2012 T. Stamovaja	cap 1-10	Mondadori / Machaon	12102	10843
G. D'Agata	1997 <i>Il ritorno dei templari / Vozvrascenie tamplierov</i>	2005 I. Kostantinova	cap 1-6	Newton Compton / Azbuka-klassika	14824	13223

2.2.3. Saggistica

Per quanto concerne la saggistica, trovare testi tradotti in entrambe le direzioni è senz'altro più difficile. Per avere un quadro generale su quanto effettivamente tradotto dal russo all'italiano, e viceversa, ci siamo basati sull'Index Translationum. Nonostante la sua incompletezza - tralascia alcune opere e attesterebbe solo i libri tradotti entro il 2012, ma per il russo e l'italiano, in realtà copre solo il periodo fino al 2009 -, questo strumento si è rivelato particolarmente utile proprio in quest'ambito, poiché suddivide le opere tradotte in base al tema trattato. Ciò ha permesso di diversificare il nostro repertorio

fin da subito, selezionando, per iniziare, quattro testi per ognuna delle categorie tematiche proposte dall'Index (due dal russo all'italiano, due dall'italiano al russo), che sono:

1. Arte / Giochi / Sport; 2. Critica letteraria / Letteratura¹²⁰; 3. Filosofia / Psicologia; 4. Religione / Teologia; 5. Legge / Scienze sociali / Formazione; 6. Scienze esatte e naturali; 7. Scienze applicate; 8. Storia / Biografia.

La classificazione proposta dall'Index è stata dunque utile per proporre testi non letterari appartenenti a diverse categorie tematiche. Tuttavia non è impeccabile, soprattutto per alcune categorie, come le Scienze applicate, comprendenti ad esempio anche libri di cucina. Pertanto, nella tabella 12, per alcune delle opere selezionate abbiamo nominato i campi in modo diverso (cfr. “epistolario” per Florenskij e “Memorie” per Stanislavskij e Čukovskaja).

Di seguito presentiamo in due tabelle le opere di saggistica selezionate, ordinate per categoria (le categorie seguono l'ordine alfabetico).

Tab 12. Opere di saggistica russe tradotte in italiano

CATEGORIA	AUTORE	ANNO e TITOLO (originale e traduzione)	ANNO TRADUZIONE e TRADUTTORE	CASA EDITRICE (originale e traduzione)
Arte	E. Trubeckoj	1916 <i>Umozrenie v kraskach. Etjudi po russkoj ikonopisi / Contemplazione nel colore. Tre studi sull'icona russa</i>	1989 P. Cazzola	Izd-vo Moskovskoj patriapchii Russkoj pravoslavnoj cerkvi (2012) / La Casa di Matriona
Critica letteraria	M. Bachtin	1929 <i>Problemy poetiki Dostoevskogo / Dostoevskij: Poetica e stilistica</i>	1968 G. Garritano	Priboj / Einaudi
Epistolario	P. Florenskij	1933-37 (pubbl. 1998) <i>Pis'ma s Dal'nego Vostoka i Solovkov / Non dimenticatemi: dal gulag staliniano le lettere alla moglie e</i>	2001 G. Guaita, L. Charitonov	Mysl' / Mondadori

¹²⁰ L'index etichetta questo gruppo con il nome “Letteratura” comprendendo anche la fiction. Per distinguere opere di saggistica sulla letteratura dalle opere letterarie, abbiamo aggiunto la denominazione “critica letteraria”.

		<i>ai figli del grande matematico, filosofo e sacerdote russo</i>		
Filosofia	V. Solovev	1874 <i>Krizis zapadnoj filosofii / La crisi della filosofia occidentale</i>	1980 A. Dell'Asta	Egloss Edišniz (2012) / La casa di Matriona
Legge/scienze sociali	R. Medvedev	1972 <i>Kniga o socialističeskoj demokratii / La democrazia socialista</i>	1977 S. Rapetti	Fond im. Gercena / Vallecchi Editore
	A. Solženicyn	<i>L'errore dell'Occidente: gli ultimi interventi su comunismo, Russia e Occidente</i> ¹²¹	1980 I. Ilovajskaja Alberti, S. Rapetti	YMCA – Press / La casa di Matriona
Memorie	K. Stanislavskij	1924 <i>Moja žizn v iskusstve / La mia vita nell'arte</i>	1963 M. Borsellino di Lorenzo	Iskusstvo (1962) / Einaudi
	L. Čukovskaja	1976 <i>Zapiski ob Annoj Achmatovoj / Incontri con Anna Achmatova, 1938-1941</i>	1990 G. Moracci	YMCA – Press / Adelphi
Religione/teologia	N. Berdjaev	1927 <i>Filosofija svobodnogo ducha / Filosofia dello spirito libero: problematica e apologia del cristianesimo</i>	1997 G. L. Giaone	YMCA – Press / Edizioni San Paolo
	S. Bulgakov	1922 <i>U sten Chersonisa / Presso le mura di Chersoneso</i>	1998 M. Campatelli	AO “Dorval”: AO “Liga” /

¹²¹ L'opera è un'antologia che raccoglie tre saggi di A. Solženicyn: “*Kommunizm: u vseh na vidu – i ne ponjat*” [Il comunismo: quando si smetterà di chiudere gli occhi di fronte all'evidenza?] (in *Sobranie sočinenij*, YMCA - PRESS, Vermont - Pariž, t. IX, 1981, pp. 298-304.); “*Čem grozit Amerike plochoe ponimanie Rossii*” [I pericoli che incombono sull'Occidente a causa della sua ignoranza della Russia] (in *Sobranie sočinenij*, YMCA - PRESS, Vermont - Pariž, t. IX, 1981, pp. 305-344); “*Reč' v Garvarde*” [Un mondo in frantumi, «Discorso di Harvard»] (in *Sobranie sočinenij*, YMCA - PRESS, Vermont - Pariž, t. IX, 1981, pp. 280-297).

				La casa di Matrona
Scienze applicate	A. Lurija	1975 <i>Osnovnye problemy nejrolingvistiki / Problemi fondamentali di neurolinguistica</i>	1978 G. Cossu, A. Miazzi	Librokom (2009) / Armando Editore
	I. Pavlov	1973 <i>Dvadcatiletnij opyt ob'ektivnogo izučenija vysšej nervnoj dejatel'nosti (povedenija) životnyh / I riflessi condizionati</i>	2011 M. Silvestri-Lapenna	Biomedgiz / Bollati Boringhieri
Scienze esatte e naturali	V. Arnol'd	1983 <i>Teorija katastrof / Teoria delle catastrofi</i>	1990 F. Aicardi	Nauka (1990) / Bollati Boringhieri
	L. Laundau, E. Lifšic	1938 <i>Statističeskaja fizika / Fisica statistica</i>	1978 A. Machov	ONTI / Editori Riuniti
Storia/biografia	A. Politkovskaja	2004 <i>Putinskaja Rossija / La Russia di Putin</i>	2010 C. Zonghetti	Novaja gazeta / Adelphi
	A. Dostoevskaja	1971 (prima pubblicazione quasi completa). Versione di riferimento: 1987 <i>Vospominania A. G. Dostoevskoj / Dostoevskij marito</i>	1939 Anna Milazzo Lipschutz	Azbuka (2011) / Bompiani

Tab 13. Opere di saggistica italiane tradotte in russo

CATEGORIA	AUTORE	ANNO e TITOLO (originale e traduzione)	ANNO TRADUZIONE e TRADUTTORE	CASA EDITRICE (originale e traduzione)
Arte	M. Tibaldi Chiesa	1944 <i>Paganini: la vita e l'opera / Paganini</i>	1981 I. Kostantinova	Garzanti / Mol. gvardija
	P. Grassi	1997 <i>Quarant'anni di palcoscenico / Moj teatr</i>	1982 Ju. Dobrovol'skaja	Mursia / Iskusstvo
Critica letteraria	F. De Sanctis	1954 <i>Storia della letteratura italiana/Istorija italjanskoj literatury v dvuch tomach</i>	1963 Ju. Dobrovol'skaja, N. Elinaja	Laterza / Izd-vo inostr. lit.
	E. Lo Gatto	1976 <i>I miei incontri con la Russia / Moi vstreči s Rossiej</i>	1992 K. Gladyš, I. Dergačevaja	Mursia / Krug'
Filosofia/psicologia	A. Meneghetti	2000 <i>La psicologia del leader/Psihologija lidera</i>	1996 A. Nikolaev	Psicologica editrice /
	U. Eco	1968 <i>La struttura assente / Otsutstvjuščaja struktura</i>	2006 V. Reznik, A. Pogonjajlo	Bompiani / Ontopsichologija
Legge/Scienze sociali/Formazione	A. Gramsci	1948 <i>I quaderni del carcere / Tjuremnye tetradi</i>	1991 G. Smirnov	Carocci / Politizdat
	M. Montessori	1909 <i>Il metodo della pedagogia scientifica applicato all'educazione infantile nelle case dei bambini / Dom rebenka</i>	2005 S. Zajmovskij	Editrice Lapi / AST: Astrel'
Religione/Teologia	A. Sicari	1993 <i>Breve catechesi sul matrimonio/ o brake</i>	1993 L. Charitonova	Jaca book / Christian. Rossija

	A. Donini	1975 <i>Storia del cristianesimo / U istokov hristianstva</i>	1989 I. Kravčenko,	Teti / Politizdat
Scienze applicate	G. Cella al-Chamali	2002 <i>Riflessologia / Refleksologija</i>	2008 ¹²²	Fabbri / Kontent
	E. Bettelli	<i>Il libro d'oro della casa / Zolotaja enciklopedija hozjajki</i>	2005 N. Belokon	De Agostini / AST-Press kn.
Scienze esatte e naturali	T. Regge	1981 <i>Cronache dell'universo / Etjudy o Vselennoj</i>	1985 D. Pontekorvo	Boringhieri / Mir
	G. Careri	1982 <i>Ordine e disordine nella materia / Porjadok i besporjadok v strukture materii</i>	1985 B. Kerbikov	Laterza / Mir
Storia/biografia	F. Cardini	1982 <i>Alle radici della cavalleria medievale / Istoki srednevekovogo rycarstva</i>	1987 V. Gajduk	La Nuova Italia / Progress
	G. Boffa	1995 <i>Dall'URSS alla Russia / Ot SSSR k Rossii</i>	1996 L. Haustova	Laterza / Meždunarod. otnošenija

La maggior parte delle opere nelle due tabelle soprastanti sono state acquisite e digitalizzate, ma non ancora allineate. L'allineamento sarà completato nel 2017. Al momento il CP ru-ita contiene 5 delle opere di saggistica russe elencate nella tabella 12, di cui presentiamo qui di seguito un elenco con alcuni metadati aggiuntivi.

¹²² Le informazioni sul traduttore al momento non sono reperibili.

Tab 14. Opere di saggistica disponibili sul NKRJa (dicembre 2015)

AUTORE	ANNO e TITOLO (originale e traduzione)	CAPITOLI ALLINEATI	ANNO TRADUZIONE e TRADUTTORE	N. PAROLE ru	N. PAROLE ita
Bachtin M.	1929 <i>Problemy poetiki Dostoevskogo / Dostoevskij: Poetica e stilistica</i>	cap 1	1968 G. Garritano	12375	15724
Cukovskaja L.	1976 <i>Zapiski ob Annoj Achmatovoj / Incontri con Anna Achmatova, 1938-1941</i>	manca suddivisione in capitoli	1990 G. Moracci	11130	14828
Solovev V.	1874 <i>Krizis zapadnoj filosofii / La crisi della filosofia occidentale</i>	cap 1	1980 A. Dell'Asta	9465	12810
Bulgakov S.	1922 <i>U sten Chersonisa / Presso le mura di Chersoneso</i>	cap 1-2	1998 M. Campatelli	13010	16473
Politkovskaja A.	2004 <i>Putinskaja Rossija / La Russia di Putin</i>	manca suddivisione in capitoli	2010 C. Zonghetti	12605	14679

2.3. Le fasi di costruzione del corpus

Concludiamo la nostra descrizione, con una presentazione più dettagliata delle fasi di costruzione del corpus. Daremo così spiegazione di eventuali errori o mancanze che gli utenti potrebbero riscontrare durante l'utilizzo dello strumento.

Le fasi che seguono la scelta dei titoli e il reperimento dei testi online o nelle biblioteche sono tre:

- 1) La conversione in file di testo txt., che per i testi a stampa significa anche la previa digitalizzazione.
- 2) L'allineamento tramite un software reso disponibile dai programmatori di Mosca.
- 3) La compilazione dei metadati.

Nonostante l'ausilio di macchinari e appositi programmi, sia la conversione dei testi in formato digitale, sia l'allineamento richiedono un successivo intervento di

correzione e controllo manuale, ed è per questo motivo che l'eventualità di incontrare degli errori durante l'utilizzo del corpus non è da escludere. Vediamo nel dettaglio i due errori più frequenti:

- 1) Refusi: parole scritte non correttamente, errori di punteggiatura e spaziatura.
- 2) Incongruenze nell'allineamento, per cui una frase tradotta non corrisponde all'originale soprastante.

Il primo tipo di errore si verifica generalmente nei testi acquisiti dalla versione cartacea. Il processo di digitalizzazione di opere cartacee consiste in due fasi: a) scansione delle pagine; b) conversione dell'immagine scansionata in un file di testo modificabile. Questo secondo passo è realizzabile automaticamente mediante uno scanner con opzioni avanzate, oppure attraverso programmi disponibili gratuitamente sul web che permettono di convertire un pdf o un'immagine in file Word. In entrambi i casi i risultati ottenuti esigono però un controllo per verificare che ogni singola parola sia stata correttamente convertita. Un segno di qualsiasi tipo presente sulla pagina scansionata (un appunto, una sottolineatura, un granello di polvere, così come lo scorretto posizionamento della pagina durante la scansione) può compromettere la corretta elaborazione del documento. Inoltre, vanno eliminati i numeri di pagina, eventuali intestazioni, trattini di sillabazione a fine riga e così via. A titolo di esempio, mostriamo qui di seguito un esempio di pagina appena convertita da scansione in pdf a documento in Word, dove le parti sottolineate sono da eliminare o correggere:

"I pugnatori 285 di Palermo ha dichiarato risultargli bastano oggi duecen-tocinquantamila lire a pagare il servizio di fare uccidere un uomo: che corrispondono ai tre tarì di allora, tenuto conto dell'attuale leggerezza della moneta e della legge-rezza con cui la si spende. Ci furono altre riunioni: ma non si concludeva nulla, tanto che qualcuno si sentiva rimordere la coscienza, per quei tre tarì al giorno che prendeva senza "soddisfare" co-lui che li pagava. Finalmente, la sera del 1° ottobre, Ca-stelli disse: "questa sera cí sarà tonnina", cioè un massacro come quello che si faceva nelle tonnare, nei giorni della passa dei tonni. Ad ora di avemaria, agli ordini di Castelli, si trovarono presso il palazzo delle Finanze D'Angelo e Termini (dove andassero gli altri due gruppi D'Angelo non seppe). Ri-masero in quella zona per tre volte eseguendo gli ordini di

Castelli. E il primo da ferire se lo giocarono, D'Angelo e Termini, a pari e dispari: e toccò al Termini. Il secondo spettava al D'Angelo: e più vilmente dell'altro, avvicinan-dosi alla vittima con la richiesta della presa di tabacco, eseguì il comando. Il terzo, che sarebbe spettato al Termini, il Castelli — forse per meglio educarlo — lo assegnò al D'Angelo. Creduto pienamente sui nomi degli undici esecutori e sul racconto dei fatti, D'Angelo non fu creduto per niente sul nome del mandante. Si credette cioè che quel nome fosse stato sì detto dal Castelli, in accordo con Masotto e Calì: ma tanto per farne uno che garantisse il soldo e facesse da schermo al mandante vero. Naturalmente, Castelli negò; negò tutto, negò sempre. E così gli altri. Sí ritenne dunque che il principe di Sant'Elia fosse la quattordicesima vittima: non dí coltello, ma dí calunnia. E questo fino al momento del processo ai dodici pugnalatori, e anche da parte del procuratore Giacosa che sostenne l'accusa. Ma nella violenza con cui, nell'arringa, respinge il sospetto che il principe di Sant`.1 - .lia possa avere avuto mano in quei delitti, è da intra-vedere la volontà di liberarsi appunto dí quel sospetto: che gli si insinuava, che lo inquietava.”

Per questo tipo di lavoro è necessario un buon compromesso tra efficienza e velocità: è impossibile rileggere i testi parola per parola, ma allo stesso tempo bisogna ridurre il più possibile il margine di errore. La segnalazione automatica degli errori compiuta da Word è di grande aiuto per identificare le parole scorrette. In caso di frasi interamente mancanti (succede anche questo talvolta) il testo cartaceo a fronte consente di colmare le lacune. I problemi maggiori sono dati da traduzioni italiane di testi russi contenenti molti segni diacritici. In questo caso l'intervento manuale deve essere più rigoroso, ma soprattutto richiede più tempo.

Il secondo tipo di errore, fortunatamente più raro, poiché di gran lunga più facile da cogliere prima della pubblicazione, riguarda la fase di allineamento. Anche in questo caso, a titolo di esempio, presentiamo un'immagine che esemplifica l'interfaccia del programma di allineamento utilizzato. Una volta che i testi (originale e traduzione) sono stati caricati nel programma, quest'ultimo propone un allineamento che, come si può vedere, in alcuni casi risulta scorretto, poiché frase originale e frase tradotta non coincidono (qui, nello specifico, alla frase italiana numero 2, il russo aggiunge un

ulteriore frammento che corrisponderebbe all'italiano numero 3, causando così un errore di allineamento non solo per la frase 2, ma anche per quelle successive). Il nostro lavoro è stato correggere manualmente questi errori, sempre cercando di ottimizzare il compromesso tra precisione e velocità. Nell'esempio sottostante sono inoltre visibili i problemi di punteggiatura menzionati nel primo capitolo, per cui la suddivisione in frasi tramite punto fermo non sempre coincide tra originale e traduzione. Qui nell'esempio 1, dove l'italiano propone i due punti, il russo opta per un punto fermo.

Fig 3. Esempio di testi allineati automaticamente dal programma

0	Schiaccia un piede Cipollone al gran principe Limone	0,264286	В которой Чиполлоне отдал ногу принцу Лимону
1	Cipollino era figlio di Cipollone e aveva sette fratelli: Cipolletto, Cipolotto, Cipolluccio e così di seguito, tutti nomi adatti ad una famiglia di cipolle.	-0,0132353	Чиполлино был сыном Чиполлоне. ~~~ И было у него семь братьев: Чиполлетто, Чиполлотто, Чиполлочча, Чиполлучча и так далее – самые подходящие имена для честной луковой семьи.
2	Gente per bene, bisogna dirlo subito, però piuttosto sfortunata.	0,278947	Люди они были хорошие, надо прямо сказать, да только не везло им в жизни. ~~~ Что ж поделаешь: где лук, там и слезы.
3	Cosa volete, quando si nasce cipolle, le lacrime sono di casa.	0,118868	Чиполлоне, его жена и сыновья жили в деревянной лачуге чуть побольше ящичка для огородной рассады.
4	Cipollone ed i suoi figli abitavano in una baracca di legno, poco più grande di una cassetta di quelle che si vedono dall'ortolano.	0,264545	Если богачам случалось попадать в эти места, они недовольно морщили носы, ворчали: «Фу, как несёт луком!» – и приказывали кучеру ехать быстрее.

La terza fase, ossia l'inserimento dei metadati, avviene al termine dell'allineamento grazie ad un'apposita finestra presente nel programma di allineamento stesso. In questo caso, un'attenta compilazione e rilettura dovrebbe garantire l'assoluta precisione.

Il portale del NKRJa offre uno spazio per segnalare qualsiasi errore riscontrato dagli utenti. Come già accennato nel primo capitolo, durante l'utilizzo del corpus, cliccando su ogni parola nella pagina delle occorrenze compare una finestra con le informazioni grammaticali e semantiche sulla parola in questione. In questa stessa finestra è possibile cliccare sulla voce “*soobščit' ob ošibke*” (*comunicare un errore*) nel caso si riscontrasse un errore nella parola o nell'annotazione. Cliccando poi sul titolo dell'opera o sul nome dell'autore, comparirà una finestra con i metadati. Anche in quest'ultima vi è la voce “*soobščit' ob ošibke*”, da utilizzare in caso di errori riscontrati nelle informazioni extralinguistiche inserite.

CAPITOLO 3

I VERBI DI MEMORIA IN RUSSO E IN ITALIANO: INDAGINE CONTRASTIVA *CORPUS-BASED*

Il presente capitolo è dedicato all'analisi *corpus-based*, in chiave contrastiva, dei verbi russi cosiddetti di memoria¹²³, che si differenziano semanticamente fra di loro mediante prefissazione.

La prefissazione verbale del russo è un fenomeno molto esteso e, di conseguenza, altrettanto studiato. Pertanto, nella prima parte di questo capitolo (par. 3.1.) presenteremo brevemente i problemi centrali legati a questo tema così come li hanno evidenziati alcuni importanti lavori della russistica degli ultimi vent'anni. Per ogni argomento citato forniremo in nota una bibliografia fondamentale di riferimento. Rimandiamo invece a Krongauz (1998: 40-49) per una rassegna esaustiva riguardante gli studi relativi alla prefissazione verbale nella lingua russa, anche quelli più datati. Al termine di questo primo paragrafo esporremo quindi il quadro teorico a cui faremo riferimento per alcune distinzioni tipologiche tra il russo e l'italiano. I par. 3.2. e 3.3. introdurranno la nostra ricerca, esponendo le motivazioni a favore di un'analisi dei prefissi verbali *corpus-based* e descrivendo in modo più preciso i verbi di memoria russi formati tramite prefissazione. Infine (par. 3.4.), mediante alcuni esempi tratti dal corpus parallelo ru-ita, presenteremo le strategie di resa di questi verbi in italiano, mettendo in evidenza i casi di perdita di informazione riscontrati. Seguiranno alcune osservazioni a conclusione dell'analisi proposta (par. 3.5.).

Il sotto-corpus utilizzato per la nostra analisi, ora presente anche nel NKRJa, è composto da 30 testi letterari russi con la rispettiva traduzione italiana (per due opere si dispone di più traduzioni) e conta 3.144.533 parole. Come precisato nel secondo capitolo, in alcuni casi si tratta di estratti, in altri di opere intere. Di seguito presentiamo la lista degli autori e dei testi compresi nel corpus, affiancati dall'indicazione del titolo italiano, del traduttore e dell'anno di traduzione. I testi sono elencati nell'ordine cronologico decrescente proposto dal NKRJa:

¹²³ Non esiste una classificazione univoca dei verbi in base alla classe semantica, pertanto ci avvaliamo qui della denominazione adottata dai linguisti che si sono occupati di questo gruppo in russistica (in russo “glagoly pamjati”). Cfr. Rebrina (2008) e Zaliznjak (2013).

1. A. Kurkov. *Zakon ulitki* (2005) – *I pinguini non vanno in vacanza*. B. Osimo (2006).
2. V. Erofeev. *Chorošij Stalin* (2004) – *Il buon Stalin*. L. Montagnani (2008).
3. B. Akunin. *Azazel'* (1998) – *La regina d'inverno*. P. Pera (2000).
4. L. Ulickaja. *Veselye pochorony* (1997) – *Funeral party*. E. Guercetti (2004).
5. V. Pelevin. *Žizn' nasekomych* (1993) – *La vita degli insetti*. V. Piccolo (2000).
6. Ju. Družnikov. *Angely na končike igly* (1988) – *Angeli sulla punta di uno spillo*. F. Aceto (2006).
7. S. Dovlatov. *Filial* (1987) – *La filiale*. L. Salmon (2010).
8. S. Sokolov. *Škola dlja durakov* (1973) – *La scuola degli sciocchi*. M. Crepax (2007).
9. V. Nabokov. *Lolita*. (1955-1967) – *Lolita*. G. Arborio Mella (1993).
10. Č. Ajmatov. *Proščaj, Gul'sary!* (1966) – *Addio Gul'sary!* C. Di Paola, S. Leone (1988).
11. V. Šalamov. *Kolymskie rasskazy* (1954-1962) – *I racconti della Kolyma*. M. Binni (1999) / *I racconti di Kolyma*. S. Rapetti (2005).
12. M. Bulgakov. *Master i Margarita* (1929-1940) – *Il maestro e Margherita*. V. Drisdo (1967).
13. V. Nabokov. *Dar*. (1937) – *Il dono*. S. Vitale (1988).
14. A. Platonov. *Sčastlivaja Moskva* (1933-1936) – *Mosca Felice*. O. Discacciati, S. Vitale (1996).
15. I. Bunin. *Mitina ljubov'* (1923) – *L'amore di Mitja*. R. Kufferle (1934).
16. V. Zazubrin. *Ščepka* (1923) – *La scheggia*. S. Vitale (1990).
17. A. Čechov. *Rasskazy* (1885-1903) – *Racconti*. F. Malcovati (2004).
18. M. Gor'kij. *Byvšie ljudi* (1897-1898) – *Gli ex-uomini*. E. W. Foulques (1905).
19. M. Gor'kij. *Ozornik* (1897-1898) – *Il burlone*. E. W. Foulques (1905).
20. M. Gor'kij. *Suprugi Orlovy* (1897-1898) – *I coniugi Orlof*. E. W. Foulques (1905).
21. M. Gor'kij. *Toska* (1897-1898) – *L'angoscia*. E. W. Foulques (1905).
22. A. Čechov. *Palata № 6* (1892) – *La corsia n. 6*. F. Malcovati (2004).
23. F. Dostoevskij. *Velikij inkvizitor* (1879) – *Il grande inquisitore*. S. Vitale (2010).
24. L. Tolstoj. *Anna Karenina* (1873-1877) – *Anna Karenina*. M. B. Luporini (1961).
25. F. Dostoevskij. *Prestuplenie i nakazanie* (1866) – *Delitto e Castigo*. G. Kraiski (1983).
26. I. Turgenev. *Otcy i deti* (1860-1861) – *Padri e figli*. F. Verdinois (1908).
27. I. Gončarov. *Oblomov* (1849-1858) – *Oblomov*. A. Michettoni (2000).

28. N. Gogol'. *Mertvyje duši* (1835-1852) – *Le anime morte*. P. Nori (2009).
29. F. Dostoevskij. *Belye noči* (1848) – *Le notti bianche*. P. Biloni (2006).
30. N. Gogol'. *Šinel'* (1842) – *Il Cappotto*.¹²⁴ C. Rebora (1922) / Duchessa D'Andria (1937) / *Il Mantello*. T. Landolfi (1941) / O. Del Buono (1949) / L. Pacini-Savoj (1957) / G. Pacini (1963) / A. Julovic (1964) / P. Zveteremich (1967) / E. Bazzarelli (1980) / S. Beffa (1986) / F. Mariano (1986) / L. De Nardis (1993) / *La Mantella*. S. Prina (1994) / E. Guercetti (1995) / F. Legittimo (2001).

D'ora in poi, tutti gli esempi, ove non indicato altrimenti, saranno tratti da questi testi. Ogni esempio originale russo sarà accompagnato dall'indicazione del nome dell'autore tra parentesi. Rimandiamo a questo elenco per le restanti informazioni relative ad anno, titolo e traduzione. Il titolo dell'opera sarà segnalato solo nei casi in cui ad un singolo autore dovesse corrispondere più di un'opera (precisamente: Nabokov, Čechov, Gor'kij, Dostoevskij e Gogol').

3.1. Lo studio dei prefissi verbali nella russistica: dagli anni '90 ad oggi

Il sistema di prefissazione verbale russo rientra tradizionalmente in due campi di indagine: la formazione delle parole (*slovoobrazovanie*) e l'aspettologia (Krongauz 1998: 41). Nel primo ambito i prefissi verbali russi si segnalano per la loro polisemia, nel secondo per la polifunzionalità (Krongauz 1994: 34-35); ciò contribuisce a creare un quadro complesso, a cui si aggiunge un ulteriore elemento, ossia la difficoltà nel distinguere fra semantica del prefisso e semantica del verbo (Zaliznjak 2006: 299-300) e l'interrogativo sull'esistenza o meno di una relazione motivazionale fra i vari significati di un unico prefisso (fermo restando che nella maggior parte dei casi, come afferma la stessa Zaliznjak, si tratta di polisemia e non di omonimia)¹²⁵. Inoltre, lo studio della semantica dei prefissi è spesso complicato dall'esistenza di verbi prefissali la cui radice,

¹²⁴ Dove non segnalato altrimenti, anche gli altri traduttori hanno optato per questo titolo.

¹²⁵ I lavori dedicati alla semantica di uno o più prefissi sono innumerevoli. Tra quelli più recenti, rimandiamo agli svariati saggi dello stesso Krongauz, redattore, tra l'altro, di uno dei numeri del *Moskovskij Lingvističeskij Žurnal* del 2001, interamente dedicato ai prefissi verbali, ai lavori di Anna Zaliznjak sul prefisso *za-* (cfr. 2006: 298-341), e in generale agli studi del gruppo di ricerca guidato da Laura Janda dell'università di Tromsø, fondati sull'analisi dei prefissi in chiave cognitiva (cfr., ad esempio, Nessel, Bayadimirova e Janda (2009), impegnate in un confronto tra la semantica di *iz-* e *vy-*).

se presa singolarmente, presenta tutt'altro significato rispetto a quello del verbo prefissale (es.: *zabyt'* – *dimenticare*, composto da *za-* + *byt'* – radice che di per sé significa *essere*).

Per quanto riguarda il processo di perfettivizzazione tramite prefissi, ricordiamo che questi ultimi in alcuni casi specifici vengono definiti “vuoti” nella convinzione che abbiano un valore puramente aspettuale e che non apportino alcun cambiamento semantico alla base. Come ricorda Krongauz (1998: 65), questa tradizione, risalente alla prima metà del XX secolo, è stata anche recepita dai manuali scolastici¹²⁶. Già nel secolo scorso, tuttavia, la tesi dei prefissi desemantizzati cominciò ad essere messa in discussione (cfr. Thelin 1984), e il dibattito è oggi ancora aperto.

Tra i linguisti contemporanei impegnati nel discernere l'esistenza o meno di prefissi puramente aspettuali o semanticamente vuoti citiamo Laura Janda, che descrive questo fenomeno anche nei suoi lavori più recenti (Janda 2012: 4, Kuznecova, Janda 2013: 91-92). I verbi perfettivi risultanti dall'aggiunta di un prefisso vuoto alla forma imperfettiva sono definiti dalla studiosa “*Natural perfectives*” (*perfettivi naturali*) e comprendono verbi quali, ad esempio, *napisat'* (*scrivere*, perfettivo di *pisat'*) o *sdelat'* (*fare*, perfettivo di *delat'*) (Janda 2007)¹²⁷. Janda e il suo gruppo di lavoro dimostrano come i prefissi che la tradizione definiva vuoti (cfr. nota 125) non possono essere considerati tali: partendo dalla “*Overlap Hypothesis*”, secondo cui i prefissi aspettuali non sono vuoti, ma se mai il loro significato si sovrappone al significato del verbo base, la studiosa si spinge oltre, affermando che anche prefissi come *-na* in *napisat'* e *s-* in *sdelat'* non solo hanno un valore semantico, ma possono essere intesi come classificatori verbali, simili a quelli esistenti in altre lingue, come il Mandarino o l'Urdu. Queste idee, raggruppate nel volume “*Why Russian aspectual prefixes aren't empty. Prefixes as verb classifiers*” (Janda et al. 2013), sono state in parte criticate da Zaliznjak e Mikaeljan (2012), sebbene anche queste ultime confermino che l'idea dei prefissi vuoti sia ormai superata.

Molti altri sono gli studi intesi a dimostrare che alcuni dei sopracitati prefissi vuoti apportano in realtà un cambiamento semantico al verbo base; citiamo - a titolo

¹²⁶ In “*Russkij jazyk*”, nel '47, V. Vinogradov parla di “prefissi vuoti” rifacendosi a A. Šachmatov (*Očerki sovremennogo russkogo literaturnogo jazyka*, 1941), il quale a sua volta riprende G. Ul'janov (*Značenijs glagol'nych osnov v litovsko-slajanskom jazyke*, 2, 1891), specificando però che la distinzione tra prefissi “puramente aspettuali” e prefissi con significato proprio era stata introdotta già nella *Praktičeskaja grammatika russkogo jazyka* di N. Greč (1827) e nella *Russkaja grammatika* di A. Vostokov (1831) (Vinogradov 2001: 420-21).

¹²⁷ Cfr. anche i lavori di Janda e del gruppo dell'Università di Tromsø relativi al database “*Exploring Emptiness*” (<http://emptyprefixes.uit.no>) (gennaio 2017).

esemplificativo e senza alcuna pretesa di fornire un elenco completo – Volochina, Popova (1997) e Pozolotina (2009, 2010), concentratesi sul prefisso *vy-*, e Gjervold (2013), che porta l'esempio di *marat'* (sporcare) e dei suoi quattro perfettivi *vymarat'*, *zamarat'*, *izmarat'* e *namarat'*, definiti naturali: l'autore dimostra che, pur presentando un significato molto simile, questi quattro verbi non sono pienamente interscambiabili, e pertanto si deve ammettere che i prefissi abbiano di volta in volta un ruolo chiave nella variazione semantica dei verbi in cui compaiono.

Sempre sul tema della perfettivizzazione, si segnalano indagini che esplorano la regolarità o meno di certi prefissi. Nettet, Janda et al. (2008), ad esempio, si interrogano sul perché i verbi di percezione attiva, come *smotret'* (guardare) e *slušat'* (ascoltare), formino il perfettivo con il prefisso *po-*, mentre i corrispondenti di percezione “passiva” *videt'* (vedere) e *slyšat'* (sentire) preferiscano *u-*. Le autrici ipotizzano che la scelta del prefisso sia determinata dalla semantica del verbo¹²⁸.

Aggiungiamo, inoltre, che in russo la prefissazione può essere doppia o tripla, come ricorda Tatevosov (2009) in uno dei suoi saggi sui verbi prefissali basato sugli esempi del NKRJa (cfr. ad esempio, *po-pere-za-pisyvat'* o *do-pere-za-pisat'*, verbi presentati dall'autore stesso).

A completare il quadro delle questioni relative alla prefissazione, vi è infine la ricorrenza nella lingua russa di verbi prefissali creati occasionalmente da parlanti o autori e quindi non registrati nei dizionari. Come ricorda Krongauz (1998: 4), verbi prefissali di questo tipo ricorrono molto spesso nella prosa letteraria. Comprendere tali verbi, tuttavia, non costituisce un problema per i parlanti madrelingua, i quali inferiscono per analogia anche il significato delle varianti più improbabili. Le difficoltà in questo caso – come ricorderemo anche in seguito nel corso del capitolo – sono sentite piuttosto dagli apprendenti stranieri. Riportiamo qui alcuni esempi proposti dal linguista:

(1) *Рупь последний в Сочи трачу, / Телеграмму накатал: / «Шлите денег — отбатрачу, / Я их все прохохотал»* (V. Vysockij)

(2) *Любовь надо **вылюбить**, а то свербить, сверлит, ныть будет* (V. Kornilov).

Nell'esempio (1), il destinatario è portato ad associare i due verbi prefissali rispettivamente a *otrabotat'* (compensare, ripagare con il lavoro) e *propit' / proest'*

¹²⁸ Cfr. anche Martinowsky (1998), per uno studio sulla perfettivizzazione in generale.

(*spendere tutti i soldi in alcool / in cibo*), ossia a verbi con lo stesso prefisso e il cui significato - chiaro e attestato - in tale contesto può servire da modello per comprendere i verbi *otbtračit'* e *prochohotat'*, che significherebbero, dunque, “*ripagare facendo il bracciante*” e, nel secondo caso, “*spendere tutti i soldi ridendo a crepelle*”. Nell'esempio (2), secondo Krongauz, è difficile trovare verbi prefissali analoghi che consentano di inferire il significato di *vyljubit'*; eppure anche in questo caso la frase è assolutamente comprensibile per i parlanti di lingua russa. In realtà, interrogando alcuni madrelingua russi, emerge un'analogia con il verbo *vypit'* (*bere fino in fondo*), per cui *ljubov' nado vyljubit'* potrebbe significare *l'amore va consumato fino in fondo*.

Ma l'importanza della semantica dei prefissi è ancora più rilevante di fronte a verbi prefissali con basi inesistenti (3): è chiaro come in questi casi sia il prefisso a trasmettere completamente la semantica del verbo, poiché la base a cui si associa non esiste come verbo a sé stante nella lingua russa, e non ha dunque un significato proprio (Krongauz 1998: 11).

(3) *Если этот твой дружок... Рома... будет продолжать интересоваться тем... что абсолютно его не касается... мы его с орлятами так отволохаем – век не забудет!* (V. Попов).

In questo caso, non solo il verbo prefissale *otvolochat'*, ma nemmeno il verbo semplice *volochat'* è attestato nei dizionari. Eppure, di nuovo, la frase è comprensibile, sia grazie al contesto, sia poiché a fungere da modello intervengono altri verbi prefissali quali *otdelat'* (*ridurre male*), *otlupit'* (*picchiare*), nonché ulteriori forme verbali tipiche del turpiloquio.

Infine, la particolarità dei verbi prefissali russi si delinea in modo molto evidente se si adotta un'ottica tipologica. In generale vediamo che, secondo Leonard Talmy (2001, vol II), russo e italiano appartengono a due gruppi di lingue tipologicamente distinte: lingue *satellite-framed*, da un lato, e lingue *verb-framed*, dall'altro. Questa distinzione, inizialmente introdotta da Talmy per i soli verbi di moto, è stata in seguito estesa anche ad altri gruppi di verbi. Per “satellite” si intende una componente distinta, ma dipendente dalla radice del verbo, espressa sotto forma di affisso o di parola a sé stante: per esempio, nei *phrasal verbs* dell'inglese i satelliti sono le preposizioni, mentre in russo sono proprio i prefissi verbali. Le lingue *satellite-framed* sono dunque quelle lingue in cui il satellite

porta la componente semantica distintiva, come nei verbi di moto russi, in cui la direzione del movimento (caratteristica distintiva) non è espressa dal verbo base in sé, ma dal satellite. Il verbo base contiene invece altre informazioni, come ad esempio il mezzo o il modo del moto. Al contrario, nelle lingue *verb-framed* come l'italiano e, in generale, le lingue romanze, l'informazione sulla direzione del movimento è espressa dal verbo, che tuttavia non è in grado di veicolare tutte le specifiche informazioni contenute nei verbi prefissali che esprimono il moto nella lingua russa (Talmy 2001, vol II: 102-106)¹²⁹.

Come il russo e le altre lingue slave¹³⁰, anche le lingue germaniche e il latino rientrano nel gruppo delle lingue *satellite-framed*, nonostante poi tutte le lingue romanze si siano evolute altrimenti. Nelle lingue romanze i verbi prefissali si sono semplificati: in italiano, ad esempio, le forme non sono trasparenti e verbi che in latino erano prefissali ora non vengono percepiti come tali: per es. il verbo italiano *uscire*, derivante dal latino “*ex + ire*” (Alisova, Repina, Tariverdieva 1987: 241), non viene più riconosciuto come forma derivata tramite prefisso. È pur vero che in italiano, soprattutto nella lingua colloquiale e in alcune varianti regionali (specialmente del nord Italia), si attestano verbi sintagmatici formati da verbo + preposizione, anche con significato figurato e non necessariamente di moto (cfr.: *stare su, correre dietro, andare su, andare giù, lavar via* ecc.); come ricorda Cordin (2011), tale procedimento è un'evoluzione del sistema di prefissazione latino¹³¹, tuttavia, trattandosi per lo più di regionalismi, non si può parlare dell'italiano come lingua *satellite-framed*.

¹²⁹ Riportiamo alcuni esempi di traduzione dal russo all'italiano in cui il traduttore ha due possibilità: tralasciare alcuni dettagli (4b) o esprimerli per mezzo di strategie alternative (5b):

(4a) *Иван Николаевич как упал на скамейку, не добежав до турникета (...)* (M. Bulgakov)

(4b) *Ivan Nikolaevich, che si era lasciato cadere su una panchina senza arrivare fino all'uscita (...)*

(5a) (...) *тут вбежали люди.* (M. Bulgakov)

(5b) (...) *entrarono di corsa degli uomini.*

In (4a) e (5a) il verbo russo *bežat'* (*correre*) rappresenta la radice e contiene le informazioni sul modo in cui avviene il movimento. I prefissi *do-* e *v-* sono invece i satelliti, ed esprimono la direzione del movimento: *do-* indica l'arrivo fino a un punto, mentre *v-* indica l'ingresso in un luogo. In (4b) un tratto semantico del verbo *dobežat'* è omesso; qui la traduttrice conserva il dettaglio sulla direzione dell'azione, tralasciando però il modo (correndo). In (5b), invece, aggiunge il complemento di modo “*di corsa*” al verbo *entrare* garantendo così la trasmissione di tutti i tratti semantici del verbo russo, pur distribuiti tra più elementi linguistici (verbo + complemento).

¹³⁰ Per una breve rassegna sulla prefissazione in tutte le lingue slave, rimandiamo al saggio di presentazione di un volume speciale di *Nordlyd*, tutto dedicato ai prefissi verbali slavi, a cura di Svenonius (Svenonius 2004: 192-200).

¹³¹ Per un approfondimento sui verbi sintagmatici dell'italiano, cfr. i lavori di Masini, come ad esempio (2008) e (2012), che offrono un'ottima panoramica del fenomeno anche in chiave contrastiva con le lingue germaniche.

3.2. Lo studio dei prefissi verbali in chiave corpus-based

Se i prefissi verbali non destano problemi ai parlanti di lingua russa, inferirne la semantica è molto più complesso per l'apprendente straniero, poiché difficilmente un comune dizionario soddisferà le sue esigenze di comprensione. E ciò vale non solo per i verbi prefissali di fantasia, come negli esempi (1), (2) e (3), ma anche per i casi più regolari. Come nota Dobrušina, “il comportamento di questi verbi sta su un piano di astrazione più profondo rispetto alla definizione lessicografica”¹³² (Dobrušina 2010: 100). Ecco perché uno studio sui prefissi in chiave contrastiva e fondato su una base empirica abbastanza solida, come quella offerta da un corpus parallelo, diventa particolarmente utile sia in senso teorico, sia per la didattica del russo come L2. La linguistica dei corpora arriva laddove l'intuizione dei parlanti non è più sufficiente. E non solo. I materiali compilati con metodologie *corpus-based* integrano e migliorano le grammatiche e i dizionari tradizionali che, compilati sulla base dell'introspezione e di esempi non autentici, non offrono una visione esaustiva dei comportamenti linguistici¹³³.

Riferendosi all'utilità dei corpora in lessicografia, Meyer (2002: 17) ricorda come Fillmore (1992: 39-45), lavorando alla teoria della *frame semantics*, avesse scoperto che solo due dizionari su dieci evidenziavano tutti e tre i significati della parola inglese “*risk*” (da lui meglio definiti come “*frame elements*”): il significato principale di *risk* infatti varia, secondo Fillmore, in dipendenza di tre fattori: 1) se il rischio non è causato dall'azione di qualcuno; 2) se il rischio è dovuto a un'impresa di qualcuno; 3) se il rischio risulta dalla decisione di qualcuno di compiere l'impresa. Ricercando il termine tra i testi di un corpus di 25 milioni di parole (un'entità relativamente esigua rispetto ai corpora che contano centinaia di milioni di parole), il linguista trovò invece che “*risk*” non solo era usato in tutti e tre i significati, ma compariva anche in alcuni esempi non corrispondenti al suddetto “*risk frame*”, cosa che, a detta di Fillmore stesso, non si sarebbe mai potuta scoprire senza un corpus.

L'analisi che presentiamo parte dunque dalla constatazione di alcune lacune lessicografiche per quanto riguarda i verbi prefissali aventi la medesima radice e un

¹³² “Поведение этих глаголов существует на более глубоком уровне абстракции, чем словарное толкование”.

¹³³ Nel primo capitolo abbiamo ricordato che già negli anni '80 Sinclair e i colleghi dell'Università di Birmingham avevano inaugurato l'era dei dizionari di lingua inglese *corpus-based* con il loro COBUILD. Ma la strada per far sì che tutti i dizionari vengano compilati sulla base dei corpora sembra essere ancora lunga.

significato molto simile, ma non affatto identico; tali lacune sono riscontrabili sia in dizionari bilingui (nel nostro caso di italiano-russo), sia in dizionari monolingui¹³⁴.

Nel paragrafo 3.1. abbiamo citato l'esempio di *marat'* (*sporcare*), con i suoi quattro perfettivi, ognuno dotato di un prefisso differente: le definizioni lessicografiche di questi verbi denotano effettivamente una certa superficialità nel loro trattamento. Ožegov, ad esempio definisce l'imperfettivo *marat'* in questo modo:

МАРАТЬ, -аю, -аешь; маранный; несовер. (разг.).

1. кого (что). **Пачкать, грязнить.** *М. одежду в грязи. М. руки* (также перен.: делать что-н. предосудительное или недостойное. *М. честное имя* (перен.: позорить, порочить)).

2. что. **Вычёркивать из написанного.** *М. строки в рукописи.*

3. что. **Плохо рисовать, писать, сочинять.** *М. стихи. М. акварелью.*

Nella stessa voce fornisce poi la lista dei perfettivi prefissali, specificando che tutti tranne *namarat'* si riferiscono al primo significato:

совер. **вымарать**, -аю, -аешь; -анный (к 1 знач.), **замарать**, -аю, -аешь; -аранный (к 1 знач.), **измарать**, -аю, -аешь; -аранный (к 1 знач.) и **намарать**, -аю, -аешь; -аранный (к 3 знач.).

Da ciò si potrebbe concludere che almeno tre di questi quattro perfettivi sarebbero perfettamente interscambiabili, a differenza di quanto dimostrato da Gjervold (2013). Se ricercati singolarmente nello stesso dizionario, in effetti, non si otterranno né definizioni, né esempi, ma si verrà rimandati all'imperfettivo *marat'*. L'unica eccezione è *vyumarat'*, per cui Ožegov attesta un ulteriore significato:

ВЫМАРАТЬ, -аю, -аешь; -анный; совер.

1. см. мара́ть.

2. что. **Зачеркнуть, вычеркнуть** (прост.). В. строчку.

¹³⁴ Più precisamente, i dizionari di lingua russa a cura di D. N. Ušakov (*Tolkovyj slovar' russkogo jazyka. V 4 tomach*), di S. I. Ožegov e N. Ju. Švedova (*Tolkovyj Slovar' Ožegova*), di A. P. Evgeneva (*Malyj Akademičeskij slovar*) e il dizionario bilingue a cura di V. Kovalev (*Il Kovalev. Dizionario russo-italiano, italiano-russo*).

Il dizionario bilingue Kovalev attribuisce quattro significati all'imperfettivo *marat'* (a differenza di Ožegov, suddivide il primo significato di quest'ultimo in due semi distinti):

1. **sporcare, insudiciare;**
2. **macchiare, disonorare;**
3. **scarabocchiare, impiasticciare** (imbrattare la carta, scrivendo cose prive di talento);
4. **cancellare** (spec. ciò che è scritto).

Tuttavia, menziona solo due perfettivi: *zamarat'* (nei sign. 1-3) e *vymarat'* (nel sign. 4). In conclusione, ne risulta un quadro abbastanza confuso, che richiederebbe ulteriori approfondimenti, ma fin qui il nostro intento è semplicemente volto a mostrare come le descrizioni lessicografiche spesso siano contraddittorie per quanto riguarda l'analisi dei verbi prefissali.

I verbi di memoria, nostro oggetto specifico, presentano un trattamento lessicografico analogo; riportiamo pertanto le voci dizionariali tratte da Ožegov¹³⁵ e da Kovalev del verbo *pomnit'* e dei suoi derivati prefissali *vspomnit'*, *zapomnit'* e *pripomnit'*:

1) Ožegov

ПОМНИТЬ, –ню, –нишь; несовер., кого (что) и о ком (чём).

Сохранить, удерживать в памяти, не забывать. II. *своё детство.* II. *стихотворение наизуст.* II. *о своих обязанностях.*

ВСПОМНИТЬ, –ню, –нишь; совер., кого (что) и о ком (чём).

1. **Возобновить в памяти, вернуться мыслью к прошлому.** В. *свою молодость.*

2. **Внезапно вернуться мыслью к забытому, упущенному.** В. *о важном деле.*

Вспомнил, что обещал позвонить.

¹³⁵ II *Malyj akademičeskij slovar'* (MAS) e il dizionario di Ušakov si comportano allo stesso modo: ПОМНИТЬ è definito, rispettivamente, come “Удерживать в памяти, не забывать” (MAS) e “Держать в памяти, не забывать.” (Ušakov); ВСПОМНИТЬ: “Восстановить, возобновить в памяти” (MAS) e “Воспроизвести в памяти, привести себе на память” (Ušakov); ЗАПОМНИТЬ: 1. “Сохранить, удержать в памяти”. 2. (с отрицанием „не“: не запомнить). “Не помнить; не быть в состоянии припомнить” (MAS) e 1. “Удержать, сохранить в памяти”. 2. Только с отриц. “Помнить” (разг.) (Ušakov).

ЗАПОМНИТЬ, -ню, -нишь; совер.

1. КОГО (ЧТО). **Сохранить в памяти.** 3. *стихи. 3. своего попутчика.*

2. ЧТО КОМУ. **То же, что припомнить (в 1 знач.).** 3. *Обиду. Берегись, он тебе это запомнить.*

ПРИПОМНИТЬ, -ню, -нишь; совер.

1. КОГО (ЧТО). **Вспомнить, восстановить в памяти, вернуться мыслью к прошлому.**

П. чьи-н. советы. Имя этого человека не припомню (забыл).

2. ЧТО КОМУ. **Не забыть сделанного кем-н., чтобы отомстить, наказать, попомнить (разг.).** *Я это тебе припомню!*

Le descrizioni di *pomnit'* e *zapomnit'* sono molto simili tra loro, sebbene i due verbi siano usati in contesti differenti e con diverse sfumature di significato. Per quanto riguarda invece *vspomnit'* e *prpomnit'*, le descrizioni date nella voce evidenziano tratti distintivi rispetto al semplice *pomnit'* (il ricordo non è statico, ma improvviso; il ricordo si riferisce a fatti o informazioni temporaneamente dimenticati/e), tuttavia gli esempi riportati da Ožegov non li mettono in evidenza (cfr. *pomnit' svoe detstvo* e *vspomnit' svoju molodost'*) e dunque la differenza semantica tra il ricordo costante di *pomnit'* e il ricordare improvviso di *vspomnit'* non risulta chiara. Lo stesso accade nel dizionario di Ušakov, a cui Ožegov si è ispirato. Fa eccezione il MAS, che pur non differenziandosi da questi due dizionari nelle definizioni (cfr. nota 134), per il verbo *vspomnit'* porta alcuni esempi più significativi: *vspomnit' zabytoe stichotvorenje; Vse eto uže kogda-to bylo v ee žizni. No gde i kogda – ona ne mogla vspomnit'* (Kataev).

Per quanto riguarda i tratti distintivi di *zapomnit'* rispetto a *pomnit'* - trattenere un ricordo o un'informazione nella memoria a partire da un dato momento -, più utile ci sembra una descrizione proposta da Vasil'ev, che in un lavoro dedicato a vari gruppi di verbi prefissali definisce *pomnit'* come “avere in mente una qualche conoscenza”¹³⁶ e *zapomnit'* come “consolidare, ritenere nella memoria”¹³⁷ (Vasil'ev 1981: 165-168).

Esaminiamo ora le voci di *pomnit'* e dei suoi derivati nel dizionario bilingue di V. Kovalev.

¹³⁶ “Иметь в памяти какие-либо знания”.

¹³⁷ “Закреплять в памяти”.

2) Kovalev:

ПОМНИТЬ *imperf.* (-ню, -нишь)

ricordare, rammentare, tenere a mente. *n. детство* ricordare l'infanzia; *n. o друзьях* ricordarsi degli amici; *не п. себя от возмущения* essere fuori di sé dall'indignazione.

ВСПОМНИТЬ: *perf.* (-ню, -нишь) (*imperf.* вспоминать)

1. (*возобновить в памяти*) **ricordare, rievocare** *в. свою молодость* ricordare la propria giovinezza.

2. (*внезапно вернуться мыслью*) **rammentare, ricordarsi;** *в. о важном деле* ricordarsi di una faccenda importante.

ЗАПОМНИТЬ: *perf.* (-ню, -нишь) (*imperf.* запоминать)

ricordare, ritenere a mente, memorizzare *он хорошо запомнил своего попутчика* (lui) ricorda bene il suo compagno di viaggio; *я запомнил эти слова* ho ricordato quelle parole.

ПРИПОМНИТЬ: *perf.* (-ню, -нишь) (*imperf.* припоминать *nel sign.* 1)

1. (*восстановить в памяти*) **ricordare, rammentarsi** *n. советы врача* ricordare i consigli del medico.

2. (*попомнить*) **non perdonare, ricordare per vendicarsi** *я тебе это припомню* (*угроза*) non te lo perdonerò, me la lego al dito.

Le descrizioni di Kovalev si differenziano parzialmente l'una dall'altra, ma il primo significato è sempre "ricordare". Inoltre, anche qui, gli esempi non permettono di evincere le informazioni necessarie all'uso attivo contestualizzato. L'incompletezza del trattamento lessicografico ci ha spinti dunque a indagare più a fondo questo gruppo di verbi mediante una ricerca *corpus-based*, la quale risulta particolarmente agevole nel NKRJa, dal momento che il motore di ricerca avanzato consente di fare la ricerca sia per prefissi (es.: tutti i verbi con il prefisso X), sia per basi verbali. Nel nostro caso abbiamo cercato tutti i verbi con base *pomnit'* (*ricordare*) o *-pominat'* (variante imperfettiva). Ciò è possibile nell'area di ricerca "lessico-grammaticale", antepoendo un asterisco alla base.

Nel paragrafo 3.4. esporremo l'analisi dei tre derivati di *pomnit'* condotta sulla base del CP ru-ita, non senza aver prima introdotto brevemente nel paragrafo che segue la semantica dei tre prefissi che li caratterizzano, così come viene descritta dalla lessicografia russa.

3.3. Prefissi che caratterizzano i verbi di memoria nella lingua russa

Come afferma Rebrina (2008: 285), i linguisti sono sempre più impegnati nel tentativo di ricostruire l'immagine dell'uomo, delle sue attività e delle sue capacità in base ai dati offerti dalla lingua, e una delle capacità fondamentali dell'essere umano è proprio la memoria, ricondotta da Apresjan (1995b: 48) all'ambito dell'intelletto. Il concetto di memoria si lega infatti all'azione del “*sapere*”: “*ricordare*” e “*sapere*” spesso si sovrappongono, tanto che *ne pomnju* (*non ricordo*) di fatto significa *bol'she ne znaju* (*non lo so più*) (Zaliznjak 2013: 234-238).

Ci concentriamo ora sulla verbalizzazione dell'atto di ricordare nella lingua russa, espresso di volta in volta da verbi differenti, le cui sfumature di significato sono rese dalla combinazione con diversi prefissi.

Presentiamo dunque innanzitutto una descrizione dei tre prefissi *pri-*, *vs-* e *za-* associati a *pomnit'*, dal momento che il significato di un verbo prefissale dipende da tre fattori: 1) il significato del prefisso; 2) le caratteristiche della base senza prefisso; 3) il modello di reggenza del verbo prefissale (Zaliznjak 2006: 3). Il significato del prefisso sarebbe inoltre una costante, mentre gli altri due fattori variano e la loro interdipendenza dà luogo di volta in volta a un significato differente per il verbo prefissale nel suo complesso.

Per quanto riguarda i prefissi che stiamo analizzando, la loro invariante semantica può essere sinteticamente espressa nel seguente modo:

1) *VS-*: il significato principale di questo prefisso è “movimento verso l'alto” (Barychina et al. 1989: 10) e nel verbo *vspomnit'* questo valore sembra mantenersi in senso figurato: un pensiero che era stato dimenticato giace sul fondo della nostra mente e insorge in un preciso momento. Il processo può essere incontrollato o volontario. Zaliznjak (2013: 330) propone di considerare la memoria umana come un baule pieno di oggetti: i ricordi. Tra questi i meno utili si trovano sul fondo e metaforicamente sono quindi paragonabili agli

oggetti che non usiamo mai: si trovano nelle baule delle cose utili, ma non li vediamo, tant'è che talvolta non sappiamo di averli. Il movimento dal basso verso l'alto espresso dal prefisso *vs-* *va* immaginato come il moto compiuto dall'oggetto dimenticato nel momento in cui viene estratto di tanto in tanto dal fondo del baule, per caso o volontariamente.

2) **ZA-**: questo prefisso è caratterizzato da un'elevata polisemia; tra i significati considerati nella nostra analisi, ci soffermiamo su quello di *polučit'* (*ottenere*) e *zachvatit'* (*afferrare*), identificati da (Barychina et al. 1989: 18): nel verbo *zapomnit'* infatti il prefisso sembrerebbe indicare l'atto di ottenere un'informazione, afferrarla e conservarla, fissandola nella mente. Anche Zaliznjak (2006: 317-320) attribuisce a *zapomnit'* il significato GET (*afferrare*) o FIX (*fissare*).

In realtà, la semantica di *zapomnit'* è molto più complessa di quanto possa sembrare a prima vista, come nota la stessa Zaliznjak. In alcuni contesti, infatti, può sembrare che questo verbo assuma il significato di “cominciare a ricordare” e che quindi sia paragonabile a verbi come *zagovorit'* (*iniziare a parlare*) con significato incoativo (BEGIN - *iniziare*). La studiosa nota tuttavia che, per essere incoativo, un verbo prefissale deve rispondere a tre criteri: il prefisso deve 1) essere quanto più possibile indipendente dal punto di vista semantico; 2) esprimere l'inizio della stessa azione espressa dal verbo base senza prefisso (es: *plakat'* = *piangere*; *zaplakat'* = *iniziare a piangere*); 3) esprimere un “processo omogeneo”, che si svolge cioè in modo identico durante tutta la sua durata, senza presentare una fase iniziale o finale diversa da quella centrale. La studiosa porta l'esempio del movimento di un'altalena: senza fine e senza meta. I verbi con *-za* incoativo, infatti, spesso indicano un'azione legata alla percezione sensoriale (es.: *zašumet'* – *udito*, *zablagouchat'* – *olfatto*). *Zapomnit'*, dunque, non rientrerebbe nel gruppo dei verbi con *za-* incoativo, non solo perché il suo tratto principale e inequivocabile è il già citato FIX, ma anche perché non risponde al secondo criterio, per cui *za-* dovrebbe indicare “*načat' pomnit'*” (*iniziare a ricordare*) (Zaliznjak 2006: 329-333). *Zapomnit'* significa invece non tanto “iniziare a ricordare”, quanto “iniziare a sapere”.

Interessante, infine, è notare che Dal' nel suo dizionario registra un ulteriore significato di *zapomnit'*, opposto a quello odierno, ovvero *dimenticare*. In questo caso, il prefisso *za-* significherebbe “annullare l'azione espressa dal verbo base” (ANNIHIL -

annullare). La studiosa ritiene che questa vera e propria “polarizzazione” di significati non sia altro che una casualità, analoga a quella di altri verbi, quali *zarisovat’* che può essere usato come “imprimere” (*zarisovat’ slona* – disegnare un elefante) e “coprire” e quindi “annullare” (*zarisovat’ list bumagi slonami* – riempire un foglio di elefanti disegnati).

3) **PRI-**: uno dei principali significati di *pri-* è “condurre un’azione o un oggetto fino a un punto d’arrivo”¹³⁸ (Barychina et al. 1989: 51). Quando qualcuno “*pripominaet*” (*ricorda, richiama alla mente*) un fatto è come se lo riconducesse alla propria mente. Jakunina (2001: 144) associa *pripomnit’* al gruppo di verbi per cui *pri-* assume il significato di “*pojavlenie videnij/myslej*” (*comparsa di idee/pensieri*). Si puntualizza però che anche in questo caso si tratta di un movimento, se pur un movimento di pensieri, e pertanto Jakunina lo colloca come sotto-gruppo all’interno del macro-gruppo dei verbi che indicano “*pomestit’(sja) v mesto Z*” (*spostare(si) in un posto Z*). Tra questi, oltre a *pripomnit’*, troviamo *pridumat’* (*inventare/avere un’idea*), ma anche verbi come *prisnits’ja* (*sognare*) (Jakunina 2001: 144).

3.4. Analisi delle traduzioni in italiano

Prima di passare all’analisi di *vspomnit’*, *zapomnit’* e *pripomnit’* (incluse le forme imperfettive *vspominat’*, *zapominat’* e *pripominat’*), chiariamo alcune nostre scelte metodologiche.

3.4.1. Osservazioni metodologiche preliminari

Per svolgere la nostra analisi abbiamo ricercato le occorrenze dei verbi sopra citati nel CP ru-ita ed esaminato le rispettive traduzioni in italiano, ma dovendo restringere il campo di indagine abbiamo escluso le varianti riflessive (es.: *vspomnitsja*) (in futuro potrà essere inclusa anche questa categoria). Non abbiamo considerato nemmeno il verbo *napomnit’/napominat’* (*ricordare qualcosa a qualcuno*, 6b) per via del suo significato inequivocabilmente causativo, anche con soggetto inanimato (7a):

¹³⁸ “*доведение действия до конечного пункта*”.

(6a) Я **напомнил** Тасе, который час. (Dovlatov, *Filial*)

(6b) **Ricordai** a Tasja che ora fosse.

(7a) От нее шел легкий, едва уловимый запах ладана, и **это напомнило** ему время, когда он тоже веровал в бога (...). (Čechov, *Rasskazy*)

(7b) Emanava da lei un leggero odore d'incenso, appena percettibile; e **ciò gli ricordava** il tempo in cui credeva anche lui in Dio (...).

Infine, non abbiamo analizzato gli esempi con *upomnit'* e *popomnit'*, per via della loro ridottissima frequenza, sia nel NKRJa principale, sia nel nostro corpus parallelo. Per dare un'idea delle proporzioni numeriche, riportiamo il numero di occorrenze di ciascuno di questi verbi nel corpus principale e nel corpus parallelo (le cifre comprendono anche i relativi imperfettivi, tranne nel caso di *popomnit'* - che non presenta l'imperfettivo *popominat'* - e *upomnit'*, il cui imperfettivo rimane esclusivamente *pomnit'*. *Upominat'*, infatti, significa *menzionare* e si associa al perfettivo *upomjanut'*).

Le occorrenze dei derivati prefissali di *pomnit'* nel corpus generale sono 139.975, così distribuite:

- *vspomnit'* / *vspominat'* – 80.215
- *napomnit'* / *napominat'* – 39.198
- *zapomnit'* / *zapominat'* – 11.035
- *pripomnit'* / *pripominat'* – 8.501
- *upomnit'* – 681
- *popomnit'* – 345

Per quanto riguarda il CP (escludendo *napomnit'*, per i motivi sopra esposti) abbiamo registrato:

- *vspomnit' / vspominat''* – 622
- *pripomnit' / pripominat'* – 113
- *zapomnit' / zapominat'* – 52
- *upomnit'* – 3
- *popomnit'* – 2

Se *popomnit'* e *upomnit'* sono quasi assenti nel CP, *vspomnit'* è al contrario preponderante.

Gli esempi riscontrati nel CP sono tratti da testi letterari russi con la rispettiva traduzione italiana, per un totale di oltre 3 milioni di parole. Abbiamo però deciso di non esaminarli tutti, in modo da mantenere l'indagine bilanciata: il CP contiene infatti anche alcune opere in versione integrale¹³⁹, a cui, comprensibilmente, si riferisce quasi la metà delle occorrenze totali: nel caso di *vspomnit'*, per esempio, su 468 occorrenze, 224 si trovano nel romanzo di L. Tolstoj *Anna Karenina*. Per evitare dunque di esaminare una percentuale troppo elevata di esempi attribuibili ad uno stesso autore e traduttore, abbiamo deciso di considerare un massimo di 30 occorrenze per ogni opera, precisamente le prime 30¹⁴⁰.

Chiarite le nostre scelte metodologiche, passiamo ora all'analisi degli esempi.

3.4.2. Le strategie traduttive della resa in italiano

Come abbiamo visto nel dizionario di V. Kovalev, spesso la lessicografia bilingue propone come primo traduce di tutti i derivati prefissali di *pomnit'* il solo verbo *ricordare*. Ciò significa che nella traduzione di questi verbi in italiano può avvenire una certa perdita di informazione rispetto all'originale russo. Oltre all'indagine *corpus-based* dei verbi prefissali abbiamo pertanto voluto interrogarci anche su questo aspetto. Le domande che ci hanno guidato nell'analisi sono due: 1) traducendo in italiano tutti i verbi russi considerati indistintamente con l'equivalente *ricordare* si verifica una perdita di informazione? 2) Quando questa perdita non si verifica, ciò accade perché i traduttori hanno adottato strategie particolari per rendere la specificità di ciascun verbo? Se sì, quali?

¹³⁹ Cfr. capitolo 2 per i dettagli sulla dimensione dei testi.

¹⁴⁰ Questa precisazione ci rimanda alle riflessioni del primo capitolo sulla scelta di inserire in un CP testi in versione integrale o estratti: se il nostro CP contenesse lo stesso numero di opere letterarie attualmente disponibili (30), ma in versione integrale, gli esempi sarebbero troppi per essere analizzati dal primo all'ultimo. Pertanto, per garantire eterogeneità di autori e traduttori, si opterebbe per estrarre 20-30 esempi da ogni opera. Tuttavia, resta chiaro che un corpus di dimensioni troppo esigue non garantisce la validità scientifica auspicata. L'utilità di avere un corpus di volume consistente emerge soprattutto per la ricerca di parole più rare o arcaiche. Ma anche appurato che la validità scientifica di un'analisi è direttamente proporzionale alle dimensioni di un corpus, riteniamo che sia più rappresentativo avere, a parità di numero di parole, più estratti di opere diverse, rispetto alla versione integrale di un minor numero di opere. In quest'ultimo caso, infatti, come per *Anna Karenina*, si rischierebbe sempre di dover scartare un gran numero di dati contenuti nella stessa opera.

L'analisi ci ha portato a concludere che, effettivamente, in alcuni casi, si verifica una perdita di informazione, mentre in altri si possono identificare delle macro-strategie di resa che consentono di evitarla. Torneremo al tema della perdita d'informazione nel paragrafo successivo. Qui esporremo invece le strategie impiegate dai traduttori per la resa dei verbi prefissali di memoria considerati.

A. VSPOMNIT'

Del verbo *vspomnit'*, il più frequente, abbiamo analizzato 251 esempi. Innanzitutto è bene distinguere i vari esempi a seconda del modo e del tempo verbale in cui si presenta il verbo russo, verificando se è possibile riscontrare, oltre a specifiche collocazioni, anche fenomeni di “*colligation*”, se vi sono, cioè, forme morfo-sintattiche che i verbi esaminati assumono di preferenza¹⁴¹. Ciò può essere importante per precisare alcuni tratti semantici del verbo in questione.

Abbiamo riscontrato i seguenti casi, in ordine di frequenza:

- passato → *vspomnil / vspomnila / vspomnili* (158 occorrenze)
- gerundio passato → *vspomniv* (31 occorrenze)
- infinito → *vspomnit'* (25 occorrenze)
- imperativo (20 occorrenze)
- futuro (12 occorrenze)
- futuro (perfettivo) ma con valore temporale di presente (5 occorrenze)

Altri *pattern* lessico-grammaticali (collocazioni) ricorrenti sono:

- la presenza di avverbi o locuzioni temporali quali *v to že vremja* (in quello stesso momento), *sejčas* (adesso), *vdrug* (all'improvviso) ecc.;
- la presenza di verbi modali (es.: *moč' – potere, riuscire; starat'sja – cercare di*), avverbi (es.: *medlenno i trudno – lentamente e con difficoltà*) o altre espressioni che testimonino l'idea di sforzo nel ricordare. In questi casi il verbo è spesso alla forma negativa, come in “*ne mogu vspomnit'*” (non riesco a ricordare).

¹⁴¹ (Partington 2004: 131) porta l'esempio del sostantivo inglese *case*, che nella forma plurale *cases* è preceduto più spesso da quantificatori indefiniti, in sintagmi preposizionali come “*in some cases*”, “*in many cases*”.

La frequente occorrenza di *vspomnit'* al passato perfettivo in parte è da attribuire allo stile narrativo tipico della prosa letteraria del nostro corpus, mentre gli avverbi e le locuzioni temporali che accompagnano di solito questo verbo nei nostri esempi indicano che l'azione di ricordare e richiamare alla memoria è in questo caso improvvisa e puntuale. Anche nelle descrizioni delle voci dizionariali troviamo l'avverbio *vnezapno* (cfr. par. 3.2.) che di solito si usa con azioni puntuali: propriamente significa “*inaspettatamente*” e segnala spesso l'involontarietà del ricordo. La puntualità dell'azione in questo caso è un elemento molto importante per definire la semantica del verbo *vspomnit'*, in quanto è il tratto che lo distingue chiaramente da *pomnit'*, verbo stativo. Anche l'imperfettivo *vspominat'* non esprime mai un'azione costante e prolungata come *pomnit'*, bensì un'azione realizzata in un arco di tempo circoscritto e, eventualmente, iterata.

Dai nostri esempi emerge un secondo tratto tipico di *vspominat'/vspomnit'*, ignorato nelle voci dizionariali, e precisamente il tratto conativo, cioè lo “sforzo” volontario compiuto dal soggetto per richiamare qualcosa alla memoria. A seconda del contesto, dunque, l'azione di *vspomnit'* può essere volontaria (tratto conativo) o involontaria (quando il ricordo è improvviso e generato da uno stimolo esterno).

Per quanto riguarda le traduzioni estratte dal corpus, la maggior parte delle volte, com'era prevedibile, i traduttori optano per il verbo *ricordare*; a volte invece ricorrono a diverse strategie di resa che, combinate con “ricordare” o da sole, consentono di trasmettere in modo più completo la semantica di *vspomnit'*. Le abbiamo suddivise in cinque gruppi.

1) **Uso di perifrasi** come *venire in mente, tornare alla mente*, che sembrerebbero il modo più chiaro ed esplicito per rendere l'idea del ricordo che riaffiora. In questi casi, nella traduzione italiana il pensiero, la cosa o l'evento che ritornano alla mente assumono il ruolo di soggetto sintattico, mentre nell'originale russo il soggetto rimane pur sempre la persona che ricorda. In questo modo l'italiano esprime la mancanza di controllo sull'azione da parte del soggetto/paziente. Come detto, tuttavia, il più delle volte in traduzione si riscontra il verbo *ricordare*, di conseguenza su 251 esempi questo tipo di resa compare soltanto in 21 casi (8,3% delle volte):

(8a) «Утомленное солнце!» – **вспомнил** Виктор, узнав наконец музыку этого танго. (Kurkov)

(8b) «*Utomlénnoe solnce!*» **venne in mente** a Viktor, quando infine riconobbe la musica di quel tango.

(9a) Я **вспомнил**, как лет пятнадцать назад заболел мой отец. (Dovlatov, *Filial*)

(9b) **Mi era tornato in mente** quando, quindici anni prima, si era ammalato mio padre.

2) **Espressioni di tempo** che indicano l'**immediatezza** e la **puntualità** dell'azione (29 occorrenze – 11,5%). Parole o locuzioni di questo tipo affiancano il verbo *ricordare* (o gli eventuali sinonimi impiegati¹⁴²) e, cosa molto importante, non ricorrono soltanto quando espresse già nell'originale russo (10b), ma anche quando nel testo fonte sono assenti (11b). Nel primo caso (e questo vale per tutte le strategie descritte da qui in poi), è chiaro che non si può parlare di “strategia di resa”, dal momento che il traduttore si limita a riportare quanto espresso nel testo russo senza nulla aggiungere; nei casi come (11b), invece, emerge l'intento del traduttore di trasmettere l'insorgere improvviso di un evento che interrompe uno stato precedente e contrario, tipico della semantica di *vspomnit'*:

(10a) - Да я, Рустид Фомич, **уж вспомнил**. (Gorkij, *Вуvs'ie ljudi*)

(10b) - Ma, Aristide Fomitsc.... **ora mi ricordo**....

(11a) - **Вспомнил! Вспомнил!** (Bulgakov)

(11b) - **Adesso mi ricordo!**

3) La strategia più frequente dei traduttori dal russo in italiano (121 occorrenze – 48,2%) consiste nel **rendere il verbo ricordare con tempi verbali che indichino**, ancora una volta, la **puntualità** dell'azione. In questo caso, a seconda del contesto, l'uso di tempi quali **passato remoto**, **passato prossimo** e **trapassato prossimo** permette di comprendere che l'atto del ricordare è improvviso (oltre ad essere la resa classica dell'aspetto perfettivo russo, presente nell'originale). Nella narrazione ricorre più spesso il passato remoto (12b). Il passato prossimo è invece più tipico dei dialoghi (13b):

¹⁴² Sui sinonimi cfr. strategia 5.

(12a) (...) *только когда, на другой день утром, увидел мелькнувший мимо окон пакет брата, он **вспомнил** про доверенность (...)* (Gončarov)

(12b) (...) *solo quando, la mattina dopo, vide passare sotto le finestre il solito fascio di carte del fratello, **ricordò** la procura (...)*

(13a) *Увидал я сегодня вас и **вспомнил**, что ведь и вы товарищем мне были...*
(Gorkij, Ozornik)

(13b) *Vi ho veduto questa sera e **mi sono ricordato** che, anche voi, siete stato il mio compagno...*

Includiamo in questo gruppo anche i casi in cui *vspomnit'* compare nella forma del gerundio perfettivo ed è reso quindi in italiano con un gerundio passato o con un participio passato (14b):

(14a) - *Rrrr... - заворчала она, но, **вспомнив** про вчерашний обед, завильяла хвостом и стала нюхать* (Čechov, *Rasskazy*)

(14b) «*Rrrr...*» *incominciò a ringhiare, ma, **ricordatasi** del pranzo della sera prima, si mise a dimenare la coda e ad annusare.*

4) Uso di **espressioni** che indichino **il tratto conativo** (15 occorrenze – 5,9%). Come già notato, *vspomnit'* può indicare un atto volontario fatto per vincere una certa difficoltà a richiamare un pensiero o un fatto alla memoria. Come per le locuzioni di tempo, anche in questo caso il tratto conativo può essere esplicitato già nella versione russa - come in (15a e 15b) con “*постараться вспомнить*”, giustamente reso come “sforzarsi di ricordare” -, oppure può essere implicito in *vspomnit'* (16a) e semantizzato in italiano mediante un verbo dedicato, come in (16b), dove l’espressione “arrivarci” indica lo sforzo compiuto per richiamare alla mente il nome dimenticato dal soggetto:

(15a) (...) *рассуждения здесь ни к чему и самое главное – **постараться вспомнить*** (...) (Šalamov)

(15b) (...) *tutte queste considerazioni sono qui fuori luogo, e la cosa principale è **sforzarsi di ricordare*** (...) (Trad. Binni)

(16a) — *Причем здесь Вульф? Вульф ни в чем не виноват! Во, во... Нет! Так не вспомню!* (Bulgakov)

(16b) *Che c'entra Wulf? Wulf non ha nessuna colpa! Wo, Wa... No, così non ci arriverò mai!*

5) **Sinonimi di ricordare o altri verbi** (16,7% dei casi). Tra questi vi sono traduzioni pertinenti, che da sole o in combinazione con le sopracitate strategie consentono di trasmettere in modo più o meno adeguato la semantica di *vspomnit'* (gruppo A), oppure varianti che non rispecchiano la semantica del verbo originale e quindi rientrano nei casi di “perdita di informazione” menzionati nel paragrafo 3.4.2. (gruppo B). Riportiamo l'elenco completo indicando tra parentesi il numero delle occorrenze:

- **gruppo A:** *rammentare*(si) (16), *sovvenire* (3) *risovvenirsi* (1), *ripensare* (3), *non dimenticare* (2), *non scordare* (1), *riacquistare improvvisamente la memoria* (1).
- **gruppo B:** *pensare* (10), *rifarsi* (1), *dire* (1), *menzionare* (1), *rendersi conto* (1), *tener conto* (1)

Tra le varianti del gruppo A, possiamo distinguere tra sinonimi di *ricordare* (*rammentare*, *sovvenire*), casi di litote (*non dimenticare*, *non scordare*) e varianti più precise, che si avvicinano maggiormente alla semantica di *vspomnit'*, come il ricordo che torna alla mente dopo un periodo di assenza: *ripensare*, *risovvenirsi* e *riacquistare improvvisamente la memoria*. Notiamo poi la relativa frequenza con cui *vspomnit'* è tradotto con *pensare*: ciò rimanda all'analogia evidenziata da Zaliznjak (2013) tra memoria e pensiero (cfr. par. 3.3.), che sarà ancora più evidente quando analizzeremo l'imperfettivo *vspominat'*. Riportiamo di seguito alcuni esempi dal corpus per il gruppo A (17) e per il gruppo B (18):

(17a) *Аня вспомнила, как года три назад на балу (...)* (Čechov, *Rasskazy*)

(17b) *Anna si rammentò che una sera, tre anni prima (...)*

(18a) — *Приезжала ваша жена, — вспомнила сестра.* (Družnikov)

(18b) «È passata sua moglie» *disse l'infermiera.*

Se *rammentare* in (17b) è usato come *ricordare*, e per di più al passato remoto, diverso è il caso di *dire* in (18b), dove il traduttore sceglie di omettere un tratto, quello del ricordo improvviso dell'infermiera, per esprimere la sua conseguenza, ossia la verbalizzazione del pensiero¹⁴³.

Altri casi: il 9,1% delle occorrenze, infine, comprende:

- 6 esempi in cui *vspomnit'* nella traduzione è stato completamente omissso, come in (19b):

(19a) *Я вспомнил, что Анюта Благово за всё время не сказала со мною ни одного слова.* (Čechov, *Rasskazy*)

(19b) *Anjùta Blàgova non m'aveva detto una sola parola.*

- 17 esempi in cui il traduttore non adotta una strategia particolare, ma il contesto ha comunque una funzione chiarificatrice e lascia intendere che l'originale russo, tra i vari verbi di memoria, non potrebbe presentare altro che *vspomnit'*, come in (20b), dove è chiaro che l'esortazione si riferisce a ricordare qualcosa contestualmente al momento dell'enunciazione e non a ritenere nella memoria qualcosa per sempre (come potrebbero voler dire gli imperativi *pomni* o *zapomni*):

(20a) *Вспомни, подумай. Где твои книги, переводы?* (Gončarov)

(20b) «*Ricordalo, pensaci. Dove sono i tuoi libri, le traduzioni?*»¹⁴⁴.

Aggiungiamo che, talvolta, il traduttore può combinare diverse strategie¹⁴⁵, come in (21b), dove troviamo, nell'ordine, le strategie 3 e 2 (uso del passato remoto e della locuzione avverbiale “*di colpo*”) e in (22b) in cui riscontriamo le strategie 4 e 1 (il tratto

¹⁴³ L'uso di *vspomnit'* in (18a) è legato alla possibilità del russo di omettere un verbo dichiarativo. *Vspomnit'*, come molti verbi russi, possono quindi essere usati al posto di un verbo dichiarativo e introdurre un discorso diretto, cosa che in italiano generalmente non è ammissibile. In questo modo il russo può esprimere in un'unica unità linguistica molti tratti semantici assenti nei verbi dichiarativi italiani. Ciò comporta eventuali aggiunte o omissioni nella traduzione dal russo all'italiano (Sarracco 2012). Cfr.: *vspomnila ona: disse lei, dopo essersi ricordata; smejalsja on: disse lui ridendo.*

¹⁴⁴ In questo contesto, nonostante non si verifichi una perdita di informazione, suona strano l'uso all'imperativo di “*ricordare*”. Il più delle volte, infatti, “*ricordare*” esprime un atto involontario della mente. La volontarietà è più comunemente realizzata con locuzioni di tipo conativo o in contesti circoscritti (*me lo ricorderò oppure ricordare qualcosa a qualcuno*).

¹⁴⁵ Questo spiega perché la somma delle percentuali, includendo anche i casi di perdita di informazione che vedremo nel prossimo paragrafo, va oltre il 100%

conativo, espresso con *fare* causativo riflessivo - dove il causatore coincide con il soggetto -, e l'uso della perifrasi “*tornare a mente*”¹⁴⁶:

(21a) *Во всех отношениях приятная дама **вспомнила**, что выкройка для модного платья еще не находится в ее руках (...)* (Gogol', *Mertvye duši*)

(21b) *La signora amabile da tutti i punti di vista **si ricordò di colpo** che il cartamodello per il vestito alla moda non era ancora in mano sua (...).*

(22a) ***Вспомнил** номер комнаты, где проживали Рябов и Лепко.* (Dovlatov, *Filial*)

(22b) ***Mi ero fatto tornare a mente** il numero di stanza in cui vivevano Rjabov e Lepko.*

Infine, nei casi in cui *vspomnit'* ricorre al futuro, l'italiano talvolta impiega una delle strategie sopra elencate (23b), oppure opta per un semplice “*ricorderò*”, causando talvolta una perdita di significato:

(23a) *Что такое я хотела сказать?.. Ну, все равно, после **вспомню**.* (Gončarov)

(23b) *Cos'era che volevo dirti? ... **Be'**, fa lo stesso, **mi verrà in mente** poi.*

V. VSPOMINAT'

Ci occuperemo ora dell'imperfettivo *vspominat'*, presente 220 volte nel nostro sotto-corpus. Anche in questo caso, per i romanzi molto lunghi sono state analizzate solo le prime 30 occorrenze, riducendo così il numero degli esempi analizzati a 151.

La voce lessicografica di *vspominat'*, sia in Ožegov che in Kovalev, rimanda semplicemente al perfettivo *vspomnit'*, mentre Kovalev aggiunge un esempio, peraltro con un uso fraseologico del verbo:

ВСПОМИНАТЬ: *imperf.* (-ню, -нишь) v. **вспомнить.** *He будем в. прошлого* – non andiamo a rivangare il passato.

¹⁴⁶ I casi di doppia strategia, così come le situazioni di “traduzione-zero” o di ruolo chiarificatore del contesto, sono stati registrati anche per gli altri verbi analizzati, ma data la loro minore frequenza verranno solo menzionati senza presentare esempi specifici.

Se dunque la lessicografia non distingue *vspominat'* da *vspomnit'*, l'analisi dei testi mostra invece come la semantica del verbo sia molto complessa.

Dal punto di vista morfo-sintattico, non si rileva alcun fenomeno particolare di *colligation*: il verbo si comporta in modo simile a *vspomnit'*, e da questo punto di vista sembra di essere davanti a una coppia puramente aspettuale.

Nel nostro corpus *vspominat'* occorre:

- 22 volte all'infinito
- 25 volte al presente
- 64 volte al passato imperfettivo
- 36 volte al gerundio presente
- 4 volte al futuro imperfettivo

Anche in questo caso la prevalenza del passato è il riflesso dello stile narrativo dei testi analizzati. Il numero rilevante di gerundi indica la contemporaneità tra il *ricordare* espresso da *vspominat'* e un'altra azione concomitante.

Esaminando le strategie di resa, sebbene vi siano delle differenze dovute all'aspetto, si nota parimenti un'analogia rispetto a quelle impiegate nella traduzione di *vspomnit'* (perifrasi, locuzioni di tempo, uso dei tempi verbali, significato conativo, altri verbi). Le indichiamo qui di seguito.

1) **Perifrasi** del tipo *venire in mente, tornare alla mente* (12 occorrenze – 7,9%). Analogamente a *vspomnit'*, anche qui l'uso di simili espressioni consente di trasmettere la semantica del prefisso *vs-*, che associato alla radice *pomnit'* dà l'idea di un ricordo che riemerge in maniera improvvisa, ma in questo caso iterata. Presentiamo un esempio:

(24a) *И она опять **вспоминала** все, что она видела.* (Tolstoj)

(24b) *E di nuovo **le tornava in mente** quello che aveva visto.*

Nell'esempio (24b) l'uso dell'espressione di tempo “*di nuovo*” rafforza il senso iterativo del ricordo e ci rimanda alla seconda strategia.

2) **Espressioni di tempo** indicanti sia l'iteratività dell'azione, sia, meno frequentemente, il suo carattere **improvviso** (28 occorrenze – 18,54%): in *vspominat'* il ricordo del soggetto va e viene (*di nuovo, talvolta, spesso, ogni volta che, quando* ecc.) e, a differenza di *pomnit'*, non è costante. In questo gruppo includiamo anche i casi in cui *vspominat'* è tradotto con “*al ricordo (di)*”, espressione temporale di tipo fraseologico (6 occorrenze – 3,97%) (27b):

(25a) (...) *и когда он, как сейчас, ни с того ни с сего падал так, то есть вспомнил эту полусотню только-что вышедших стихотворений, он в один миг мысленно пробежал всю книгу (...)* (Nabokov, *Dar*)

(25b) (...) *e quando, come adesso, faceva simili capitomboli interiori, e cioè ricordava all'improvviso la cinquantina di poesie che avevano appena visto la luce, scorreva mentalmente in un solo attimo l'intero libro (...)*

(26a) *Да, говорят, что и до сих пор критик Латунский бледнеем, вспоминая этот страшный вечер (...)* (Bulgakov)

(26b) *Già, si dice che ancora adesso il critico Latunskij impallidisca al ricordo di quella terribile sera (...)*

In (25a) sono presenti sia la congiunzione *quando*, che rende il russo *kogda* e in questo contesto indica il ripetersi dell'azione, sia la locuzione avverbiale *all'improvviso*, aggiunta dalla traduttrice per sottolineare la repentinità del ricordo.

3) **Tempi verbali** (17 occorrenze – 11,25%). Abbiamo due diverse strategie:

A) Si verificano casi in cui la **consecutio temporum** svolge un ruolo decisivo nel chiarire la semantica del verbo. Ciò avviene quando una proposizione X, il cui soggetto A “*vspominaet*” (VSP), si trova in rapporto subordinativo o coordinativo con una seconda proposizione Y esprime un'azione Z, e le due azioni (Z e VSP) sono simultanee:

Y [Tutte le volte che/mentre accade Z] → X [il soggetto A VSP]

Z può essere una condizione esterna, un'azione compiuta da un altro soggetto o un'azione compiuta da A stesso. In questo caso la strategia dell'italiano consiste nel rispettare il parallelismo temporale del russo attraverso i tempi verbali a sua disposizione, che possono essere:

- gerundio presente + imperfetto (corrispondenti a gerundio pres. /pass. imperf. + pass. imperf. russi) (10 occorrenze) (27b)
- imperfetto + imperfetto (pass. imperf. + pass. imperf.) (5 occorrenze) (28b)
- presente + presente (pres. imperf. + pres. imperf.) (1 occorrenza) (29b)

In queste costruzioni il verbo italiano *ricordare* non determina una perdita di informazione rispetto all'originale russo *vspominat'*.

(27a) *Думая об интрижке, он **вспоминал** про свою непобедимую робость (...)*
(Čechov, *Rasskazy*)

(27b) *Pensando a questo capriccio **si ricordava** della sua timidezza invincibile (...)*

(28a) *Танабай шел, **вспоминал** все связанное с иноходцем за долгие годы (...)*
(Ajtmatov)

(28b) *Танабай camminava, e **ricordava** le cose che in tanti lunghi anni che lo avevano legato a quel cavallo (...)*

(29a) *И что ж? Теперь, когда прошли года, я **вспоминаю** и удивляюсь, как это могло огорчить меня. (Tolstoj)*

(29b) *Ebbene? ... Ora che gli anni sono passati, **ricordo** e stupisco come abbia potuto addolorarmene tanto.*

Si osservano poi casi, nello specifico 8 occorrenze (5,29%), in cui il gerundio presente di *ricordare* traduce il gerundio presente *vspominaja*, ma sia in russo che in italiano è affiancato ad una principale con verbo puntuale (pass. remoto/prossimo in italiano e pass. perfettivo in russo) (30a), (30b):

(30a) *Она порывисто встала из-за стола и ушла к себе, **вспоминая**, как часто отец бывал к ней несправедлив. (Čechov, *Rasskazy*)*

(30b) *Ella si alzò con impeto dalla tavola e se ne andò nella sua camera fortemente adirata, ricordando quanto spesso suo padre fosse ingiusto con lei.*

Anche in questo caso, il russo *vspominaja* indica che vi è contemporaneità tra le azioni di alzarsi e andarsene e quella di ricordare, sebbene le prime due siano puntuali; ma soprattutto, a differenza di *vspomniv*, *vspominaja* non indica in nessun modo che il ricordo è la causa dell'alzarsi e dell'andarsene del soggetto. In ciò consiste la differenza principale tra i due gerundi *vspomniv* e *vspominaja*. Deduciamo quindi che il gerundio presente *ricordando* possa essere spesso l'unica alternativa valida per la resa del gerundio imperfettivo *vspominaja*, ma che non sia sempre appropriato nella traduzione di *vspomniv*. Il modo più esplicito per rendere *vspomniv* è attraverso un gerundio passato o un participio passato, come visto in (14b).

B) Quando *vspominat'* è usato al **presente**, indicando un'azione contemporanea all'atto comunicativo, nella traduzione italiana l'uso del gerundio presente di *ricordare* (30b) non comporta una perdita di informazione, come invece avverrebbe se si usasse un presente indicativo. Il gerundio, infatti, permette di trasformare un verbo stativo come "ricordare" in verbo di attività (secondo la terminologia di Vendler): "sto ricordando" indica l'azione di richiamare alla memoria nel momento in cui si parla. "Io ricordo" indica invece una condizione. Alcuni verbi sembrano appartenere a classi verbali diverse a seconda del tempo verbale in cui ricorrono (cfr. ad es. l'inglese "*I have two dogs*" vs "*I'm having a shower*"). Rimandiamo al par. 3.5. per considerazioni più approfondite sul tema del rapporto fra semantica del verbo e classi verbali. Purtroppo abbiamo registrato solo un caso di questo tipo nel nostro sotto corpus. È infatti più facile trovare *vspominat'* usato in questo modo nel discorso diretto, le cui occorrenze sono minori rispetto alle parti narrative.

(31a) *Ну, что вы, я просто **вспоминаю** всю многогранность этой игры.* (Šalamov)

(31b) *Ma no, che dite, **sto solo ricordando** tutti gli aspetti interessanti di questo gioco.*

4) Uso di **espressioni indicanti il tratto conativo** nel richiamare qualcosa alla mente. Le segnaliamo anche se con questa accezione l'imperfettivo *vspominat'* compare meno frequentemente del perfettivo *vspomnit'* (5 occorrenze – 3,3%). Anche in questo caso è

stato rilevato qualche esempio in cui in russo questa accezione rimane sottintesa mentre nella versione italiana viene esplicitata proprio con lo scopo di trasmettere nel modo più completo possibile la semantica del verbo russo:

(32a) *Алексей Александрович смотрел на Вронского с неудовольствием, рассеянно вспоминая, кто это (...)* (Tolstoj)

(32b) *Aleksej Aleksandrovič guardava Vronskij con disappunto **cercando** distrattamente di ricordarsi chi fosse (...)*

5) Uso di altri verbi o espressioni sinonimiche di ricordare (23,84%). Anche la traduzione di *vspominat'*, oltre al verbo *ricordare*, può avere alcune alternative:

- **gruppo A:** *rammentare* (8), *ripensare* (7), *rievocare* (4), *sovvenirsi* (1), *riandare* (1), *riandare col pensiero* (1), *evocare il ricordo* (1), *il solo pensiero* (1);
- **gruppo B:** *pensare* (6), *ripetere/si* (3), *parlare di* (2)

Inoltre, troviamo alcune forme sporadiche, usate solo una volta come: *conservare un buon ricordo*, *rendersi conto*.

Ancor più che per il verbo *vspomnit'*, si nota qui una vicinanza tra ricordo e pensiero¹⁴⁷, ma se *pensare*, in questi contesti, può determinare una perdita di informazione (33b), *ripensare*, invece, grazie al prefisso *ri-* conserva l'idea del ritorno e quindi si avvicina maggiormente alla semantica di *vspominat'* (34b). In generale, tutte le varianti con questo prefisso sottolineano questo tratto semantico, anche se, come per *ricordare*, vi è sempre la possibilità di una perdita di informazione in mancanza di una delle strategie sopra citate (35b). Infine, vediamo che in alcuni casi *vspominat'* può essere percepito come sinonimo di “ricordare ad alta voce”, “parlare di qualcosa” (36b).

(33a) (...) *он все чаще вспоминал свой героический поступок и размышлял, как бы его приспособить к делу.* (Družnikov)

¹⁴⁷ Ricordiamo infatti che il concetto di memoria si lega soprattutto all'azione del “sapere”, ma è connesso anche all'azione di *pensare*: si confronti l'analogia delle due frasi (esempi tratti da Zaliznjak 2013: 328): “Он редко вспоминал о своей первой жене” e “Он редко думал о своей первой жене”. L'italiano potrebbe infatti tradurre *vspominal* nel primo esempio sia con *ricordare*, sia con *pensare* o *ripensare*: “(Lui) di rado (ri)pensava alla sua prima moglie”.

(33b) (...) *gli capitava di **pensare** sempre più spesso al suo atto eroico e iniziò a meditare su come potesse usarlo per trarne vantaggio.*

(34a) *Эту метку счастья поэт **вспоминал** много раз (...)* (Šalamov)

(34b) *Spesso il poeta **aveva ripensato** a quel marchio di fortuna (...)*

(35a) *Сейчас он **вспоминал** китайца без злобы и без иронии (...)* (Šalamov)

(35b) *Adesso **ripensava** al cinese senza rancore e senza ironia (...)*

(36a) *Мы с ним не то что дружили, а просто любили **вспоминать** Москву (...)*
(Šalamov)

(36b) *Non si può dire che io e lui fossimo amici, semplicemente ci piaceva **parlare** di Mosca (...)* (Trad.: Rapetti)

6) Altri casi: analogamente a *vspomnit'*, si aggiungono, infine, 3 casi in cui il verbo russo non viene tradotto in italiano e 4 esempi in cui, pur senza una strategia di resa particolare, grazie al contesto non si rileva perdita di informazione (37b).

(37a) (...) *и там, идя, спеша, ни на кого не глядя, ничего не замечая, думать, **вспоминать**, соображать, каждое сказанное слово, каждое обстоятельство.*
(Dostoevskij, *Prestuplenie i nakazanie*)

(37b) (...) *poi, camminando svelta senza guardare nessuno e senza vedere nulla, avrebbe pensato, **ricordato**, riflettuto su ogni parola detta, su ogni circostanza.*

С. ЗАПОМНИТЬ/ЗАПОМИНАТЬ

I due verbi sono trattati insieme, dal momento che *zapominat'* occorre solo in due casi, e in uno (38b) non viene tradotto in italiano:

(38a) (...) *она слушала внимательно, **запоминая** новые подробности.* (Družnikov)

(38b) (...) *ascoltò con molta attenzione ogni singolo dettaglio.*

Pertanto, parleremo da qui in avanti solo del perfetto *zapomnit'*, che occorre nel sottocorpus 50 volte. Le forme verbali che assume sono perlopiù due e rivelano poco della semantica del verbo:

- passato (24 occorrenze)
- infinito (13 occorrenze)
- presente (8 occorrenze)
- imperativo (3 occorrenze)
- participio passato (2 occorrenze)

Per quanto riguarda le macro-strategie di resa, si nota che, in un certo senso, ripropongono quanto rilevato per la coppia *vspomnit'-vspominat'* (uso di perifrasi, locuzioni temporali e uso dei tempi verbali).

1) Perifrasi, che qui sottolineano l'idea del ricordo che si fissa e **permane nella memoria** (*rimanere impresso, scolpire nella memoria, tenere a mente, imparare a memoria*) e di cui abbiamo registrato 6 occorrenze (11,7%):

(39a) *Он запомнил улыбку этого человека навсегда (...)* (Šalamov)

(39b) *Il sorriso di quell'uomo gli era rimasto impresso per sempre (...)*

2) Locuzioni temporali quali *per sempre*, oppure *mai* e *più* (in caso di resa negativa, attraverso l'espressione *non dimenticare*): in questo caso esprimono il ricordo come ciò che rimane impresso nella memoria per un periodo di tempo illimitato (5 occorrenze – 9,8%).

(40a) (...) *и бабушка запомнила царя навсегда.* (Erofeev)

(40b) (...) *e la nonna ricordò per sempre lo zar.*

(41a) (...) *он запомнил капитана Шнайдера навек.* (Šalamov)

(41b) (...) *non dimenticò mai il capitano Schneider.*

Nei casi come (41b), con il verbo *dimenticare*, la locuzione temporale (*mai*) è aggiunta dai traduttori, senza che sia esplicitata nell'originale russo. Al russo basta il prefisso per aggiungere questo tratto semantico.

3) Uso del tempo presente per tradurre esempi al passato perfettivo (8 occorrenze – 15,68%). In molti casi tradurre un'espressione come “*ja zapomnil*” con il presente del verbo *ricordare* non comporta una vera e propria perdita di informazione poiché, di fatto, ciò che abbiamo fissato nella nostra mente in passato, si ricorda anche nel presente.

(42a) *Как же-с, претлично его запомнили.* (Akunin)

(42b) «Come no, me lo **ricordo** benissimo.

4) Altri casi: vi sono casi sporadici in cui *zapomnit'* è tradotto con verbi diversi da *ricordare* a seconda del contesto (es.: *riconoscere, sentire, registrare*), ma non trattandosi di un fenomeno regolare nel nostro corpus¹⁴⁸, non se ne terrà conto.

D. ПРИПОМНИТЬ

Del verbo prefissale *pripomnit'* sono state registrate 67 occorrenze. Come è emerso dalle definizioni lessicografiche, si tratta, nel primo significato, di un sinonimo di *vspomnit'* (cfr. par. 3.2.); il secondo significato, per cui però non disponiamo di esempi, ha invece il valore traslato di “non dimenticare il torto subito” e sembrerebbe il modo in cui è maggiormente impiegato nel russo contemporaneo. *Pripomnit'*, infatti, nel suo primo significato, compare più di frequente fino alla fine del XIX secolo. Le strategie di resa in italiano di *pripomnit'* dovrebbero dunque essere simili, in quanto sinonimo, a quelle osservate per *vspomnit'*. L'analisi ha confermato questa ipotesi: in generale, sono state evidenziate le medesime macro-strategie impiegate per la traduzione di *vspomnit'*, su cui, pertanto, non ci dilungheremo, limitandoci a riportare alcuni esempi. Segnaliamo tuttavia che, sulla base degli esempi a nostra disposizione, *pripomnit'* viene usato più spesso per evidenziare la volontarietà dell'azione di ricordare, assumendo quindi significato conativo.

¹⁴⁸ Probabilmente per via del basso numero di occorrenze registrate.

1) Perifrasi: “**venire in mente**” (4 occorrenze – 5,97%):

(43a) (...) *однако ж все, как поразмыслили каждый с своей стороны, как **припомнили**, что они еще не знают, кто таков на самом деле есть Чичиков (...)*
(Gogol', Mervnye duši)

(43b) (...) *e tuttavia, quando ognuno rifletté per conto suo, quando a tutti **venne in mente** che non sapevano ancora chi fosse Cicikov in realtà (...)*

2) Uso di **locuzioni temporali** che indicano l'immediatezza e la repentinità del ricordo (7 occorrenze – 10,44%):

(44a) *Она **припомнила вдруг**, что Раскольников сам хотел к ней сегодня зайти (...)*
(Dostoevskij, Prestuplenie i nakazanie)

(44b) *All'improvviso **ricordò** che Raskòlnikov stesso voleva passare da lei quel giorno (...)*

3) **Passato remoto** (o **passato prossimo**) per indicare la puntualità dell'azione (12 occorrenze – 17,91%) [cfr. anche (43b)]:

(45a) *Он где-то видал эту улыбку; он **припомнил** какую-то картину, на которой изображена женщина с такой улыбкой (...)* (Gončarov)

(45b) *Aveva già visto da qualche parte quel sorriso; **ricordò** un quadro nel quale una donna era raffigurata con quel sorriso (...)*

4) Uso di espressioni indicanti il **tratto conativo** nel richiamare alla mente un ricordo (22 occorrenze – 32,83%):

(46a) *Но как их звали? **Припомни-ка!*** (Čechov, Rasskazy)

(46b) *E come si chiamavano? **Cerca un po' di ricordarti.***»

4) Altri casi: anche per *pripomnit'* abbiamo registrato sporadici casi in cui la traduzione italiana rende il verbo russo per mezzo di altri verbi, alcuni sinonimi di *ricordare* (es.: *rammentare*), altri meno vicini al campo semantico della memoria (es.: *ripercorrere*,

usato in senso figurato “con la mente”). Vista la scarsa frequenza di questi ultimi casi, non li abbiamo distinti dal verbo ricordare e li abbiamo considerati all’interno delle categorie sopracitate, dove opportuno (es.: *sforzandosi di rammentare* – categoria 4), oppure nei casi di perdita di informazione.

Per concludere le riflessioni sulla semantica di questo verbo, segnaliamo un caso in cui *pripomnit’* compare nella versione russa con il significato di *zapomnit’*. Si tratta di un uso molto obsoleto e riscontrato una volta nel nostro sotto-corpus:

(47a) (...) *еще раз окинувши все глазами, как бы с тем, чтобы хорошо **припомнить** положение места, отправился домой прямо в свой номер, поддерживаемый слегка на лестнице трактирным слугою.* (Gogol’, Mertvye duši)

(47b) (...) *poi, gettato un ultimo sguardo all’intorno, come per **imprimersi bene in mente** la disposizione dei luoghi, rientrò dritto nella sua stanza, leggermente sostenuto sulla scala dal cameriere dell’albergo.*

Е. ПРИПОМИНАТЬ

Di *pripominat’*, forma imperfettiva di *pripomnit’*, abbiamo rilevato 46 occorrenze. La voce del dizionario bilingue Kovalev, come per *vspominat’*, rimanda semplicemente a quella del perfettivo. In Ožegov *pripominat’* non compare come voce a sé stante, ma è inserito nella voce *pripomnit’*, dove è indicato come suo imperfettivo, ma solo in relazione al primo significato (*vosstanovit’ v pamjati* – *richiamare alla mente*). Come per *vspominat’*, l'imperfettività si esprime nei valori iterativo o limitativo; e così come *pripomnit’* (rispetto a *vspomnit’* e *vspominat’*) è usato più frequentemente con valore conativo¹⁴⁹ (lo sforzo volontario nel ricordare).

Le strategie di resa in italiano riscontrate negli esempi sono tre, dal momento che nel nostro corpus non si è registrato alcun caso in cui l’uso dei tempi verbali potesse essere considerato come decisivo per la resa della semantica del verbo. Di seguito proponiamo gli esempi più significativi per ogni tipo.

1) **Perifrasi** che indicano il **ritorno alla mente** del ricordo (8 occorrenze – 17,39%):

¹⁴⁹ Più spesso (in percentuale) rispetto agli altri verbi analizzati.

(48a) *Направляясь между толпой в залу первого класса, она помнемногу **припоминала** все подробности своего положения (...)* (Tolstoj)

(48b) *Dirigendosi fra la folla verso la sala d'aspetto di prima classe, ella **riandava con la mente** a tutti i particolari della sua situazione (...)*

(49a) *Он стал **припоминать себе**: кто бы это был (...)* (Gogol', *Mertvye duši*)

(49b) *Cercò **di farsi venire in mente** chi fosse (...)*

In (49b) vediamo anche l'uso di “*cercare di*” e, in secondo luogo *pripominat'* è seguito dal dativo, e sembrerebbe sinonimo di *napomnit'-napominat'*, ossia “ricordare qualcosa a qualcuno”, in questo caso “a sé stesso”.

2) **Locuzioni di tempo** - tra cui anche le espressioni “*al ricordo di*” (già registrata nelle traduzioni di *vspominat'*) o “*nel ricordare*” - che indicano la frequenza del ricordo che va e viene o che accade simultaneamente a un'altra azione o stimolo (11 occorrenze – 23,91%):

(50a) *Впоследствии, когда он **припоминал** это время (...) его до суверия поражало всегда одно обстоятельство (...)* (Dostoevskij, *Prestuplenie i nakazanie*)

(50b) *Più tardi, quando gli capitava di ricordare quel tempo (...) ogni volta lo colmava di stupore superstizioso un fatto (...)*

3) Espressioni che indicano il **tratto conativo** nel ricordare (12 occorrenze – 26%); qui nella maggioranza dei casi, più che per i tre sinonimi precedentemente analizzati, nella traduzione italiana si esplicita questo tratto, come in (51b):

(51a) *Тут перед глазами Римского возник циферблат его часов... Он **припоминал**, где были стрелки... (...)* (Bulgakov)

(51b) *(davanti a Rimskij apparve l'immagine del quadrante del suo orologio). Rimskij **cercava di ricordare** dove fossero le lancette... (...)*

4) **Altri casi:** *pripominat'* in alcuni casi viene reso con verbi, quali *rammentare*, *ripensare*, o per mezzo dell'espressione “*andare ricordando/rammentando*”, di cui riportiamo un esempio in (52b). Quest'ultima compare 4 volte:

(52a) *Полный удовольствия, он слегка **припоминал** все веселые места прведенного вечера (...)* (Gogol', *Šinel'*)

(52b) *D'ottimo umore, egli **andava ricordando** i momenti simpatici della serata appena trascorsa (...)*

3.4.3. Perdita di informazione

Esaminiamo ora i casi in cui l'uso del semplice verbo *ricordare* nella traduzione italiana non consente di cogliere in modo esaustivo la semantica del verbo prefissale russo, comportando così una perdita di informazione¹⁵⁰.

Indichiamo di seguito la frequenza dei casi di questo tipo per ogni singolo verbo preso in considerazione:

- *Vspomnit'*: 43 casi di perdita di informazione (17,13%)
- *Vspominat'*: 45 casi (29,80%)
- *Zapomnit'*- *zapominat'*: 13 casi (26,9%)
- *Pripomnit'*: 30 casi (44,77%)
- *Pripominat'*: 4 casi (8,69%)

Riportiamo alcuni estratti in italiano senza segnalare, per il momento, l'originale di cui sono la traduzione:

(53) *Probabilmente **ricordo** queste minuzie con tanta chiarezza perché pochi minuti più tardi avrei passato attentamente in rassegna le mie impressioni;*

¹⁵⁰ Anche in alcuni degli esempi sopra esposti si verifica una certa perdita di informazione, nonostante essa sia attenuata dall'uso delle strategie compensative che abbiamo evidenziato. In (10b) e (11b), per esempio, nonostante sia stata adottata la strategia n. 2 (locuzioni temporali) per tradurre *vspomnit'* (*ora ricordo, adesso ricordo*) ed esprimere un pensiero o un evento che viene ricordato in un preciso momento (dopo essere stato dimenticato), il fatto che nella traduzione si utilizzi il presente indicativo ci dà una resa italiana equivalente a quella che si userebbe per l'espressione “*sejčas pomnju*”, in cui il verbo non è prefissato.

(54) Nessuno **ricordava** più quando fosse entrato in quel ministero e chi lo avesse assunto.

(55) Neanche **si ricordavano** più chi ce le avesse messe.

In nessuno di questi casi emergono i tratti semantici specifici dei verbi prefissali presenti nell'originale russo. Sembrerebbe infatti che in tutti questi esempi il verbo italiano *ricordare* corrisponda all'imperfettivo *pomnit'* senza prefisso. In realtà, non è così, giacché (53) traduce (53b), (54) corrisponde a (54b) e (55) a (55b):

(53b) *Я **припоминаю** так ясно эти мелкие подробности потому, вероятно, что мне пришлось так основательно проверить свои впечатления несколько мгновений спустя (...)* (Nabokov, *Lolita*)

(54b) *Когда и в какое время он поступил в департамент и кто определил его, этого никто не мог **припомнить**.* (Gogol', *Šinel'*)

(55b) *Кто поставил, так потом и не **вспомнили**.* (Ulickaja)

Notiamo che in (54) il tratto semantico specifico di *pripomnit'* - la volontarietà dell'azione, il significato conativo - è stato omesso, sebbene nell'originale (54b) la presenza del modale *ne mog* (*non riusciva*) presupponga chiaramente la volontà e lo sforzo per riuscire. Dal momento che de *Il Cappotto* di Gogol' si dispone di 15 versioni, abbiamo avuto modo di controllare la resa di questa frase in tutte le traduzioni, constatando che 8 su 15, esattamente come in (54), non rendono questa sfumatura.

Visto il numero abbastanza elevato di casi di perdita di informazione, riportiamo di seguito altri esempi. In (56b) *si ricordavano* sembrerebbe tradurre *pomnili* perdendo l'idea del ricordo iterato e limitato nel tempo presente nell'originale.

(56a) ***Вспоминали** его смех, его лицо (...)* (Gorkij, *Suprugy Orlov*)

(56b) ***Si ricordavano** del suono della sua voce, dell'intonazione del suo riso, della conformazione della sua persona (...)*

In (57b) e (58b) per rendere la semantica specifica di *pripominat'/pripomnit'* sarebbe necessario esplicitare il tratto conativo, ossia l'atto volontario attuato dal soggetto per richiamare alla mente quanto vuole ricordare.

(57a) *Она называла их и **припоминала** не только имена, но года, месяцы, характеры, болезни всех детей (...)* (Tolstoj)

(57b) *Nominava e **ricordava** non soltanto i nomi, ma gli anni, i mesi, i caratteri, le malattie di tutti loro (...)*

(58a) *Довольно трудно ответить на этот вопрос, я не **припомню** в точности, что носил Павел Петрович.* (Sokolov)

(58b) *È un po' difficile rispondere a questa domanda, io non mi ricordo con precisione che cosa portasse indosso Pavel Petrovič.*

In (59b) il contesto può essere parzialmente d'aiuto, ma senza testo russo a fronte non è del tutto chiaro che *ricordarsi* qui significa “fissare nella memoria tutto ciò che si osserva”; una perifrasi come quelle descritte nella strategia n. 1 (es: *fissare nella memoria, tenere a mente*) sarebbe stata più esplicita:

(59a) *У мельника от его дум явилась какая-то особая, незнакомая ему раньше, способность всё подмечать и **запоминать** и ко всему прилагать вопрос (...)* (Gorkij, *Toska*)

(59b) *Una facoltà del tutto speciale, e sconosciuta fin'allora al mugnaio, si è sviluppata in lui da quei pensieri; una facoltà di osservare tutto e di **ricordarsi**, e di rispondere ad ogni questione.*

Anche in (60b) si tralascia l'idea del ricordo fissato nella memoria a partire da un preciso momento e l'espressione *per quanto ricordava* potrebbe essere la traduzione di *skol'ko mogla pomnit'/pripomnit'/vspomnit'*.

(60a) *Правда, сколько она **могла запомнить** свое впечатление в Петербурге у Кареных, ей не нравился самый дом их;* (Tolstoj)

(60b) *Però, per quanto ricordava, l'impressione da lei riportata a Pietroburgo dei Karenin, non era stata favorevole (...)*

In (61b) *ti ricordi* può equivalere a *pomniš'*. Qui per rendere la semantica del verbo russo sarebbe stata utile un'espressione di tempo, quale *adesso*, ad indicare che il ricordo non è costante, ma giunge in un preciso momento.

(61a) – *Вспомнил ты меня, краснопогонник?! (Dovlatov)*

(61b) – *Ti ricordi di me, questurino?!*

3.5. Osservazioni conclusive

L'analisi degli esempi ci ha permesso di rintracciare le sfumature di significato assunte di volta in volta dai diversi verbi prefissali con base *pomnit'*. Proponiamo quindi una tabella riassuntiva che consenta di confrontare i tratti semantici evidenziati per ogni verbo prefissale, le strategie di resa in italiano e la loro distribuzione in percentuale.

Tab 15. semantica dei verbi prefissali di memoria, strategia di resa e frequenza

Verbo	Tratti semantici rilevati	Strat. 1: perifrasi	Strat. 2: uso di locuzioni temporali	Strat.3: uso dei tempi verbali	Strat. 4: esplicitazio -ne del tratto conativo	Strat. 5: sostituzione con sinonimi di ricordare o con altri verbi	Perdita di informaz.
VSPOMNIT'	1) puntualità dell'azione (ricordo improvviso)	8,30%	11,50%	48,20%	5,90%	16,70%	17,13%
	2) significato conativo						
	3) ricordo involontario che ritorna dopo un periodo di assenza e generato da uno stimolo esterno						
VSPOMINAT'	1) ricordo iterato e limitato nel tempo	7,90%	18,54%	11,25%	3,30%	23,84%	29,80%
	2) significato conativo						

ZAPOMNIT' / ZAPOMINAT'	1) informazione che si fissa nella memoria	11,70%	9,80%	15,68%	/	/	26,90%
	2) ricordo che si fissa a partire da un determinato momento e permane per sempre						
PRIPOMNIT'	1) puntualità dell'azione (ricordo improvviso)	5,97%	10,44%	17,91%	32,83%	/	44,77%
	2) significato conativo						
	3) ricordo involontario che ritorna dopo un periodo di assenza e generato da uno stimolo esterno						
	4) uso obsoleto equivalente a <i>zapomnit'</i> : imprimere un'informazione e nella memoria						
PRIPOMINAT'	1) ricordo iterato	17,39%	23,91%	/	26%	/	8,69%
	2) significato conativo						
	3) Uso più raro come sinonimo di <i>napominat'</i> : far ricordare qualcosa a qualcuno (48a)						

Per sintetizzare i risultati della nostra ricerca può essere utile riprendere la nozione vendleriana di classe verbale¹⁵¹. Infatti, il verbo italiano *ricordare* può essere interpretato

¹⁵¹ In *Linguistics in philosophy* (1967: 97-121) Vendler suddivide i verbi in quattro classi: di azione, stativi (non ammettono la forma progressiva), e all'interno dei verbi di azione distingue quelli di compimento,

come verbo stativo (*pomnit'*), d'attività (*vspominat'*, *pripominat'* e *zapominat'*) oppure culminativo (*vspomnit'* o *pripomnit'*), e in base all'interpretazione (suggerita dal contesto) corrisponderà a diversi derivati prefissali di *pomnit'*. Se si considerano *pomnit'* e i suoi derivati prefissali, si nota che i prefissi *vs-*, *pri-* e *za-* permettono allo stativo *pomnit'* di diventare un verbo culminativo (*vspomnit'*, *pripomnit'*), mentre l'aspetto marca il successivo passaggio da verbi culminativi a verbi di attività (*vspominat'*, *pripominat'*), quando si passa da perfetto a imperfetto secondario. Nell'italiano, invece, alcuni esempi mostrano che la differenza tra l'interpretazione di *ricordare* come stativo o come verbo d'azione non è marcata morfologicamente tramite prefissi, ma dipende dall'aggiunta di locuzioni che esprimano la volontarietà dell'azione [es: *cercare di ricordare*, cfr. (46b)], dall'uso di espressioni considerate sinonimiche in determinati contesti (cfr.: “pensarci”, “memorizzare” ecc.), o dall'uso del tempo e del modo verbale: per esempio, il gerundio permette di trasformare un verbo stativo come *ricordare* in verbo di attività (cfr. *io ricordo vs sto ricordando*). Questa seconda opzione è certamente più rara, ma possibile in alcuni contesti [cfr. esempio (31b)]¹⁵².

In conclusione, la nostra analisi *corpus-based* ha messo in luce le sottili differenze semantiche fra i vari verbi prefissali di memoria russi, trattati spesso come quasi-sinonimi dagli strumenti lessicografici. Il fatto che il nostro sotto-corpus, di soli 3 milioni di parole, abbia permesso di rilevare - oltre alla relazione fra prefissi e classi di azione - alcune sfumature di significato ignorate dai dizionari, in particolare l'uso conativo di *vspomnit'*-*vspominat'* e *pripomnit'*-*pripominat'* dimostra l'importanza di questo tratto semantico e la necessità di includerlo nelle definizioni lessicografiche e, eventualmente, nelle grammatiche specializzate nella descrizione dei prefissi.

Per quanto riguarda invece il confronto tra russo e italiano, abbiamo visto che, la differenza tipologica fra queste due lingue (l'una *satellite-framed* e l'altra *verb-framed*) non permette all'italiano di esprimere in un solo verbo tutti i tratti semantici espressi da un verbo prefissale russo, tanto più quando il verbo russo presenta numerosi derivati prefissali formati dalla medesima radice, come è il caso dei verbi di memoria. Per questo motivo anche le strategie compensative utilizzate dai traduttori dal russo in italiano non

rivolti a un fine, da quelli di culminazione, il cui fine si realizza in modo istantaneo (in inglese: *stative*, *activity*, *accomplishment* e *achievement*). La classificazione di Vendler ha subito critiche e correzioni durante i decenni successivi. In particolare, i criteri di appartenenza alle classi da lui identificati risultano in realtà variabili in base al contesto e al modello di reggenza del verbo (Bonomi, Zucchi 2001: 145-150).

¹⁵² Ma no, che dite, **sto solo ricordando** tutti gli aspetti interessanti di questo gioco.

sempre riescono ad evitare perdite di informazione nel passaggio dal testo fonte russo a quello di arrivo.

Bibliografia

- Alisova T. B., Repina T. A., Tariverdieva M. A., (1987). *Vvedenie v romanskuju filologiju*. Moskva: Vysšaja škola.
- Apresjan Ju. D. (1995b). *Obraz čeloveka po dannym jazyka: popytka sistemnogo opisanija*. In *Voprosy Jazykoznanija*, 1995, № 1. Moskva: Nauka, pp. 37-67.
- Barychina A. N., Dobrovolskaja V. V., Merzon S. N., (1989). *Izučenie glagolnych pristavok*. Moskva: Russkij Jazyk.
- Bonomi A., Zucchi A. (2001). *Tempo e linguaggio. Introduzione alla semantica del tempo e dell'aspetto verbale*. Milano: Paravia Bruno Mondadori.
- Cordin P. (2011). *From verbal prefixes to direction/result markers in Romance*. In *Linguistica*, 51, 1, pp. 201-216.
- Dobrušina E. P. (2010). *V zaščitu edinoj abstraktnoj semantiki glagol'nych pristavok*. In *Vestnik PSTGU: Filologija*, 4, 22, pp. 99-102.
- Evgeneva A. P. (1957-1984). *Malyj Akademičeskij slovar'*. Moskva: Institut russkogo jazyka Akademii nauk SSSR.
- Fillmore C. (1992). *Corpus Linguistics or Computer-Aided Armchair Linguistics*. In Svartvik (ed.), *Directions in Corpus Linguistics*. Berlin: Mouton de Gruyter, pp. 35-60.
- Gjervold J. (2013). *Prefix variation: a comparison between slang and contemporary standard Russian*. In *Poljarnyj Vestnik*, 16, pp. 33-46.
- Jakunina D. V. *Pristavka pri-: postroenie semantičeskoj seti*. In Krongauz M. A. (ed) (2001). *Glagolnye prefiksy i prefiksalye glagoly*. Mosckovskij Lingvističeskij Žurnal, 5, 1, RGGU: Moskva, pp. 125-160.
- Janda L. (2007). *Aspectual clusters of Russian verbs*. In *Studies in language*, XXXI, 3, pp. 607-648.
- Janda L. (2012). *Russkie pristavki kak sistema glagol'nych klassifikatorov*. In *Voprosy Jazykoznanija*, 2012, 6. Moskva: Nauka, pp. 3-47.
- Janda L., Endresen A., Kuznetsova J., Ljashevskaja O., Makarova A., Nessel T., Sokolova S. (2013). *Why Russian aspectual prefixes aren't empty. Prefixes as verb classifiers*. Indiana University: Slavica.
- Kovalev V. (2000). *Il Kovalev. Dizionario italiano-russo, italiano-russo*. Bologna: Zanichelli.

- Krongauz M. A. (1994). *Pristavki i glagoly: grammatika sočetaemosti*. In *Semiotika i informatika*, 34, Moskva, pp. 32-57.
- Krongauz M. A. (1998). *Pristavki i glagoly v rusckom jazyke: semantičeskaja grammatika*. Škola “Jazyki rusckoj kul’туры”: Moskva.
- Krongauz M. A. (ed) (2001). *Glagolnye prefiksy i prefiksals’nye glagoly*. Mosckovskij Lingvističeskij Žurnal, 5, 1, RGGU: Moskva.
- Kuznecova Ju. L., Janda L. (2013). *Pristavki v svete kognitivnoj lingvistiki i tipologii (otlik na stat’ju A.A. Zaliznjak i I. L. Mikaeljan)*. In *Voprosy Jazykoznanija*, 2013, 4. Moskva: Nauka, pp. 87-96.
- Martinowsky G. (1998). *Glagol’naja pristavka i vid v rusckom jazyke*. In *Revue des études slaves*, 70, 1. Communications de la délégation française au XIIe Congrès international des slavistes (Cracovie, 27 août - 2 septembre 1998). pp. 79-84.
- Masini F. (2008). *Verbi sintagmatici e ordine delle parole*. In Cini M. (a cura di.), *I verbi sintagmatici in italiano e nelle varietà dialettali. Stato dell’arte e prospettive di ricerca. Atti delle giornate di studio, Torino, 19-20 febbraio 2007*, Frankfurt: am Main, Peter Lang, pp. 83-102.
- Masini F. (2012). *Parole sintagmatiche in italiano*. Roma: Caissa Italia.
- Meyer C. F. (2002). *English corpus linguistics. An introduction*. Cambridge University Press: Cambridge.
- Neset T., Janda L., Kuznetsova J., Lyashevskaya O., Makarova A., Sokolova S. (2008). *Why послушать, but услышать?* In *Poljarnyj Vestnik*, 11, pp. 39-46.
- Neset T., Baydimirova A., Janda L. (2009). *Getting out: A preliminary analysis of вы- and из-*. In *Poljarnyj Vestnik*, 12, pp. 72-84.
- Ožegov S. I., Švedova N. Ju. (a cura di) (1992). *Tolkovyj Slovar’ Ožegova*. Moskva: Russkij jazyk.
- Partington A. (2004). *“Utterly content in each other’s company”: semantic prosody and semantic preference*. In *International Journal of Corpus Linguistics* 9, 1, pp. 131-156.
- Pozolotina T. (2009). *Emptiness is not so empty: The meaning of the Russian prefix вы- in the natural perfective verbs*. In *Poljarnyj Vestnik*, 12, pp. 85-105.
- Pozolotina T. (2010). *Are prefixes always or never “empty”?* In *Poljarnyj Vestnik*, 13, pp. 49-62.

- Rebrina L. N. (2008). *Glagoly pamjati nemeckogo jazyka. Paradigmatičeskie otnošenija edinic*. In Aktual'nye voprosy sovremennoj nauki: sbornik naučnych trudov. Novosibirsk: Centr razvitija naučnog sotrudničestva, pp. 281-292.
- Sarracco R. (2012). Metafonologia de *I demoni* di F.M. Dostoevskij: analisi e problemi di traduzione. In "L'Analisi Linguistica e Letteraria", Anno XX, 1/2012, Facoltà di Scienze Linguistiche e Letterature Straniere, Università Cattolica del Sacro Cuore, Milano, pp. 53-69.
- Svenonius P. (2004). *Slavic Prefixes and Morphology: an introduction to the Nordlyd volume*. Nordlyd 32, 2, CASTL, Tromsø, pp. 177-204.
- Talmy L. (2001). *Towards a cognitive semantics. Typology and Process in Concept Structuring*. Vol II. Cambridge, London: MIT Press.
- Tatevosov S. G. (2009). *Množestvennaja prefiksacija i anatomija russkogo glagola*. In Kiseleva K. L., Plungjan V. A., Rachilina E., Tatevosov S. G. (a cura di). Moskva: Probel-2000, pp. 92-156.
- Thelin N. B. (1984). *K tipologii glagol'noj prefiksacij i ee semantiki v slavjanskich jazykach*. In Revue des études slaves, 56, 2. *Grammaire et lexique*. pp. 225-238.
- Ušakov D. N. (a cura di) (1935-1940). *Tolkovyj slovar' russkogo jazyka*. V 4 tomach. Moskva: Gosudarstvennyj institut "sovetskaja enciklopedija; OGIZ (t.1); Gosudarstvennoe izdatel'stvo inostrannyh i nacional'nych slovarej (t. 2-4).
- Vasil'ev L. M. (1981). *Semantika russkogo glagola*. Moskva: Vysšaja škola.
- Vendler Z. (1967). *Linguistics in philosophy*. London: Cornell University Press.
- Vinogradov V. V. (2001). *Russkij jazyk (grammatičeskoe učenie o slove)*. Zolotova G. A. (ed), 4-e izd. Moskva: Russkij Jazyk.
- Volochina G. A., Popova Z. D. (1997) *Kategorija glagol'nogo vida v svete semantičeskogo ustrojstva glagol'nych pristavok*. In Trudy aspekto-logičeskogo seminaru filologičeskogo fakul'teta MGU im. M. V. Lomonosova. Moskva, pp. 34-41.
- Zaliznjak A. A. (2006). *Mnogoznačnosť v jazyke i sposoby ee predstavlenija*. Moskva: Jazyki slavjanskich kul'tur.
- Zaliznjak A. A. (2013). *Russkaja semantika v tipologičeskoi perspektive*. Moskva: Jazyki slavjanskich kul'tur.

Zalznjak A. A., Mikaeljan I. L. (2012). *O nekotorych diskussionykh momentakh aspektologičeskoj koncepcii Lory Jandy*. In *Voprocjazykoznanija*, 2012, 6, Moskva: Nauka, pp. 48-65.

CAPITOLO 4

IL CAUSATIVO ANALITICO ITALIANO “FARE + INFINITO”: ANALISI E CONFRONTO CON IL CAUSATIVO RUSSO

Oggetto di questo capitolo è il confronto *corpus-based* tra la costruzione causativa *fare + infinito* e il causativo russo.

Il verbo *fare* occupa una posizione di rilievo nel sistema verbale dell’italiano, tanto che, in riferimento a questo verbo, Rylov (2003) parla di “dominante semantica”. In effetti, secondo diversi studi sul lessico di frequenza (cfr. Bortolini et al. 1972), *fare* occuperebbe il sedicesimo posto tra le parole più usate¹⁵³.

Il verbo *fare* è trattato dai dizionari come lessema dalla ricca polisemia: il *Grande Dizionario della Lingua Italiana* di Salvatore Battaglia, nelle 21 pagine dedicate a questo lemma cita 63 significati. Eppure, secondo alcuni linguisti, sembra più lecito pensare a *fare* come ad un verbo non tanto polisemico, quanto polifunzionale, dal momento che, tolto il suo significato intrinseco, detto ipsivalente, ossia “creare”, “produrre”, “fabbricare” (Consales 2006: 80, Ponchon 1994: 11-28), nel resto dei casi interviene nel testo assumendo sì diversi tratti semantici, ma soprattutto diverse funzioni grammaticali. La prospettiva funzionale è adottata ad esempio da La Fauci e Mirto (2003: 101), secondo i quali, dal momento che *fare* può significare tutto, di fatto non significa nulla, e il tentativo di spiegarlo esclusivamente dal punto di vista semantico può rivelarsi infruttuoso. I due studiosi addirittura affermano che “è la sintassi interamente responsabile di *fare*” (La Fauci e Mirto 2003: 101). Pertanto, trattando il verbo *fare* parleremo soprattutto delle sue funzioni, piuttosto che di significati, prestando particolare attenzione a quello che Jurij Apresjan (1967: 5-7) definisce il “significato sintattico” del verbo¹⁵⁴. Come vedremo, una delle funzioni dove la sintassi di *fare* assume un’importanza preponderante è proprio quella definita “causativa”.

¹⁵³ Bartolini et al. (1972: 10-18) si basano su testi appartenenti a cinque campioni scelti (teatro, romanzo, cinema, periodici, sussidiari delle elementari), di 100.000 parole ciascuno e redatti a partire dalla fine del secondo conflitto mondiale (in particolare tra il 1947 e il 1968). Nell’elenco precedono *fare* le seguenti parole: *il, di, egli, a, essere (verbo), e, uno, in, che, non, essere (ausiliare), io, avere (ausiliare) da, che* (1972: 718-19)

¹⁵⁴ Apresjan parla di almeno quattro tipi di significati: denotativo, significativo, sintattico e pragmatico.

Il primo paragrafo (4.1.) presenterà una rassegna delle principali funzioni del verbo *fare* in italiano, per poi passare, dal par. 4.2. alla funzione causativa: in particolare si definiranno i concetti di causatività e verbo causativo. Nel par. 4.3. esporremo le differenze tra tradizione italiana e tradizione russa nel trattare questo tema. I paragrafi 4.4. e 4.5. sono dedicati all'analisi contrastiva tra la costruzione causativa italiana con *fare* e la sua resa in lingua russa. Infine, nel par. 4.6. verranno esposte alcune considerazioni conclusive.

Di seguito presentiamo, nell'ordine cronologico decrescente proposto dal NKRJa, la lista di testi ed autori che compongono il sotto-corpus da noi utilizzato per l'analisi, affiancato da titolo russo, traduttore e anno di traduzione. I testi sono 21, per un totale di 725.686 parole. Anche in questo caso vi sono sia estratti, sia testi interi (cfr. cap. 2):

1. Giorgio Faletti. *Io uccido* (2002) - *Ja ubivau*. I. Kostantinova (2005).
2. Niccolò Ammaniti. *Io non ho paura* (2001) – *Ja ne bojus'*. V. Nikolaev (2005).
3. Valerio Massimo Manfredi. *Aléxandros, III, il confine del mondo* (1998) - *Aleksandr makedonskij. Predely mira*. M. Kononov (2003).
4. Andrea Camilleri. *Il Cane di Terracotta* (1996) - *Sobaka iz terrakotty*. A. Kondjurina (2005).
5. Alessandro Baricco. *Novecento* (1994) - *Legenda o pianiste*. N. Kolesova (2005).
6. Umberto Eco. *Il pendolo di Foucault* (1988) - *Majatnik Fuko*. E. Kostjukovič (2006).
7. Umberto Eco. *Il nome della rosa* (1980) - *Imja rozy*. E. Kostjukovič (1989).
8. Giorgio Bassani. *Il giardino dei Finzi-Contini* (1962) - *Sad Finci-Kontini*. I. Soboleva (2008).
9. Italo Calvino. *Il barone rampante* (1957) - *Baron na dereve*. L. Veršinin (1965).
10. Dino Buzzati. *Eppure battono alla porta* (1942) - *I vse že stučat v dver'*. I. Smagin (2010).
11. Dino Buzzati. *I sette messaggeri* (1942) - *Sem' goncov*. F. Dvin (2010).
12. Dino Buzzati. *Il borghese stregato* (1942) - *Očarovannyj buržua*. P. Chlodovskij (2010).
13. Dino Buzzati. *Il mantello* (1942) – *Plašč*. G. Kiselev (2010).

14. Dino Buzzati. *L'assalto al Grande Convoglio* (1942) - *Napadenie na bol'shoj konvoj*. P. Chlodovskij (2010).
15. Dino Buzzati. *L'uccisione del drago* (1942) - *Kak ubili drakona*. F. Dvin (2010).
16. Dino Buzzati. *Sette piani* (1942) - *Sem' etažej*. G. Kiselev (2010).
17. Luigi Pirandello. *La signora Frola e il signor Ponza, suo genero* (1917) - *Gospoža Frola i ee zjat' gospodin Ponca*. M. Gasanova-Mijat (2006).
18. Luigi Pirandello. *Il Fu Mattia Pascal* (1904) - *Pokojnyj Mattia Paskal'*. G. Rubcova, N. Rykova (1967).
19. Edmondo De Amicis. *Cuore* (1886) – *Serdce*. V. Davidenkova (1958).
20. Carlo Collodi. *Pinocchio* (1883) - *Priključenija Pinokkio*. E. Kazakevič (1959).
21. Carlo Goldoni. *La locandiera* (1753) – *Traktirščica*. A. Dživelegov (1933).

Come per il terzo capitolo, d'ora in poi tutti gli esempi, ove non indicato altrimenti, saranno tratti da questi brani e ogni esempio originale italiano sarà accompagnato dall'indicazione dell'autore tra parentesi. Riporteremo anche il titolo dell'opera laddove ad un singolo autore ne corrisponda più di una (nei casi di Buzzati, Pirandello ed Eco).

4.1. Classificazione funzionale del verbo *fare*

Per una maggiore comprensione del tema, presentiamo innanzitutto il verbo *fare* (d'ora in poi F) in tutte le sue principali funzioni. La rassegna dei lavori dedicati a F, in prospettiva diacronica e sincronica¹⁵⁵, ci ha permesso di identificare tre funzioni principali e altre funzioni secondarie, i cui tratti e confini rimangono ancora non del tutto definiti. In tutti questi casi, come accennato, F si distanzia dal suo significato ipsivalente.

Grazie al corpus è stato possibile attingere a numerosi esempi per ogni categoria. La ricerca degli esempi sul CP ha avuto un duplice scopo: identificare le funzioni più frequenti e, grazie alle traduzioni allineate, confrontare il comportamento del russo e dell'italiano in ciascuno dei casi. La scelta di concentrarsi proprio sulla costruzione

¹⁵⁵ Cfr. Bertinetto (1991: 147), La Fauci (2009), La Fauci, Mirto (2003), Mirto (2003) e Salvi (1988: 81-82) per un approfondimento generale sul verbo F; Brambilla Ageno (1964: 468-472), Cerbasi (1998), Consales (2006), Pelo, Consales (2003), Robustelli (1993, 1994, 1995), Simone, Cerbasi (2001) per la diacronia e, più in particolare per lo studio diacronico di F causativo; Benigni, Cotta Ramusino (2011) per F come verbo di supporto. Inoltre, cfr. Bajric (2008), Gross (1968), Ponchon (1994) per degli studi sul francese *faire*, applicabili all'analogo F italiano.

causativa è dipesa non solo dalla rilevanza del tema, ma anche dal suo elevatissimo riscontro negli esempi raccolti.

Durante questa prima fase di ricerca sono state estratte 3007 occorrenze del verbo F, che dopo l'eliminazione manuale della sinonimia e delle forme non pertinenti si sono ridotte a 2576. Oltre a casi di omonimia [*fa*] nel senso di “*prima di questo momento*” come nell’espressione “*sei anni fa*” (25 occorrenze); *faccia*” come sostantivo femminile sinonimo di “*viso*” (139 occorrenze); *fatto*” come sostantivo maschile, sinonimo di “*avvenimento*” (65 occorrenze)], non sono state prese in considerazione la forma riflessiva “*farsi*” (184 occorrenze) e le forme pronominali “*farcela*” (9 occorrenze) e “*farsene*” (9 occorrenze), poiché usate esclusivamente in locuzioni fisse con significati particolari. Inoltre, data la mole dei dati, sono stati analizzati non più di 50 esempi per ogni estratto. In totale, questa prima analisi ha coinvolto 630 occorrenze di F in tutte le restanti forme.

Passiamo quindi ad una breve descrizione delle funzioni di F (ad eccezione del causativo, ripreso in seguito), seguite dai relativi esempi tratti dal corpus.

A) F vicario¹⁵⁶: F sostituisce un verbo già espresso in precedenza nel co-testo [si parla in questo caso di “vicario anaforico” (1a)], oppure un verbo che seguirà [“vicario cataforico” (2a)]. Talvolta la sostituzione non interessa un singolo verbo, ma un intero avvenimento o una serie di azioni (3a):

(1a) *Le ho slacciato la scarpa da ginnastica e l'ho sfilata con molta attenzione. Come avrebbe fatto un dottore.* (Ammaniti)

(2b) *Я расшнуровал ей башмак, осторожно извлек ногу. Как это сделал бы врач.*

(2a) *E voi, dottor Martora, che cosa fate, dormite?* (Buzzati, *Eppure battono alla porta*)

(2b) *Доктор Мартора, вы что, спите?*

(3a) – *Abbiamo fatto di tutto, generale, per ritrovare il corpo del tuo ragazzo.* (Manfredi)

(3b) – *Мы сделали все, чтобы отыскать тело твоего сына.*

¹⁵⁶ A volte definito come «*pro-verbo*», per analogia con il termine «*pronome*» [Pelo, Consales 2003: 46]. Per un approfondimento su F vicario cfr. (Pelo, Consales 2003: 45-50, Ponchon 1994: 251-341).

Più della metà delle volte (il 67,2% dei casi) F vicario è tradotto in russo con l'analogo "delat'/sdelat'" (*fare*). Quando ciò non avviene, i motivi sono prevalentemente due: il verbo viene omissso, come in (2b), oppure il traduttore opta per una variante più specifica. Anche in russo, dunque, "delat'/sdelat'" svolge la funzione di vicario, sebbene con minor frequenza.

B) F come verbo di supporto: le costruzioni con verbo di supporto (CVS) sono frequenti non solo in italiano. Ciò che però distingue l'italiano dalle altre lingue è che esse vengono realizzate prevalentemente con il verbo F (Cantarini 2004: 68). Seppure largamente studiate, i linguisti non hanno ancora raggiunto un accordo sul loro trattamento. Oltre alle discussioni di carattere terminologico¹⁵⁷, sembra mancare una chiara descrizione delle loro caratteristiche: spesso è difficile distinguere tra costruzioni con verbo di supporto vere e proprie e locuzioni apparentemente analoghe, ma non riconducibili a questo gruppo.

Le CVS sono formate dal verbo, totalmente o parzialmente desemantizzato, e dal suo complemento diretto, che funge da predicato e da centro semantico della frase (Benigni, Cotta Ramusino 2011: 9-15). A differenza di alcune scuole (cfr. Cantarini 2004: 111), Benigni e Cotta Ramusino includono nel gruppo delle CVS anche espressioni fisse o idiomatiche. Queste ultime spesso non sono considerate CVS in quanto non ammettono variazioni sull'asse sintagmatico (per esempio, la sostituzione del complemento con un pronome o il rinvio anaforico per mezzo di pronome relativo)¹⁵⁸. Le CVS andrebbero distinte anche da costruzioni sintatticamente analoghe, ma in cui F mantiene in qualche misura il suo significato primario e, pertanto, non può essere considerato di supporto¹⁵⁹. Qui abbiamo tuttavia assunto la definizione più restrittiva di CVS, in quanto il suo

¹⁵⁷ La terminologia legata ai verbi di supporto non è omogenea, come notano Benigni e Cotta-Ramusino (2011: 8). Cantarini (2004: 67), ad esempio, parla anche di "verbi funzionali", termine che richiama le funzioni lessicali di Mel'čuk (1999). In italianistica prevale la denominazione da noi adottata "verbo supporto" (ingl. *support verb*). In russistica, invece, nel corso degli anni sono state proposte varie definizioni: *glagoly podderžki* (verbi di supporto), *funkcional'nye glagoly* (verbi funzionali), *poluspomogatel'nye glagoly* (verbi semi-ausiliari), *glagoly-operatory* (verbi operatori) (Benigni, Cotta Ramusino 2011: 8).

¹⁵⁸ Allo stesso modo, Benigni e Cotta Ramusino adducono ragioni per considerare CVS anche gli esempi del gruppo "F-lavoro", nonostante vi siano delle differenze sintattiche tra questi due tipi di costruzioni (prima fra tutte l'intransitività di F nelle costruzioni F-lavoro).

¹⁵⁹ Benigni e Cotta Ramusino (2011) propongono alcuni criteri per il riconoscimento dei diversi casi (rimandiamo pertanto al saggio citato). Inoltre, per ulteriori approfondimenti sul tema dei verbi di supporto, si raccomandano (Cantarini 2004 e D'Agostino, Guglielmo 2012).

ampliamento ci obbligherebbe ad accettare forme che sono ancora oggetto di discussione e statisticamente non rilevanti rispetto al nostro corpus.

Nel nostro sotto-corpus sono state trovate 170 CVS. Gli esempi sono così eterogenei, che è impossibile in poche righe fornire un quadro esaustivo. Pertanto presenteremo solo alcuni esempi, sottolineando che molto raramente il russo traduce F con l'analogo “*delat*”. Più spesso il senso di una CVS italiana viene reso in russo tramite un singolo verbo, come in (4b); la variante sintetica è generalmente preferita nella lingua russa.

(4a) [...] *e quando lo guardo, mi fa un sorriso con gli occhi socchiusi come per dirmi: - Ebbene, Enrico, siamo amici?* (De Amicis)

(4b) *Когда я смотрю на него, он улыбается мне и щурится так, как будто хочет сказать: «Ведь мы с тобой друзья, Энрико, правда?».*

Talvolta anche il russo presenta una CVS, realizzata però con un verbo differente da *delat*:

(5a) *Fatemi il santo piacere di lasciarla dormire.* (Buzzati, *Eppure battono alla porta*)

(5b) *Окажите мне такую любезность, дайте ей поспать.*

Benigni e Cotta Ramusino (2011) mostrano come anche “*delat*” possa avere una funzione di supporto, sebbene i casi coincidano raramente con l'uso di F italiano. Nel corpus principale del NKRJa (che conta al momento 265.401.717 parole) l'espressione “*delat' ulybku*” (*fare un sorriso*) occorre solamente 7 volte. Nel CP italiano-russo, dove le parole in lingua italiana sono circa 2 milioni, lo stesso sintagma, nei testi italiani, occorre 21 volte. Normalizzando i dati¹⁶⁰, si ottiene che nel corpus generale di lingua

¹⁶⁰ Vi è una differenza tra frequenza assoluta e frequenza normalizzata (relativa). Infatti, se basta la frequenza assoluta (ossia il mero conteggio) lavorando con un solo corpus, al contrario è indispensabile una normalizzazione dei dati in caso di confronto tra due o più corpora. Si ottiene la frequenza normalizzata moltiplicando il valore assoluto dei risultati ottenuti per un totale di parole, da stabilire in base alle dimensioni dei corpora considerati (per corpora molto grandi in genere si moltiplica per un milione il valore assoluto, mentre per corpora più ridotti è accettabile moltiplicare per mille), e dividendo per il numero di parole del corpus. A livelli più avanzati, e già nel processo di formulazione di una teoria, un concetto importante è il livello di significatività statistica, che dovrebbe essere tale da rendere la teoria accettabile. I test statistici possibili, in questi casi, sono innumerevoli. Il più usato in LC, come spiegano McEnery et al., è il *chi-square test*, anche se gli autori stessi prediligono il *log-likelihood test* (LLT) (McEnery et al., 2006: 52-55). Al sito <http://corpora.lancs.ac.uk/clmtp/2-stat.php> l'Università di Lancaster (UK), offre un programma per effettuare un LLT e calcolare quindi la significatività statistica di un dato osservato. Calcoli

russo l'espressione si ritrova 0,02 volte per milione, mentre nel CP ita-ru le occorrenze ammontano a 10,5 volte per milione.

Insieme ai causativi, le funzioni di vicario e verbo supporto costituiscono gli usi più diffusi del verbo F. Le funzioni che presentiamo di seguito sono invece statisticamente secondarie.

C) **F-lavoro** (7 esempi): si tratta di una categoria molto circoscritta e linguospecifica, poiché, a differenza degli altri casi, non è riscontrabile nelle lingue romanze diverse dall'italiano. Nonostante la possibile confusione tra questo uso di fare e le costruzioni con F come verbo di supporto, La Fauci e Mirto (La Fauci, Mirto 2003: 61-75, Mirto 2003), mettono in evidenza le caratteristiche sintattiche che regolano le costruzioni con F-lavoro e le distinguono in maniera sostanziale dalle CVS, prima fra tutte la non transitività di tali costruzioni rispetto a F-supporto, che al contrario è da considerarsi inequivocabilmente transitivo¹⁶¹. Nemmeno in russo le costruzioni indicanti la professione di un soggetto sono transitive:

(6a) “*Aggiungiamo martedì*”, disse Sam Stull, che *faceva il cameriere*. (Baricco)

(6b) «Добавим вторник», – сказал Сэм Сталл, *служивший офицером*.

D) **F-inciso** (56 esempi)¹⁶²: il verbo F può essere usato nel registro colloquiale con il significato di “dire” o di qualsiasi *verbum dicendi*. In tutti questi casi, il traduttore russo deve scegliere tra numerose varianti sull'asse paradigmatico: dal più neutro

matematici più complessi (es: ARF - *average reduced frequency*) possono essere poi impiegati, ad esempio, nei casi in cui una parola sia particolarmente frequente per via dell'argomento dei testi inseriti. Durante un workshop dedicato al funzionamento di Sketch Engine (“Lexicom 2016”. Cfr. <http://www.lexmasterclass.com/>), è stato presentato l'esempio della parola inglese *dream*, le cui occorrenze erano particolarmente concentrate in uno dei testi di un corpus. Visualizzando i metadati, si notava infatti che il testo in questione era intitolato “*sleeping and dreaming*”. Il modo più immediato e semplice (che non implichi dunque l'uso di calcoli complessi) per evitare incidenti di questo tipo durante l'analisi, è conoscere il proprio corpus. Ciò è possibile, lo ribadiamo, avendo a disposizione le informazioni extralinguistiche fondamentali relative ai testi che lo compongono.

¹⁶¹ La Fauci e Mirto ritengono imprecisa la dicitura “verbo transitivo” o “intransitivo” proprio perché uno stesso verbo può essere impiegato allo stesso tempo transitivamente o meno. È per questo che la costruzione, e non il verbo, dovrebbe essere definita transitiva o intransitiva (La Fauci, Mirto 2003: 71).

¹⁶² Questa categoria non è trattata in modo esaustivo dai linguisti italiani. Per approfondimenti proponiamo uno studio di Ponchon su *faire* francese, usato analogamente all'italiano in questo senso (Ponchon 1994: 343-381).

“govorit'/skazat” (dire) (7b) a interpretazioni più inusuali quali “burknut” (grugnire) (8b):

(7a) *Così se qualcuno ci salirà, capirà che siamo arrivati prima noi*, **ho fatto io**. (Ammaniti)

(7b) *Чтобы каждый, кто туда заберется, знал, что раньше его там были мы, - сказал я*.

(8a) — *Come dici?* — **fece lui, che continuava a restarci male ogni volta**. (Calvino)

(8b) — *О чем?* — **буркнул Козимо, совсем растерявшись**.

E) Uso impersonale di F (6 esempi): F è spesso impiegato in espressioni impersonali quali «*fa caldo*», «*fa freddo*»:

(9a) *Là sotto faceva più caldo*. (Bassani)

(9b) *Там было гораздо жарче*.

Il corpus ha reso possibile identificare ulteriori usi poco ricorrenti, che possono esser raggruppati a loro volta in sotto-gruppi funzionali. Esempi di questo tipo sono talvolta menzionati negli studi su F, ma per le difficoltà di interpretazione e, dunque, di classificazione, non viene attribuita loro un’etichetta specifica e, soprattutto, univoca. Menzioniamo, ad esempio, locuzioni come “*fare il matto*”, “*fare il medico*”, “*fare l’avvocato*” con il significato di “*comportarsi come un matto, un medico o un avvocato*”, o ancora l’uso definito da La Fauci e Mirto (2003) “*Fare-ruolo*”, dove F significa “*interpretare*”, “*recitare la parte di*”, solitamente in uno spettacolo o in un film. Tutti questi usi testimoniano ulteriormente la polivalenza di F, descritta la quale possiamo ora concentrarci sulla funzione principale, quella di F causativo.

Si parla F causativo quando questo verbo, indipendentemente dalla forma che assume, è seguito da un infinito e assume il valore di – citiamo il dizionario Battaglia - “*cercare, procurare, ottenere, conseguire un determinato risultato; ordinare, comandare, ingiungere; aiutare, rendere possibile o agevole; permettere, concedere, tollerare; essere causa, motivo, pretesto*”. Cfr.: “*Vuol aver la bontà di **farlo venire** un momento qui, che gli ho da dire una parola?*” (De Amicis).

4.2. Situazioni causative e verbi causativi: introduzione

Molti linguisti hanno dedicato i loro studi alle possibili espressioni linguistiche del rapporto causa-effetto, soprattutto per via dell'universalità del concetto di "causa": come afferma Podlesskaja (1993: 165), insieme a "identità", "esistenza" o "negazione", il concetto di "causa" rientra nel gruppo delle relazioni semantiche "fondamentali". La presenza di un gran numero di ricerche, tuttavia, rende difficile adottare un metalinguaggio univoco per trattare il tema, problema sollevato da tempo da Wierzbicka in *The semantics of grammar* (1988: 237), dove la studiosa esprime la sua insoddisfazione per la mancata univocità terminologica e mette in discussione l'universalità del causativo, da molti riconosciuta. È noto infatti, secondo Wierzbicka, che le lingue differiscono a seconda dell'attenzione che rivolgono a date porzioni di realtà, e anche nel caso della causatività tali differenze emergono in modo evidente (1988: 250). Proprio questa eterogeneità interlinguistica ha spesso portato le diverse tradizioni ad affrontare il tema della causatività da prospettive differenti: come vedremo, se in russo prevale l'approccio semantico, i linguisti italiani, distaccandosi dalla tradizione lessicografica (cfr. Battaglia), si sono dedicati a studi prevalentemente sintattici.

Consapevoli di tali differenze, in questi paragrafi si cercherà di formulare alcune definizioni fondamentali e di evidenziare i parametri da tenere in considerazione nello studio di quelli che vengono definiti "verbi causativi".

La parola "*causare*", ricorda Jurij Apresjan, è un concetto universale impiegato nella definizione lessicografica di molti verbi causativi (es.: *aprire* → *causare l'apertura*) e, nonostante la sua incompletezza semantica ("*causare l'apertura*", infatti, non significa necessariamente "*aprire*"), sembra essere l'unica in grado di garantire una definizione abbastanza univoca (Apresjan 1995: 256). Secondo il linguista, "Causare P significa, per definizione, agire in modo tale che la situazione P abbia immediatamente luogo o cominci ad avere luogo"¹⁶³. Accanto a questo tipo di definizioni semantiche, si trovano anche descrizioni formali, come quella di Sumbatova (1993: 253), la quale, concentrata sul processo sintattico di causativizzazione descritto da Comrie, afferma: "La

¹⁶³ "Казуировать P значит, по определению, действовать так, что непосредственно начинает иметь или имеет место ситуация P".

causativizzazione è spesso descritta come una sorta di derivazione verbale in cui la frase originale è incorporata in posizione di secondo argomento”¹⁶⁴.

In generale, i rapporti di causa-effetto possono essere espressi linguisticamente attraverso un gran numero di costruzioni sinonimiche. Nedjalkov e Sil’nickij (1969: 5-8) definiscono l’insieme delle costruzioni causative (CC) in termini di “situazione causativa”. Una costruzione è definita causativa quando esprime una macro-situazione comprendente due micro-situazioni legate tra loro da un rapporto di causa (C), definito anche “legame causativo”. La descrizione delle CC, dunque, non si baserebbe su caratteristiche interne, ma esterne: l’indagine, secondo i due studiosi, partirebbe dal livello referenziale per arrivare al livello grammaticale. In altre parole, una costruzione può essere causativa se esprime una situazione di causa e effetto.

Una macro-situazione può essere realizzata linguisticamente da due proposizioni unite in genere da una congiunzione con funzione causativa (10)¹⁶⁵, o da una proposizione semplice (11), in cui – puntualizza Comrie – generalmente si tralasciano alcuni particolari:

(10) *Maria mi ha distratto, perciò ho sbagliato.*

(11) *Maria mi ha fatto sbagliare.*

Anche in (11), così come in (10), Maria provoca lo sbaglio di un soggetto A, ma il modo in cui ciò avviene qui non è specificato (Comrie 1989: 165).

Nedjalkov e Sil’nickij aggiungono che anche all’interno di una proposizione semplice il legame causativo può essere espresso in vari modi oltre che da causativi verbali: da preposizioni di causa (*iz-za*); da sostantivi (*Tvoja bestaktnost’ – pričina ego uchoda*¹⁶⁶) o aggettivi (*Ty vinovat v ego uchode*¹⁶⁷).

Tuttavia, secondo molti linguisti, tra cui Comrie (1989: 166) e gli stessi Nedjalkov e Sil’nickij (1969: 9), è il causativo verbale a suscitare maggiore interesse, per due motivi: la nozione di causa è contenuta nel verbo stesso; il verbo non è solo il nucleo semantico di una CC, ma anche quello sintattico.

¹⁶⁴ “Causativization is often described as a kind of verbal derivation in which the original clause is embedded in the position of the second argument”.

¹⁶⁵ Gli esempi (10) e (11) sono tratti da Comrie 1989.

¹⁶⁶ “La tua mancanza di tatto è la causa della sua partenza”.

¹⁶⁷ “Sei colpevole per la sua partenza”.

Un verbo causativo, afferma Padučeva (2003: 174), descrive una situazione in cui un partecipante (B) subisce un cambiamento provocato da un altro partecipante (A).

Riprendendo l'esempio (11), una frase con verbo causativo può essere dunque schematizzata come segue:

A (Maria) agisce su B (io) in modo che compia l'azione X (sbagliare)

Dove A è tradizionalmente definito *causer* – “causatore” - e B *causee* – “causato”¹⁶⁸.

Jurij Apresjan definisce i verbi causativi non come “predicati”, bensì come “congiunzioni” (*sojuzy*), proprio perché legano due eventi (1995: 35).

I linguisti si trovano generalmente d'accordo nel distinguere tre tipi di causativo realizzato all'interno di una proposizione semplice:

- **Morfologico**: la componente causativa viene aggiunta ad un verbo mediante suffisso. Si prenda ad esempio il greco antico, dove *kai-omai* significa “io brucio/prendo fuoco”, mentre *kai-ō* “io do fuoco a qualcosa” (Nedjalkov, Sil'nickij 1969: 21).
- **Lessicale**: la componente causativa è espressa da un singolo verbo ed è insita nel suo significato lessicale (es.: *uccidere*), ma è legata anche a quello sintattico, cioè dipende dalle relazioni sintattiche (per esempio, *andare*, per via della sua intransitività, non può essere causativo)¹⁶⁹.
- **Analitico** [definito anche “sintattico” da molti linguisti o “produttivo” da Shibatani in (1976: 2-3)]: la componente causativa è espressa da un verbo considerato in questo caso servile, preposto ad un altro verbo espresso all'infinito, come nel caso della costruzione causativa italiana con F¹⁷⁰.

¹⁶⁸ Questa distinzione tra *causer* e *causee* è talvolta resa in italiano come “iniziatore” ed “esecutore”. Tuttavia, sembra più consono in questo caso adottare la terminologia più simile all'inglese, largamente riconosciuta. Inoltre, i due termini italiani “iniziatore” ed “esecutore” si addicono maggiormente a partecipanti animati, e risultano quindi inadeguati quando manca questa premessa.

¹⁶⁹ Come puntualizza Padučeva (2003: 174), portando ad esempio *zpozomaty* (*rumoreggiare, rimbombare*) esistono verbi intransitivi classificati come “causativi”, ma i pareri sulla loro effettiva natura causativa sono discordanti.

¹⁷⁰ Si parla di causativo sintattico anche in relazione a frasi complesse legate da congiunzioni di tipo causativo o da espressioni quali “fare sì che” (in russo “*delat' tak, čtoby*”). Sul causativo realizzato mediante frasi complesse è interessante, ad esempio, la classificazione di Song (1996), che propone una nuova tipologia di costruzioni causative definite COMPACT (causativo lessicale e morfologico, ma anche costruzioni quali F+inf.), AND (causa ed effetto sono legate da una congiunzione coordinativa) e PURP (causativo sintattico con due proposizioni legate tra loro in vari modi).

Sui confini di queste distinzioni le opinioni talvolta divergono. Non vi è quasi alcun dubbio relativamente al causativo morfologico, tuttavia, come fa notare Comrie, molti verbi russi formanti la coppia “causativo/de-causativo” (ad es.: *lomat’-lomatsja* – rispettivamente *rompere-rompersi*) presentano anche una componente morfologica, benché il causativo non sia dato dall’aggiunta di un suffisso, ma dalla sua assenza e benché *lomat’* sia generalmente trattato come causativo lessicale. La traduzione italiana rispecchia una tendenza simile al russo quando l’aggiunta del pronome riflessivo *-si* determina il de-causativo¹⁷¹ (anche se gli studi di linguistica italiana sembrano non essersi interessati a questo tipo di verbi, concentrando piuttosto l’attenzione sulla costruzione F + inf).

Maggiore ambiguità si riscontra, secondo Comrie, nel caso delle costruzioni analitiche con verbo “servile” (tipo F) seguito da infinito: queste si troverebbero a metà strada tra causativo morfologico e analitico, poiché realizzate all’interno di una proposizione semplice, mentre, secondo lo studioso, si potrebbe parlare di causativo analitico solo in presenza di due proposizioni in rapporto subordinativo o coordinativo. I tre tipi di causativo andrebbero quindi collocati su un continuum (Comrie 1989: 169). In realtà esistono divergenze anche sullo status di proposizione semplice di una frase con F + inf., come si evince dal tentativo di Skytte (1976) di legittimare, con un elenco di nove motivazioni, la teoria secondo cui F + inf. sarebbe un unico predicato complesso, analitico e transitivo. Sempre Skytte, più tardi, in una monografia del 1983, esprime il suo disaccordo con la posizione assunta da alcuni generativisti-trasformazionalisti, convinti della compresenza di due proposizioni nella costruzione F + inf. – una contenuta nell’altra – e dello status dell’infinito come predicato a sé stante (Skytte 1983: 51). In questa sede, condividiamo l’opinione della studiosa danese, la quale sostiene che F + inf. sarebbe un predicato complesso realizzato in una proposizione semplice; tuttavia ci distanziamo da Comrie e dalla sua definizione “bifrasale” di causativo analitico, considerando invece la costruzione F + inf. un esempio di causativo analitico a tutti gli effetti, come ribadito anche da Nedjalkov e Sil’nickij (1969:23). Per questi ultimi, anche la distinzione tra causativo morfologico e lessicale è netta: morfologico è il causativo ottenuto da un processo regolare e produttivo, mentre lessicale è il causativo ottenuto da un processo non produttivo.

¹⁷¹ Sul tema del de-causativo cfr. Padučeva (2001, 2003), e Dolinina (1991) in Bondarko (a cura di), che parla della relazione tra causativo e riflessivo

Aggiungiamo, inoltre, una nota di Talmy, per cui un verbo può essere allo stesso tempo causativo e non causativo. Ne è un esempio il verbo inglese *break* (*rompere*), considerato non causativo non solo nella variante “*the vase broke*” [(il vaso si è rotto), analogo ai de-causativi in russo e in italiano marcati morfologicamente con *-sja(s’)* e *-si: lomat’sja* e *rompersi*], ma anche in “*I broke my arm*” (*ho rotto/mi sono rotto un braccio*), dove il soggetto coincide con il paziente (Talmy 2001, vol I: 69-70). Allo stesso modo, per la lingua russa, Dadueva (2011: 80) riporta un esempio di Gordon con il verbo *skazat’* (*dire*), di norma non causativo, ma impiegato in alcune circostanze come sinonimo di “ordinare” e quindi con funzione causativa: *On skazal emu vyjti* [*gli ha detto (ordinato) di uscire*]. In generale, afferma Apresjan in (1995: 47), ogni verbo causativo ha un corrispettivo de-causativo (es.: *soobščat’ – znat’ / informare – sapere*) e presenta, rispetto ad esso, una valenza in più, tant’è che “quasi tutti i verbi trivalenti sono causativi”.

Oltre alla differenza tra causativo morfologico, lessicale e analitico, ulteriori importanti distinzioni riguardano i seguenti parametri: 1) la forza causativa del verbo; 2) la distanza tra causa ed effetto; 3) il grado di controllo del soggetto causatore. Vediamoli nel dettaglio.

Il primo parametro è di tipo semantico e riguarda il valore causativo assunto dal verbo, che può essere collocato su di un continuum con due estremi: valore *fattitivo* o *permissivo*. La coercizione, compresa nel valore fattitivo, è il grado massimo di forza causativa. Tra il grado coercitivo e quello permissivo esistono però una serie di gradi intermedi. Come afferma anche Wierzbicka (1988: 246-247), una particolarità dell’italiano sta nella capacità di realizzare ciascuno di questi valori, o gradi intermedi, con F – definito come causativo “*all-purpose*” - pur ammettendo la costruzione *lasciare* + inf. per la variante permissiva¹⁷². Altre lingue impiegano verbi diversi, non interscambiabili, a seconda della forza causativa (cfr. l’inglese *make vs let*, o il russo, che dispone di molti sinonimi per le due realizzazioni: *zastavit’, velet’, razrešit’, pozvolit’* ecc.). Nedjalkov e Sil’nickij (1969: 31) attestano, accanto al causativo fattitivo e permissivo, un tipo di causativo “assistenziale” o “di aiuto”, reso in alcune lingue morfologicamente. In quechua, ad esempio, il suffisso *-isi* (*-usi*) indica l’aiuto: *llank’a* significa “*lavorare*”, mentre *llank’a-isi* si traduce come “*aiutare qualcuno nel lavoro*”.

¹⁷² Come spesso accade, la terminologia può essere fuorviante, poiché la costruzione causativa italiana è talvolta denominata “fattitiva” pur presentando il verbo *lasciare* (cfr. Skytte 1976: 355).

Come notato da Wierzbicka, e come si mostrerà in seguito, il verbo F ha una semantica così opaca, da assumere anche una forza causativa assistenziale.

La questione dei gradi di forza causativa è ripresa da Simone e Cerbasi (2001: 445-447), proprio in relazione a F + inf. Per la sua tendenza ad impiegare tale costruzione anche in situazioni dove la forza causativa è pressoché nulla, l'italiano, analogamente al francese – ma a differenza di altre lingue romanze come spagnolo e portoghese¹⁷³ –, è definito una lingua “a forte orientamento causativo” (Simone, Cerbasi 2001: 457). Spesso, quando la forza causativa è minima, si parla di “falsi causativi”. Moreno, in (1993: 159), ovvia a questa confusione affermando che i causativi, di qualunque tipo essi siano, per essere considerati tali devono presentare almeno due dei tre “primitivi semantici” che li caratterizzano: “forza”, “intenzione” e “transizione”¹⁷⁴. Di conseguenza, anche laddove la forza è minima, se il predicato presenta i due restanti primitivi semantici, può essere considerato causativo. In (Simone, Cerbasi 2001: 458-60), si prosegue tuttavia sulla linea del “falso” *fare* causativo, identificando per la precisione tre classi¹⁷⁵: 1) falsi causativi, dove il causato sembra in realtà un beneficiario dell'azione (cfr. “*la zia ha fatto mangiare al bambino una buonissima torta*”); 2) causativi ambigui, dove non sono chiare né la forza causativa del verbo, né l'agentività di causatore e causato (cfr. “*ho fatto studiare mio figlio*”); 3) causativi reali, dove il causatore ha un minor grado di agentività, e promuove un'azione compiuta dal causato (cfr. “*L'imperatore fece costruire una flotta ai marinai*”).

Il secondo parametro riguarda il rapporto di vicinanza temporale e logica tra causa ed effetto e il contatto diretto tra causatore e causato. Si parla in questo caso di causativo diretto e indiretto, o anche “mediato” come propone DeLancey (1984: 182). Generalmente, i causativi lessicali o morfologici sono diretti, mentre la variante analitica tende ad essere indiretta (Comrie 1989: 173, Song 1996: 4-5). Cfr.: (12) *A ha rinchiuso B* vs (13) *A ha fatto rinchiusere B*. In realtà vedremo che in italiano ciò dipende dalla transitività o meno del verbo all'infinito (cfr. par. 4.5.1.).

Aggiungiamo che in questo specifico esempio la frase (13) implica non solo una maggiore distanza temporale tra causa ed effetto, ma anche l'intervento di un ulteriore

¹⁷³ Cfr. Cerbasi (1998), dove si specifica come gli equivalenti di *fare* “*hacer*” e “*fazer*” concorrano con il verbo “*mandar*”, preferito per il grado elevato di forza causativa.

¹⁷⁴ Il concetto di “primitivi semantici” è ampiamente trattato in (Wierzbicka 1996).

¹⁷⁵ Gli esempi sono dei due autori.

partecipante¹⁷⁶. A questo proposito, un'importantissima nota riguarda la differenza tra causativo diretto e indiretto in russo. Padučeva (1997: 64) ricorda come la lingua russa sia in questo senso ambigua, non distinguendo in modo esplicito le due situazioni: la frase (14), infatti, in mancanza di alcun indizio, può essere interpretata indifferentemente come (14a) e (14b) (esempio di Padučeva, traduzione nostra):

(14) *Я сшила себе юбку*

(14a) *Mi sono cucita una gonna*

(14b) *Mi sono fatta cucire una gonna (da qualcuno)*

Per la sua rilevanza in chiave comparativa, torneremo su questo tema con l'analisi degli esempi (cfr. par 4.5.1.).

Infine, il terzo parametro (il grado di controllo del soggetto causatore) è ampiamente discusso da Padučeva (2003), le cui categorie tassonomiche per classificare i causativi sono strettamente legate all'idea di controllo¹⁷⁷. Il controllo da parte del soggetto si lega, a sua volta, ai concetti di agentività e volontà/intenzione (cfr. anche DeLancey 1984). Padučeva afferma che un elevato grado di controllo da parte del soggetto implica un'azione, mentre un basso (o assente) grado di controllo mette in scena un avvenimento (*action vs happening*, secondo la terminologia presa in prestito da Wierzbicka). In (Padučeva 1997: 62) si puntualizza inoltre che, sebbene l'idea di controllo sia parzialmente legata all'essere animato o inanimato del soggetto, il vero e proprio causatore non può essere la persona in sé, bensì l'attività da lui compiuta. Se il soggetto, al contrario, è inanimato, il causatore sarà un suo particolare stato o caratteristica. Naturalmente l'animatezza di un soggetto non implica direttamente un controllo sulla situazione causativa: da un lato interviene il concetto di intenzionalità o meno (nel caso di *rompere*, ad esempio, l'azione non è necessariamente intenzionale), dall'altro, per alcune attività, quali *rešit' problemu (risolvere un problema)* o *ugovorit' (convincere)*, il controllo da parte del soggetto sull'effetto desiderato non è totale, dal momento che, afferma Padučeva, la semantica di questi verbi implica l'idea di "fortuna".

¹⁷⁶ Da notare, però, che l'aggiunta di un nuovo partecipante alla situazione causativa non si verifica sempre col passaggio da causativo lessicale a analitico: basti pensare che il tipico esempio in lingua inglese presentato dai linguisti per distinguere tra causativo diretto e indiretto è "A killed B" vs "A caused B to die", dove il numero di partecipanti, se non esplicitato il contrario, rimane invariato.

¹⁷⁷ Padučeva (2003: 73) distingue le seguenti categorie: intenzionale in svolgimento, intenzionale realizzata, intenzionale non completamente controllata, intenzionale garantita, non-intenzionale sotto forma di evento (che evoca una certa conseguenza), non-intenzionale in svolgimento, non-intenzionale realizzata.

Sempre nell'ambito del controllo, Arkad'ev e Letučij (2009) introducono il concetto di "implicatività" (in russo *implikativnost'*), per cui solo alcuni tipi di causativi implicano che l'azione causata effettivamente si verifichi. Si tratta, ad esempio, di una differenza fondamentale tra il causativo fattitivo e permissivo (mentre per verbi esprimenti una forza causativa intermedia, dipende dalla semantica degli stessi. Cfr.: *chiedere* vs *aiutare*). In base a tale distinzione, se la frase (15) non è ammessa, poiché *zastavit'* implica necessariamente la realizzazione dell'azione, (16) invece lo è (esempi di Arkad'ev e Letučij):

(15) **Папа заставил Васю закрыть дверь, но Вася не закрыл дверь*¹⁷⁸.

(16) *Мама разрешила детям погулять во дворе, но они решили вместо этого поиграть в компьютер*¹⁷⁹.

L'esempio (16) è un tipico caso definito da Padučeva di "controllo non totale", dal momento che, nonostante la volontà del causatore, che agisce con un'azione A, il causato è in grado di scegliere se portare avanti l'azione B o rinunciarvi.

In (1973: 282), Zolotova affronta l'argomento della realizzazione o non realizzazione dell'azione da parte del causato, ricordando come in alcuni casi sia possibile esprimere morfo-sintatticamente questo parametro grazie all'aspetto: per svariati verbi il perfettivo esprime la realizzazione dell'azione. Per altri verbi invece, quali appunto *razrešit'* in (16), l'aspetto non gioca questo ruolo e il causato ha sempre libertà di scelta.

Ci chiederemo, in seguito, come si pone la costruzione F + inf. in termini di realizzazione dell'azione, controllo e agentività. Gli esempi del corpus ci forniranno i dati per esplorare questi parametri. Nel paragrafo successivo ci occuperemo invece, più nel dettaglio, della descrizione del causativo russo e italiano, proponendo un confronto preliminare tra i modi di realizzazione del causativo nelle due lingue e tra gli studi ad esso dedicati. Ulteriori osservazioni e approfondimenti seguiranno poi nella sezione dell'analisi.

¹⁷⁸ "Il papà ha costretto Vasja a chiudere la porta, ma Vasja non l'ha chiusa".

¹⁷⁹ "La mamma ha permesso ai bambini di andare in cortile, ma loro hanno deciso, invece, di giocare al computer".

4.3. Gli studi linguistici russi e italiani sul causativo

Gli studi della russistica e dell'italianistica dedicati a questo tema presentano una differenza sostanziale: in russo il causativo viene trattato prettamente dal punto di vista semantico, mentre in italiano, nonostante l'esistenza di causativi lessicali con una semantica analoga al russo (cfr. *ubit'* – *uccidere*, *lomat'* – *rompere* ecc.), prevalgono descrizioni formali e parlando di causativi si tratta quasi esclusivamente la costruzione F + infinito.

4.3.1. Il causativo italiano

La Fauci e Mirto (2003: 20-22) definiscono causativo un verbo in grado di aumentare la quantità di funzioni grammaticali di una data proposizione.

(17) [...] *e va in collera coi parenti che **fan digiunare** i bimbi per castigo.* (De Amicis)

In (17), il soggetto logico del verbo all'infinito è espresso dal SN “*i bambini*” anche se, sintatticamente, si tratterebbe del complemento oggetto diretto del sintagma verbale “*far digiunare*”. Di fatto, il causato (*i bambini*) può essere sostituito dal pronome clitico “*li*”, confermando a tutti gli effetti il suo status di complemento. F, dunque, non fa sì che il complemento oggetto abbandoni la sua funzione primaria, ma gli conferisce una funzione argomentale aggiuntiva (oggetto + soggetto). Inoltre, consente al verbo “*digiunare*”, di norma usato intransitivamente, di occorrere in una costruzione transitiva.

Una definizione così formale di verbo causativo sembrerebbe ammettere come oggetto di studio solo i causativi analitici *fare* e *lasciare*. E difatti, analogamente a La Fauci e Mirto, i linguisti italiani impegnati nell'analisi del causativo trattano soprattutto questi due verbi. Molti sono gli studi che descrivono la costruzione F + inf. in chiave diacronica¹⁸⁰, e in generale tutte le indagini, anche quelle dedicate all'italiano moderno, esaminano la costruzione dal punto di vista sintattico, interessandosi in particolare ai seguenti aspetti: posizione del clitico, realizzazione del causato (complemento oggetto, di termine o d'agente) e ordine delle parole. Gli studi diacronici su F causativo ne attestano la presenza non solo in italiano antico, ma anche in latino, come mostrato, per

¹⁸⁰ Cfr. Brambilla Ageno (1964: 468-472), Cerbasi (1998), Consales (2006), Pelo, Consales (2003), Robustelli (1993,1994,1995), Simone, Cerbasi (2001).

esempio, da Cerbasi (1998: 457-461), che descrive il passaggio dalla costruzione bifrasale latina alla monofrasale italiana. *Facio* concorreva inizialmente con altri verbi; in casi di forza causativa elevata il latino prediligeva *ibueo*, o eventualmente *cogeo* e *suadeo*, seguiti più spesso da una secondaria con AcI¹⁸¹ (18) e da *ut* + congiuntivo (19), o in altri casi da congiuntivo semplice (esempi di Cerbasi):

(18) *Caesar iussit pontem rescindi*¹⁸².

(19) *Populus Romanus iussit ut Sullae voluntas esset pro lege*¹⁸³.

In queste situazioni *facio* non costituiva la scelta preferenziale, anche se possibile; *facio* sembrava essere più usato nel caso in cui il causatore non esercitasse controllo sull'azione. Solo più tardi, grazie alla sua generalità semantica, cominciò ad essere impiegato anche per eventi di carattere coercitivo. L'uso di *facio* + inf in latino è attestato anche da Robustelli (1993: 125) che, analogamente a Cerbasi colloca la costruzione già a partire dal I secolo dopo Cristo. In particolare, nella poesia, l'uso di *facio* + inf consentiva di liberare il discorso da troppe congiunzioni, soprattutto nei casi in cui anche *facio* fosse stato introdotto, a sua volta, da congiunzione, come in (20) (Cerbasi 1998):

(20) *ut iustos vivere faciat*.¹⁸⁴

Nonostante questo esempio, *facio* e l'infinito occorreano più spesso separati, mantenendo appunto la struttura bifrasale. Il passaggio dalla costruzione latina all'attuale costruzione monofrasale italiana, dove *F* + inf. costituiscono un'unità verbale per lo più inseparabile (ad eccezione di clitici o di alcuni avverbi), avviene gradualmente nei secoli e si consolida in maniera definitiva nel diciannovesimo secolo (Robustelli 1994). In questo periodo era ancora possibile frapporre un complemento oggetto tra i due verbi (richiamando la costruzione AcI), ma solo nello stile aulico. Si trattava dunque di una struttura arcaica, impiegata in rarissimi casi.

Tornando all'italiano moderno, come precedentemente accennato, sono pochi gli studiosi che trattano *F* + inf. in termini non esclusivamente sintattici. Tra questi citeremo

¹⁸¹ Costruzione latina composta da *F* + *accusativo* + *infinito passivo* trasmessa anche all'italiano antico, ma con infinito attivo.

¹⁸² "Cesare ordinò di abbattere il ponte".

¹⁸³ "Il popolo romano ordinò che la volontà di Silla venisse resa legge".

¹⁸⁴ "Che faccia vivere i giusti".

Simone e Cerbasi (2001), che considerano anche altri parametri (forza causativa e polisemia del verbo), e Guasti (1996), che identifica le differenze semantiche inerenti ai diversi modi di esprimere sintatticamente il causato: in particolare, considera le due espressioni “*far fare qualcosa A qualcuno*” vs “*far fare qualcosa DA qualcuno*”, che infatti in inglese presentano due traduzioni alternative (*make + inf.*, nel primo caso, e *have + inf.*, nel secondo)¹⁸⁵.

4.3.2. Il causativo russo

Gli studi della russistica sul causativo presentano, come dicevamo, la situazione opposta. La classificazione dei verbi causativi russi è basata sull’analisi semantica dei verbi. Questa attenzione nei confronti di aspetti prettamente semantici si spiega con l’assenza in russo di un causativo morfologico o analitico. Come afferma Letučij (2012: 31), il russo è tra le lingue in cui il causativo non è marcato grammaticalmente, bensì resta come componente implicita della semantica verbale. Lo studio sui verbi causativi russi si rivolge infatti quasi esclusivamente al causativo lessicale, ossia a verbi che possono essere considerati causativi grazie ai tratti semantici che li contraddistinguono (in particolare, il tratto “*vozdejsťvovat’*” – *influire*).

Šorochova (2007: 139) puntualizza che in mancanza di indicatori formali (quindi non nel caso di causativo morfologico o analitico), per riconoscere l’appartenenza di un verbo al gruppo dei causativi è necessario scomporre il suo significato in componenti o semi minimi; ad esempio, il verbo *toropit’* (mettere fretta) può essere scomposto in “*zastavljat’ spešit’*” (obbligare ad affrettarsi). Come si precisa in (Dadueva 2011: 76), tutti i verbi causativi presenterebbero il tratto semantico “*vozdejsťvovat’*”, in particolare: influire su un soggetto o oggetto affinché compia un’azione, provi uno stato d’animo oppure, per citare Zolotova (1982: 166), “*modifichi una sua condizione, qualità o posizione*”¹⁸⁶.

Il tratto semantico della causatività è poi spesso fonte di discussione: per esempio, un verbo come “*stroit’*” (*costruire*) per alcuni non sarebbe causativo (Dadueva 2011, Steksova 2008), mentre per altri, ad esempio Zolotova, l’espressione *stroit’ dom*

¹⁸⁵ Anche qui il punto di partenza è la sintassi: le differenze semantiche sono veicolate da diverse strutture sintattiche.

¹⁸⁶ Riportiamo la definizione di Zolotova (1982: 166) di verbo causativo: “*Казуативные глаголы означают воздействие на лицо или предмет, в результате которого лицо совершает действия, испытывает состояния, предмет изменяет свое состояние, качество, местоположение*”.

(*costruire una casa*) è classificata come causativa in quanto significherebbe “*influire direttamente sull’esistenza della casa*”. Come riporta Steksova, esistono due approcci per la classificazione dei causativi: uno ristretto, per cui un sostantivo con funzione di oggetto diretto è portato a compiere un’azione (*mat’ kormit rebenka*), oppure a modificare una sua condizione o qualità (*krem delaet kožu uprugoj*)¹⁸⁷ e uno più ampio, che considera causativi tutti i verbi seguiti da un complemento oggetto (incluso, appunto, lo stesso “costruire”) (Steksova 2008: 332). In questa sede adotteremo l’approccio più ristretto poiché l’approccio più ampio perde la differenza tra verbi semplicemente transitivi – che dunque hanno un certo influsso sull’oggetto – e verbi che agiscono su un oggetto in modo che quest’ultimo compia un’azione (o assuma uno stato). Per i casi dubbi, presenti tra gli esempi del corpus analizzati, ci atterremo all’annotazione semantica del corpus stesso, che rileva il tratto della “causatività”.

Arkad’ev e Letučij (2009: 1) parlano inoltre di varietà semantiche del causativo: causativo fattitivo, del tipo *zastavit’* (*costringere*), permissivo, tipo *razrešit’* (*permettere*), manipolativo, dove il causatore compie un’azione insieme al causato [es.: *mat’ uložila rebenka spat’* (*la mamma ha messo a letto il bambino. Lett: a dormire*)] e causativo direttivo, dove il causatore non agisce direttamente sul causato come nell’esempio precedente, ma piuttosto induce il causato a fare qualcosa e ciò conduce in ogni caso al risultato [per esempio in *mat’ poslala rebenka spat’* (*la mamma ha mandato il bambino a dormire*)]. Esiste infine in lingua russa un particolare gruppo di verbi causativi seguiti dall’infinito, ma come puntualizza Dadueva (2011: 77) citando Gordon, pur presentando un comportamento simile a quello dei causativi analitici di altre lingue, quali *fare* per l’italiano, non possono essere equiparati a verbi servili desemantizzati, in quanto mantengono intatto il proprio significato. Nedjalkov e Sil’nickij (1969: 9) classificano questi verbi tra i causativi lessicali chiamandoli “non risultativi”, poiché l’azione causata è espressa da un altro lessema (l’infinito). Tra i verbi di questo tipo si annoverano, ad esempio, *zastavit’* (*costringere*), *pozvolit’* (*permettere*), *prikazat’* (*ordinare*), *velet’* (*ordinare*), *prosit’* (*chiedere*). L’unica eccezione è presentata dal verbo *dat’* (letteralmente “*dare*”), che seguito da un verbo all’infinito subisce un processo di desemantizzazione analogo a F, realizzando un causativo permissivo o con forza causativa debole, come nell’esempio (21a), tratto dal CP ita-ru, e reso appunto in italiano con la costruzione F + inf. (21b):

¹⁸⁷ Esempi di Steksova.

(21a) *Ваш Редька ненавидит меня и всё хочет **дать понять**, что я поступил с нею дурно.* (Čechov, Rasskazy)

(21b) *Il vostro Rèdka mi detesta e vuole continuamente **farmi capire** che ho agito male verso di lei.*

Lo stesso fare in russo (delat' - sdelat') è da considerarsi un causativo lessicale ed è usato in russo in contesti differenti rispetto al F italiano:

(22a) *Невозможность пользоваться известными видами оружия **делает нас слабыми** (...)* (Šalamov)

(22b) *L'impossibilità di ricorrere a determinati tipi di «armi» **ci rendeva deboli** (...)* (Rapetti)

In (22a) *delat* è seguito da complemento oggetto + lo strumentale della qualità o condizione risultante. In italiano costruzioni di questo tipo possono presentare F, ma più spesso sono rese con il verbo “*rendere*”¹⁸⁸.

Uno dei pochi studi sui causativi in russistica incentrato su aspetti sintattici è di G. A. Zolotova, che nel suo *Očerk funkcional'nogo sintaksisa ruskogo jazyka* (1973: 278-297) propone di analizzare il causativo russo iniziando proprio dai verbi causativi seguiti da infinito.

Zolotova parte dalle costruzioni infinitive in generale, distinguendone due tipi:

- 1) con infinito soggettivo, es.: *ja choču est'* (lett.: *voglio mangiare*), dove il soggetto del verbo principale è anche il soggetto logico dell'infinito.
- 2) con infinito oggettivo, dove il soggetto logico dell'infinitiva è espresso in forma oggettiva (tramite oggetto diretto o caso obliquo) e non coincide con il soggetto del verbo principale: es.: *prošu est'* [(ti) *prego di mangiare*].

“In russo, la costruzione con *verbo + infinito oggettivo* rappresenta il modello tipico di costruzione causativa”, dichiara Zolotova (1973: 280). La studiosa parla in questo caso anche di *modal'no-voljuntativnoe otnošenje* nei confronti di un'azione (relazione modale-

¹⁸⁸ Sulla costruzione causativa russa *delat' čto-to/kogo-to* + agg. Strum. cfr. Steksova (2009) e Letučij (2012: 33-34).

volitiva), realizzata però da un altro soggetto (il soggetto A, che agisce sul soggetto B), a differenza del verbo *chotet'* (volere).

Concentrandosi poi sulle costruzioni del secondo tipo, quindi propriamente causative, Zolotova analizza dal punto di vista semantico il rapporto tra le due azioni e i due soggetti. L'azione compiuta dal soggetto B può essere intenzionale o meno; inoltre, la relazione modale-volitiva può essere positiva o negativa, per esempio¹⁸⁹:

- relazione positiva all'azione:
 - a) verbi che reggono l'accusativo: *prosit'* (chidedere), *ugovarivat'* (persuadere), *objazyvat'* (obbligare), *zastavljat'* (costringere) ecc. → *kogo* + inf.;
 - b) verbi che reggono il dativo: *pozvolit'* (permettere), *prikazat'* (ordinare), *sovetovat'* (consigliare), *pomogat'* (aiutare) → *komu* + inf.
- relazione negativa all'azione:
 - a) verbi che reggono l'accusativo: *otgovarivat'* (dissuadere) → *kogo* + inf.;
 - b) verbi che reggono il dativo: *zapretit'* (impedire), *prepjatstvovat'* (ostacolare) → *komu* + inf.

L'analisi di Zolotova consente di rilevare un elemento necessario delle costruzioni causative: la presenza (esplicita o meno) di due soggetti. In base a ciò, Zolotova enuncia altri parametri utili per classificare il tipo di situazione causativa. Ad esempio, in caso di mancata enunciazione del soggetto dell'infinito, quest'ultimo può essere di tre tipi:

- indeterminato-personale: (23) *chozjain velel založit' koljasku (il padrone ha ordinato di preparare la carrozza)*;
- generale-personale: (24) *Ja ne razrešaju rvat' zdes' cvety (non permetto di strappare i fiori qui – a nessuno)*;
- determinato-personale: (25) *Ja sovetuju soglasit'sja (consiglio di convenire – a te, interlocutore)*.

¹⁸⁹ La differenza è riscontrabile anche nelle costruzioni con infinito soggettivo.

Le situazioni degli esempi da (23) a (25) (1973: 281) sono riscontrabili anche in italiano, sebbene in (25) sarebbe più naturale esplicitare il causato con un pronome personale (ti/vi/Le)¹⁹⁰.

Anche nel caso di causativi propriamente lessicali, Zolotova (1973: 285) sottolinea l'importanza dell'analisi sintattica. La presenza di un causativo lessicale è data dal passaggio da un "modello statico" (non causativo) a uno "dinamico": la frase (26), ad esempio, esprime un'azione statica, mentre (27) esprime lo stesso rapporto tra "libro" e "tavolo", ma da un punto di vista dinamico (esempi di Zolotova). Secondo questa definizione, si considerano causativi anche verbi come il sopracitato *stroit'* (*costruire*).

(26) Книга под столом¹⁹¹

(27) X бросает (кладет) книгу под стол¹⁹²

Prendendo poi in considerazione anche la semantica dei singoli verbi, Zolotova cita i seguenti tipi di causativo:

Intenzionale:

- causativi di azione (*dejstvija*): ossia tutti i causativi + infinito descritti sopra;
- causativi di spostamento di un oggetto (*peredviženija predmeta*): *otpravljat'* (*mandare*), *vozit'* (*portare*);
- causativi di possesso (*vladenija*): *dat'* (*dare*), *darit'* (*regalare*), *prodavat'* (*vendere*);
- causativi di possesso di un'informazione (*vladenija informacii*): *soobščit'* (*comunicare*), *rasskazat'* (*raccontare*), ma anche *učit'*, *naučit'*, *obučit'* (tutti traducibili in italiano con *insegnare*);
- causativi di qualità (*priznaka*): *belit'* (*imbiancare*), *solit'* (*salare*), *točit'* (*affilare*);
- causativi di realtà/esistenza (*bitija*) (anche in negativo: distruzione o assenza): *stroit'* (*costruire*), *peč* (*cuocere al forno*), *pisat'* (*scrivere*), *razrušit'* (*distruggere*), *razbit'* (*rompere*)¹⁹³;

¹⁹⁰ In generale in italiano l'uso del pronome personale diretto o indiretto è più frequente che in russo e spesso, laddove in italiano è obbligatorio, in russo è omissibile: *Ma ora ve lo dico* vs *Ho сейчас скажу* (N. Ammaniti, *Io non ho paura*, 2001. Traduzione russa di V. Nikolaev, 2005).

¹⁹¹ "Il libro è sotto il tavolo".

¹⁹² "X fa cadere (getta) il libro sotto il tavolo".

¹⁹³ Come già accennato, molti verbi di questo gruppo sono considerati causativi solo adottando un approccio più ampio.

- causativi di trasformazione (*prevraščeniija*) (visti come variante dei causativi di esistenza): *prevraščat'* (*trasformare*).

Non intenzionale:

- Con causatore generalmente inanimato, come in (28), ma anche animato come in (29), purché la causa primaria sia comunque inanimata (in questo caso *derzost'*).

(28) *Его дерзость ужасала ее.*¹⁹⁴

(29) *Он ужасал ее своей дерзостью.*¹⁹⁵

Questi tipi di causativo, come *užsat'*, sono definiti da Zolotova di stato (*sostojanija*), e sono analoghi ai causativi trattati da V. Apresjan (2013) e definiti “emozionali” [quali, *pugat'* (*spaventare*), *serdit'* (*far arrabbaire*), *veselit'* (*rallegrare*) ecc.], per cui il causatore compie un'azione A che suscita un sentimento o un'emozione B al causato. Anche secondo Apresjan tali verbi possono essere agentivi (con soggetto animato) o non agentivi (con soggetto inanimato), tuttavia i causativi agentivi possono distinguersi in ciò che la studiosa chiama “conativi” e “non conativi”, ossia intenzionali o meno (Apresjan 2013: 46).

Zolotova parla infine di “autocausativi” quando l'agente provoca un'azione su di sé: “*io compro un libro*” - ad esempio - dove causatore e causato (o beneficiario) coincidono¹⁹⁶. (1973: 288).

4.4. Analisi degli esempi

In questo paragrafo descriveremo il lavoro di analisi svolto sugli esempi di un sotto-corpus parallelo composto da 21 estratti di opere letterarie italiane allineate alla traduzione russa.

Per l'analisi abbiamo utilizzato un approccio *corpus-driven*, senza formulare, cioè, ipotesi a priori sul fenomeno analizzato. Le tendenze linguistiche rilevate sono dunque evinte dagli esempi e saranno la base quantitativa per ulteriori riflessioni.

¹⁹⁴ A. Puškin, *Pikovaja dama*. “La sua insolenza le incuteva terrore”.

¹⁹⁵ Esempio di Zolotova. “Le incuteva terrore con la sua insolenza”.

¹⁹⁶ Come abbiamo visto nell'esempio di break al par.4.2, secondo Talmy, casi come questi non dovrebbero considerarsi causativi.

4.4.1. Metodologia

Un confronto *corpus-based*, nel senso più ampio del termine, (cfr. introduzione), tra la costruzione causativa italiana e il causativo russo sembra essere particolarmente significativo per colmare la lacuna di basi empiriche negli studi sui causativi analitici messa in evidenza da Gilquin (2010: 25), la quale evidenzia la difficoltà di estrarre strutture sintattiche specifiche da esempi di lingua autentica e la conseguente prevalenza di studi di tipo lessicale nella LC. La ricerca di strutture sintattiche, infatti, da un lato richiede che il corpus utilizzato sia annotato grammaticalmente in modo molto preciso e, dall'altro, prevede quasi sempre un intervento manuale sugli esempi ricavati che potrebbero non corrispondere a quanto effettivamente ricercato (tra poche righe mostreremo un esempio di questo tipo con la costruzione F + inf.). Nel nostro caso, il causativo italiano con F – pur rientrando tra i causativi analitici – si presta bene per una ricerca di tipo *corpus-based*, dal momento che F e l'infinito formano un predicato complesso indivisibile (cfr. Skytte 1983, par. 4.2) e il motore di ricerca del NKRJa è sufficientemente avanzato per estrapolare sintagmi con la medesima struttura sintattica. Gli unici elementi in grado di infraporsi tra F e il verbo all'infinito sono pronomi clitici e avverbi. La presenza del clitico, tuttavia, non compromette i risultati della ricerca, poiché si aggancia a F risultando in un'unica forma di parola che il software di interrogazione riconosce come singola. Pertanto, mediante la ricerca lessico-grammaticale nel NKRJa il motore troverà gli esempi del verbo F in tutte le sue possibili forme e i risultati comprenderanno anche forme di parola quali *fammi, farlo, farglielo* ecc. L'intersezione di avverbi tra F e il suo infinito, invece, è molto più rara, ma può essere comunque individuata modificando parzialmente la ricerca.

Per ottenere gli esempi di F causativo il corpus è stato dunque interrogato in due modi: inizialmente abbiamo inserito la voce “fare” nella ricerca lessico-grammaticale abbinata a qualsiasi verbo all'infinito con distanza 1, ossia richiedendo che i due verbi fossero contigui. Abbiamo ottenuto in questo modo 493 occorrenze della costruzione F + inf. In seguito, abbiamo nuovamente interrogato il corpus, ricercando F + avverbio + qualsiasi verbo all'infinito, anche in questo caso con distanza 1. Da questa seconda ricerca sono risultate 22 occorrenze. Abbiamo poi scartato gli esempi non pertinenti, erroneamente identificati dal sistema come costruzioni del tipo F + inf., [come in (30): *Riandare ai fatti, elencarli, individuarne le cause, gli effetti* (Eco, *Il pendolo di Foucault*)], ottenendo in definitiva 486 esempi. Questi ultimi sono stati poi suddivisi (par.

4.4.2.) sulla base di alcuni dei parametri sopracitati, ossia agentività e status animato/inanimato dei soggetti coinvolti (causatore e causato), in primo luogo per verificare il comportamento di F causativo in quattro possibili situazioni (1. causatore e causato ANIMATI; 2. causatore e causato INANIMATI; 3. Causatore INANIMATO e causato ANIMATO; 4. Causatore ANIMATO e causato INANIMATO) e, di conseguenza, la resa della traduzione russa. In secondo luogo, nei casi in cui il causatore è animato, abbiamo verificato il livello di intenzionalità e il grado di controllo sulla situazione causativa, sempre in prospettiva contrastiva con la traduzione russa. Anche in questo caso abbiamo ricercato eventuali regolarità di comportamento.

Abbiamo poi identificato cinque strategie di resa della costruzione italiana in lingua russa (par. 4.4.3.), senza però constatare alcuna corrispondenza tra una particolare resa e le quattro situazioni sopra elencate: tutte le strategie sono impiegate indipendentemente dal ruolo e dallo status degli attanti. Vi sono però, in assoluto, delle strategie preponderanti, su cui ci soffermeremo più nel dettaglio.

L'analisi ci ha infine consentito di formulare alcune ipotesi e osservazioni conclusive (par. 4.5., 4.6.) relative, in primo luogo, al confronto tra causativo russo e causativo italiano e, in secondo luogo, all'uso di F causativo.

L'analisi degli esempi dall'italiano al russo verrà affiancata da quella di estratti russi tradotti in italiano, così da verificare eventuali tendenze sia negli esempi in lingua originale, sia in quelli tradotti.

4.4.2. Ruolo e status degli attanti nella situazione causativa

In questo paragrafo gli esempi saranno presi in considerazione in base ai parametri precedentemente esposti, che qui ricordiamo: 1) causatore e causato ANIMATI; 2) causatore e causato INANIMATI; 3) Causatore INANIMATO e causato ANIMATO; 4) Causatore ANIMATO e causato INANIMATO.

F causativo presenta tutte e quattro le combinazioni citate, anche se nella maggioranza dei casi entrambi i soggetti risultano animati (situazione 1):

- 1) Causatore e causato ANIMATI (338 occorrenze – 69,54%):

(31a) *Tre viaggiatori, non di meno, a forza d'insistere con le domande, riuscirono a **fargli** snodare la lingua (...)* (De Amicis)

(31b) *Несмотря на это, трем путешественникам удалось **заставить** его **разговориться** (...)*

2) Causatore e causato INANIMATI (39 occorrenze – 8%):

(32a) (...) *scoppiò un gran tuono che fece **tremare** le pareti della casa (...)* (Manfredi)

(32b) (...) *вдруг раздался гром, **заставив** **дрожать** стены здания (...)*

3) Causatore INANIMATO e causato ANIMATO (52 occorrenze – 10,69%):

(33a) (...) *stavo cadendo nel gioco che aveva fatto **uscire** di senno Jacopo Belbo (...)*
(Eco, Il pendolo di Foucault)

(33b) (...) *я ввязался в игру, через которую Якопо Бельбо **потерял** **разум** (...)*

4) Causatore ANIMATO e causato INANIMATO (57 occorrenze – 11,72%):

(34a) *Il Teschio ha afferrato il sacco dove stava il volatile e ha cominciato a **farlo girare** in aria. (Ammaniti)*

(34b) *Череп **схватил** мешок, где сидела птица, и начал **размахивать** им в воздухе.*

Ci siamo chiesti se la traduzione russa rispettasse i parametri dell'originale e se fosse possibile individuare delle tendenze nella scelta dei traduttivi. Gli esempi dimostrano che il russo rispetta generalmente questi parametri, ma solo per quanto concerne il soggetto della frase, poiché, come vedremo, nella maggior parte dei casi la struttura a “doppio soggetto” della costruzione italiana non è mantenuta in russo, come in (33b), dove la traduzione russa rende il F + inf. con un verbo non causativo, e in (34b) dove il causativo analitico è reso in russo con un causativo lessicale, e quello che in italiano è il causato, soggetto logico dell'infinito, in russo è reso come un oggetto indiretto al caso strumentale.

Successivamente ci siamo concentrati sugli esempi con causatore animato (gruppo 1 e 4 – 395 esempi) così da verificare l'intenzionalità e il grado di controllo del causatore (iniziatore dell'azione) ed eventuali tendenze emergenti dal confronto tra le due lingue.

Abbiamo constatato che nella maggior parte dei casi, se il soggetto causatore è animato, agisce in modo intenzionale. Abbiamo contato solo 32 casi su 395 (8,1%) di chiara non intenzionalità e quindi di mancato controllo sulla situazione causativa da parte del causatore (non abbiamo considerato i casi ambigui, dove l'azione potrebbe anche essere intenzionale). Eccone alcuni esempi:

(35a) *Ma **fa ridere**, grande e grosso com'è (...)* (De Amicis)

(35b) *Но вместе с тем вид у Гарроне очень **смешной** (...)*

In (35a) è la semantica del verbo *ridere* a determinare in molti casi un mancato controllo da parte del soggetto causatore che, come qui, non “fa ridere” intenzionalmente. Abbiamo rilevato altri 7 esempi con il verbo *ridere* e in 5 casi la traduzione russa è simile a (35b): il causativo si perde e il predicato è espresso sotto forma di aggettivo o avverbio (*смеšno*), come in (36b):

(36a) *In verità **mi fate ridere**.* (Goldoni)

(36b) ***Смешно**, право!*

In un sesto caso, compare l'espressione “*far sorridere*” (Bassani), tradotta in russo con “*vyzval ulybku*”; mentre il “*mi fate quasi ridere*” di Collodi è reso in tutt'altro modo: “*Vy melete vzdor*”. In generale, vediamo che la non intenzionalità del causatore in italiano dipende soprattutto dal contesto. Riportiamo ancora tre esempi in cui si nota che le azioni in questi casi risultano non intenzionali e non controllate dal causatore, sebbene le tre forme verbali con F in (37a), (38a) e (39a) potrebbero essere intenzionali in contesti differenti:

(37a) (...) *mi **faceva morire** se faceva così.* (Baricco)

(37b) (...) *он **убивал** меня, поступая так.*

(38a) *Non **mi fate andar in collera**.* (Goldoni)

(38b) *Не **сердите** меня.*

(39a) *E rise, rise, facendo sobbalzare il pancione (...)* (Pirandello, *Il Fu Mattia Pascal*)

(39b) *И он закатился смехом, от которого у него долго содрогался живот (...)*

Per quanto riguarda il russo, esaminando questi 32 esempi, non si rileva alcuna tendenza traduttiva dominante. Le strategie di resa del russo si alternano in modo bilanciato. Esse sono dunque indipendenti dallo status e dal ruolo degli attanti.

Passiamo ad una descrizione più dettagliata delle strategie impiegate per tradurre F causativo.

4.4.3. F causativo: strategie traduttive dall'italiano al russo

L'analisi ha consentito di rilevare cinque strategie impiegate dai traduttori per rendere la costruzione italiana F + inf. Le elenchiamo di seguito con i relativi esempi in italiano e in russo:

1) Verbo causativo + infinito (95 occorrenze – 19,54%) **o sostantivo** (14 occorrenze – 2,88%). Come visto nel paragrafo 4.3., in russo esistono numerosi verbi seguiti da infinito soggettivo, che Zolotova definisce “causativi di azione”. Nei nostri esempi se ne registrano diversi: *zastavljat'* (costringere), *velet'* (ordinare), *prosit'* (chiedere), *pozvoljat'* (permettere), *dat'* (lasciare, permettere), come pure altri verbi inclusi in questo gruppo dalla studiosa, ma non registrati come causativi dall'annotazione semantica del NKRJa, per esempio *pomogat'* (aiutare)¹⁹⁷. Ricordiamo che tali verbi intervengono anche con attanti inanimati [cfr. (32b) nel precedente paragrafo] e che nonostante la loro struttura analitica, analoga alla costruzione causativa italiana, essi sono da considerarsi comunque causativi lessicali, poiché mantengono intatta la semantica primaria del verbo. Come già accennato, unica eccezione è *dat'* che, usato in senso causativo seguito da infinito, perde il suo significato primario (*dare*).

(40a) *Eumene aveva fatto progettare le macchine per gli effetti scenici dall'architetto Diade.* (Manfredi)

¹⁹⁷ Come abbiamo visto, la classificazione dei verbi causativi non è univoca e dipende dal tipo di approccio adottato. Inoltre in un corpus non disambiguato, come il quello parallelo del NKRJa, l'annotazione può talvolta risultare imprecisa.

(40b) Евмен **велел** архитектору Диаду (...) **создать** машины для сценических эффектов.

(41a) Egli ha un'abilità particolare, sa fare il muso di lepre, e tutti **gli fanno fare** il muso di lepre, e ridono; (De Amicis)

(41b) Он умеет строить замечательную «заячью мордочку». Все **просят его соорудить** эту гримасу и смеются.

Notiamo che il russo, dovendo interpretare il F italiano desemantizzato, aggiunge tratti semantici assenti o impliciti nell'originale, ad es.: *ordinare (velet')*, *chiedere (prosit')*. Vi sono poi casi, molto più rari, in cui al posto dell'infinito il russo impiega un sostantivo (42b). La costruzione mantiene il significato causativo, anche se il comportamento del verbo in questi casi è più simile a quello di verbo di supporto (cfr. par 4.1., gruppo B):

(42a) *Eccolo, il mostro delle leggende la cui sola voce faceva tremare un intero paese!*
(Buzzati, *L'uccisione del drago*)

(42b) *Вот оно, легендарное чудовище, один рев которого нагонял страх на целую деревню.*

2) Preposizioni di causa → *iz-za, ot, po, za (sčet)* (a causa di) + sostantivo (15 occorrenze – 3,08%)

In questi casi il soggetto, che in italiano svolge il ruolo di causatore, è reso da un sintagma preposizionale (introdotto da una delle preposizioni sopracitate) con funzione di complemento di causa:

(43a) *Posso accettare che la tua presenza mi faccia perdere la semifinale di un torneo importantissimo (...)* (Faletti)

(43b) *Я еще могу смириться с тем, что из-за твоего появления вылетела в полуфинале страшно важного турнира (...)*

3) Causativo analitico con congiunzioni → *i (e), sdelat' tak, čtoby, dobitsja čtoby (far sì che)* ecc. (9 occorrenze – 1,85%)

La costruzione causativa italiana è tradotta con costruzioni sintattiche bifrasali, unite da una congiunzione coordinativa o subordinativa con valore causale:

(44a) (...) *e lo pregò di **far uscire** tutti (...)* (Eco, *Il nome della rosa*)

(44b) (...) *и попросил **сделать так, чтобы** все вышли (...)*

Notiamo che non sempre la presenza di *čtoby* realizza una costruzione causativa in russo. La congiunzione *čtoby* introduce una proposizione finale e in molti dei nostri esempi è impiegata per tradurre una costruzione italiana in cui F, a sua volta all'infinito, è retto da una congiunzione finale (generalmente introdotta da *per*). In tutti questi casi il russo traduce la subordinata finale, ma non mantiene il causativo. Il causatore italiano “scompare” dalla subordinata e unico agente dell'azione diviene il causato. In (45b), ad esempio, F + inf. è tradotto con la costruzione modale, non causativa, *moč' perejti* (*poter passare*):

(45a) *Alessandro fece costruire due ponti di barche **per far passare** l'esercito (...)*
(Manfredi)

(45b) ***Чтобы** войско могло **перейти** (...), Александр велел построить два моста из лодок.*

Ciò conferma la tesi di Govorucho (2010a) per cui l'italiano tenderebbe a mantenere la monoreferenzialità del soggetto – *Alessandro* in (45a) – tra principale e subordinata, al contrario del russo che aggiunge un nuovo soggetto alla subordinata (*vojsko*). Secondo Govorucho la costruzione F + inf. in questi casi avrebbe proprio la funzione di garantire la monoreferenzialità che l'italiano preferisce. Molti esempi di questo tipo sono stati qui classificati nel gruppo 5 – “resa non causativa” –, perché di fatto, considerando solo il sintagma F + inf., questo non è reso in russo con un verbo causativo.

In altri casi (46b) *čtoby* introduce un causativo lessicale. Frasi di questo tipo rientrano pertanto nella categoria numero 4:

(46a) *Quando entrarono nella fortezza di Parigi **per far uscire** alcuni loro compagni (...)*
(Eco, *Il nome della rosa*)

(46b) *Когда они ворвались в Парижскую крепость, чтобы **выпустить** несколько (...)*

4) Causativo lessicale (131 occorrenze – 29,95%)

La costruzione analitica italiana è resa in russo con un unico verbo causativo (47b).

(47a) *E secondo lei hanno (...) rischiato anni di galera solo per farsi o **farle fare quattro risate?***». (Camilleri)

(47b) *И по-вашему, кто-то (...) рисковал годы провести в тюрьме, только чтобы самому посмеяться или **вас повеселить?***

(48a) *Ora bisogna che Nelli, finalmente, abbia detto tutto a sua madre, e degli scherni dei primi giorni e di quello che **gli facevan patire** (...)* (De Amicis)

(48b) *В конце концов Нелли, должно быть, всё рассказал своей матери: и как над ним смеялись в первые дни, и как **его мучили** (...)*

A questi esempi abbiamo aggiunto 11 casi in cui i verbi impiegati dai traduttori, in base alle definizioni proposte nei precedenti paragrafi, sarebbero, a nostro avviso, causativi, sebbene l'annotazione semantica del NKRJa non li registri come tali. Ne è un esempio *mučit'* (48b), che ci sembra provochi un "cambiamento di stato" nell'oggetto, tratto che, secondo la definizione di Zolotova, caratterizza un verbo causativo.

5) Resa non causativa (197 occorrenze – 40,53%)

Il più delle volte la costruzione italiana è resa con un singolo verbo, ma a differenza della strategia precedente, non si tratta di un verbo causativo, come in (47b) e (48b). In alcuni casi la frase, oltre a perdere la causatività dell'italiano, conserva poco della struttura sintattica originale (51b):

(49a) *«Ho bisogno di **farmi curare**, sono malato».* (Camilleri)

(49b) *Мне нужно **лечиться**, болен я.*

(50a) *Lo **hai fatto avvelenare** e ora...* (Manfredi)

(50b) *Ты его **отравил**, а теперь...*

(51a) *E dal cimitero si avviava, come **ci ha fatto comprendere Berengario**, in direzione opposta al dormitorio.* (Eco, *Il nome della rosa*)

(51b) ***По показаниям Беренгара**—в сторону, противоположную спальному корпусу, то есть в направлении Храмины.*

Vi sono infine alcuni rari esempi dove F causativo o l'intera frase causativa non vengono per nulla riportati in russo; si tratta di casi che definiremo di "traduzione zero" (52b).

(52a) *"Ma come? ma come? Ma perché **li hai fatti portar via**, cara?"* (Buzzati, *Eppure battono alla porta*)

(52b) *Как же так, дорогая? Зачем?*

Il fatto che in moltissimi casi il causativo italiano venga reso in russo con una forma non causativa, lascia spazio ad alcune considerazioni a nostro avviso rilevanti, che insieme ad altre osservazioni saranno oggetto del prossimo paragrafo.

4.5. Risultati

In conclusione, l'analisi degli esempi dimostra che non esiste un modo univoco per rendere in russo la costruzione causativa F + infinito. Emergono però due tendenze principali: la prima, presente in quasi tutte le strategie di resa (esclusa la prima), è la tendenza ad esprimere un unico soggetto nella frase: se la costruzione causativa italiana mette in scena due o tre attanti che intervengono come soggetti in un enunciato, il russo, nella maggior parte dei casi, indipendentemente dalla strategia traduttiva, tende a ridurne il numero rendendo la frase a soggetto unico. In alcuni casi, il secondo soggetto è sottoposto ad un processo di ellissi sintattica, ma può rimanere sottointeso; in altri casi, scompare totalmente.

La seconda tendenza consiste nel rendere F causativo tramite verbi, o costruzioni di altro tipo, che in russo non risultano causativi (ciò riguarda quasi la metà degli esempi analizzati). Questo fenomeno, oltre a confermare la tesi di Simone e Cerbasi esposta nel

par. 4.2., per cui F non svolgerebbe sempre una funzione propriamente causativa, ci ha spinto a riflettere sulle sue reali funzioni e ad avanzare alcune ipotesi sulla sua frequente ricorrenza nella lingua italiana.

4.5.1. Riduzione degli attanti ed enunciati a soggetto unico della lingua russa

L'analisi dei nostri esempi ha messo in luce una differenza fondamentale tra russo e italiano nell'esprimere una situazione causativa. In italiano frasi come “*il re li ha fatti rinchiudere*”, “*il re li ha fatti fucilare*” implicano la presenza di tre attanti: *il re* causatore, il causato, qui non specificato, e l'oggetto/paziente, qui espresso dal clitico *li*. Un causativo lessicale ridurrebbe invece la distanza tra causatore e oggetto/paziente, togliendo dalla scena il potenziale causato e soggetto logico dell'infinito (Cfr.: “*il re li ha rinchiusi*”, “*il re li ha fucilati*”). Ciò avviene quando abbiamo un infinito transitivo, in presenza del quale l'italiano distingue sintatticamente le due situazioni [(*far fucilare* (qualcuno da altri) vs *fucilare qualcuno* (direttamente)]. Se invece all'infinito vi è un verbo intransitivo, come vedremo anche nel prossimo paragrafo (4.5.2), si aprono due possibilità: 1) gli attanti sono tre, come con predicati transitivi quali *rinchiudere* o *fucilare*. 2) il numero degli attanti è ridotto a due: il causatore, soggetto grammaticale della frase, e il causato, soggetto logico dell'infinito. In molti di questi casi si può quindi generare una situazione ambigua simile a quella del russo (ad es. in “*il re li ha fatti salire*”): come abbiamo visto grazie agli esempi di Padučeva (cfr. par 4.2.), infatti, in russo la distinzione tra causativo diretto e indiretto non è marcata sintatticamente e un causativo lessicale può sottintendere la presenza di un secondo soggetto partecipante all'azione e quindi di 3 attanti. Nel momento in cui il russo traduce la costruzione italiana per mezzo di un causativo lessicale, come *zaperet'* (rinchiudere) o *zastreljat'* (fucilare), riduce apparentemente il numero degli attanti da 3 a 2. In realtà questo avviene solo formalmente, ma non dal punto di vista logico-semantic. La cosa interessante è che questo tipo di ellissi sintattica non ricorre solo con verbi causativi, ma con moltissimi verbi transitivi e intransitivi che potenzialmente possono riflettere l'una e l'altra situazione (ossia con 2 o 3 attanti per i transitivi e con 1 o 2 attanti per gli intransitivi). Abbiamo visto l'esempio di *lečit'sja* (*farsi curare-curarsi*) e di “*otravit*” (*avvelenare*), che qui riproponiamo:

(49b) *Мне нужно лечиться, болен я.*

(50b) *Ты его **отравил**, а теперь...*

Questi verbi non sono classificati come causativi, eppure l'ambiguità si mantiene, tant'è che (49b) traduce la frase originale italiana a 2 attanti “*ho bisogno di farmi curare*” (Camilleri) e (50b) la frase originale a 3 attanti “*l’hai fatto avvelenare*” (Manfredi).

Alla luce di ciò, ci sembra di poter affermare che questa ambiguità rispecchi una preferenza generale della lingua russa ad esprimere un unico soggetto grammaticale in qualunque situazione, causativa e non¹⁹⁸. Abbiamo infatti rilevato che il passaggio da frase a due soggetti italiana a frase a soggetto unico russa avviene nella maggioranza dei casi: 385 su 486 (79,21%). Di fatto, da questa quota si escludono solo le traduzioni con verbo + infinito soggettivo (strategia 1). Vediamo dunque alcuni esempi che testimoniano questo fenomeno:

(53a) *Il movimento fa uscire un rivolo d'acqua dalla bocca semiaperta.* (Faletti)

(53b) *От легкого движения из полураскрытого рта **льется** вода.*

(54a) *Speravo solo che il Teschio **non mi facesse scontare** una penitenza troppo dura.*
(Ammaniti)

(54b) *Оставалось надеяться, что **Череп не придумает** для меня слишком суровое наказание.*

(55a) *Si informò, guardò il modulo, **si fece spiegare** dal Corte.* (Buzzati, *Sette piani*)

(55b) *Узнав, о чем речь, **он** взглянул на бланк и **выслушал** Корте.*

(56a) *Si dovrebbe prenderla a parte e **farle dire** a quattr'occhi la verità.* (Pirandello, *La signora Frola e il signor Ponza suo genero*)

(56b) *Следовало бы отвести ее в сторону и **выяснить** правду с глазу на глаз.*

¹⁹⁸ Ciò conferma la tesi di R. Govoruchо (2010b: 187) per cui il russo “a livello di gerarchia sintattica risulta essere meno coeso e più implicito” rispetto all’italiano, anche se qui l’autore si riferisce soprattutto all’aggiunta nel testo italiano di verbi di percezione, del dire o di altri verbi semanticamente simili. Es: *Эраст Петрович подождет Ø, не последует ли выстрел...* (Акунин) vs *Erast Petrovič per un po’ attese per vedere se sarebbe seguito lo sparo...* (esempi di Govoruchо) (Govoruchо 2010b: 199).

(57a) *Per cagion sua, quante speranze non feci concepire alla mamma: ch'io stessi cioè per metter senno (...)* (Pirandello, *Il fu Mattia Pascal*)

(57b) *Благодаря ей моя мать уже начала надеяться, что я образумился (...)*

Come si può notare, in alcuni casi l'ellissi di un attante non è solo sintattica, ma anche logico-semantica, come in (57b), dove il soggetto sottinteso (*io*) della frase (57a) sparisce completamente. In (53b), invece, il soggetto di (53a) viene trasformato in complemento introdotto da *от*.

4.5.2. Le sotto-funzioni di F causativo

Il fatto che spesso F causativo venga reso in russo con verbi o costruzioni non causative ci ha portato ad interrogarci ulteriormente su eventuali sotto-funzioni che la costruzione F + infinito potrebbe svolgere nel processo comunicativo. Naturalmente, nella maggior parte degli esempi, F + inf è impiegato proprio nel senso di “*far fare qualcosa a qualcuno*”, “*far sì che un soggetto B compia l'azione X o assuma lo stato Y*” e mantiene dunque intatta la sua funzione primaria (58a). Ribadiamo, tuttavia, che in questi esempi la forza causativa varia, così come il grado di agentività e di controllo dei soggetti coinvolti e, anche in caso di forza causativa molto debole (59a), una resa non causativa non compromette la buona riuscita della traduzione (59b):

(58a) (...) *e ti fanno tenere i sandali pure se fa freddo.* (Ammaniti)

(58b) *Еще тебя заставляют ходить в сандалиях, даже если очень холодно.*

(59a) *Mia madre mi condusse questa mattina alla Sezione Baretti a farmi inscrivere per la terza elementare (...)* (De Amicis)

(59b) *Утром мама повела меня в школу Баретти, чтобы записать в третий класс (...)*

L'analisi degli esempi si è rivelata molto utile, quindi, per identificare, oltre all'uso propriamente causativo, due ulteriori funzioni, dove la componente causativa permane, ma sembra non essere più dominante: una prevalentemente stilistica e una sintattico-grammaticale. Infine, abbiamo rilevato casi in cui F + inf. è impiegato in locuzioni fisse, la cui componente causativa è quasi impercettibile. Tra queste citiamo “*far venire la pelle*

d'oca", "farla pagare" (2 occorrenze), "far venire voglia di", "far crescere la barba/i capelli" (2 occorrenze), "farsi prendere da", "far venire l'acquolina in bocca".

Per quanto riguarda la prima sotto-funzione (stilistica), la costruzione F + inf. presenta sì il suo significato causativo intrinseco, ma principalmente sembra essere impiegata con lo scopo di sostituire una variante sintetica e più dotta in italiano. Tale uso prevale nei dialoghi e riguarda un ristretto gruppo di verbi, che non a caso vengono tradotti in russo quasi sempre allo stesso modo per mezzo di un causativo lessicale. In tutti questi casi, infatti, la variante sintetica russa non è stilisticamente marcata. Il caso più dimostrativo è *far vedere* (o *farsi vedere* e, più raramente, *far guardare*) sostitutivo di "mostrare", tradotto in russo nella maggior parte dei casi (17 su 24) con *pokazat'* (*mostrare*) (60b). In un solo caso il traduttore predilige la variante analitica che riprende l'italiano (61b) anche se il secondo verbo è coniugato alla prima persona singolare, anziché all'infinito. Nei restanti 6 esempi, i traduttori rendono la costruzione italiana, rigirando la frase, con un verbo non causativo (62b):

(60a) «**Fammi vedere**» disse Montalbano di colpo interessato. (Camilleri)

(60b) – *А ну-ка покажи, – сказал Монтальбано, вдруг заинтересовавшись.*

(61a) "Vieni qua. **Fammeli vedere**. (Ammaniti)

(61b) – *Подойди. Дай погляжу.*

(62a) *E con la coda di lucertola **fai vedere** tutte le cose intorno come d'argento.* (Eco, *Il nome della rosa*)

(62b) *От испарений ящеричьего хвоста все кругом **становится** как серебряное.*

Tra i causativi motivati stilisticamente, anche se meno frequenti nel nostro corpus, ricordiamo anche:

- *Far sapere*, sostitutivo di "informare", "comunicare".
- *Far venire*, nel senso di "suscitare"
- *Far cuocere*, sostitutivo della variante "cuocere", più rara.

Per verificare e, possibilmente, rafforzare la validità delle nostre osservazioni, abbiamo interrogato il corpus in direzione opposta, selezionando un sotto-corpus di estratti

originali russi allineati alla traduzione italiana, che conta 3.348.735 parole complessivamente. Inserendo nel motore di ricerca il verbo *pokazat'*, abbiamo ottenuto come risultato 326 occorrenze, di cui 82 sono dialoghi dove *pokazat'* ha il significato da noi qui esaminato (questo verbo presenta infatti una certa polisemia: *mostrare, indicare, dimostrare* ecc.). Gli esempi ottenuti dimostrano che *pokazat'* è tradotto in italiano come “*mostrare*” solo in testi narrativi (63b) o in dialoghi formali, in cui i personaggi interagiscono dandosi del “Voi”, e perlopiù datati (i romanzi del XIX secolo contenuti nel corpus)¹⁹⁹ (64b). “*Mostrare*” non ricorre in dialoghi informali e colloquiali, dove la variante analitica F + inf. è preferita in traduzione (65b), (65c):

(63a) (...) *И парень **показал** свою крученую папироску.* (Kurkov)

(63b) (...) *e il ragazzo **mostrò** la sigaretta arrotolata.*

(64a) – *А где ваш больной? – спросил Андрей Ефимыч. У меня в больнице. Мне уж давно хотелось **показать** вам...* (Čechov, *Palata № 6*)

(64b) «*E dov'è il vostro malato?*» domandò *Andrèj Efìmyč*. «*Da me, nell'ospedale. Già da tempo ve lo volevo **mostrare**...*

(65a) – *Вот я тебе сейчас **покажу**. Как твоя фамилия?* (Šalamov)

(65b) *Adesso ti **faccio vedere**. A proposito, come ti chiami?* (Rapetti)

(65c) «*Adesso ti **faccio vedere!** Com'è che ti chiami?*». (Binni)

Come seconda controprova, abbiamo ricercato un'altra tra le sopracitate espressioni, sempre in direzione italiano-russo, in genere molto diffusa: *far sapere*. Nella maggior parte dei casi *far sapere* traduce ciò che in russo è espresso con un singolo verbo - *soobščit'* (*informare, comunicare*), *rasprostranjat'sja* (lett: *diffondersi*), *izveščat'* (*annunciare, informare*), in alcuni casi anche *skazat'* (*dire*) -, il cui equivalente letterale italiano risulterebbe troppo formale, perlomeno in un dialogo. Inoltre, alcuni di questi verbi non sono causativi (secondo l'annotazione semantica del corpus, nessuno di questi lo è). Quanto avviene nel passaggio dall'italiano al russo si verifica dunque anche nella direzione opposta: una forma italiana causativa è associata ad una forma russa non

¹⁹⁹ Osservazioni di questo tipo sono possibili solo dove siano disponibili informazioni extra-linguistiche precise. Ecco perché in alcuni casi un corpus tradizionale, anziché un web corpus, può rivelarsi molto utile per garantire un maggior grado di accuratezza durante la ricerca.

causativa. Ciò testimonia che anche in italiano espressioni di questo tipo sono entrate talmente nell'uso comune, da aver perso le peculiarità originarie della costruzione causativa, confermando che il motivo della loro ampia diffusione sembrerebbe attribuibile a motivazioni perlopiù stilistiche.

La seconda sotto-funzione, da noi definita sintattico-grammaticale, consiste nel rendere transitiva una costruzione intransitiva che, come noto, non può reggere un complemento oggetto. Abbiamo già menzionato questa proprietà di F causativo, citando La Fauci e Mirto (2003), ma qui la illustreremo con i nostri esempi, dimostrando che essa spesso non solo accompagna la funzione propriamente causativa, ma addirittura prevale su di essa. Ciò accade quando il causato è un oggetto inanimato²⁰⁰. In questi casi, infatti, F + inf. sembra svolgere la mera funzione di predicato complesso transitivo, seguito da un complemento oggetto (66a). La forza causativa risulta qui minima e il causato subisce senza che vi sia alcun tipo di agentività. In questi casi, nella traduzione russa, si attestano due tipi di strategie di resa: tramite un causativo lessicale transitivo con significato analogo (66b), ove presente, oppure tramite una “resa non causativa”, in cui il causato italiano diventa soggetto grammaticale (67b). A differenza dell'italiano, abbiamo rilevato che il russo il più delle volte riesce a rendere la transitività tramite un'unica unità lessicale.

(66a) *Era diventato celebre perché a Fossanova era riuscito a **far discendere** un corpo lungo una scala a chiocciola.* (Eco, *Il nome della rosa*)

(66b) *Он добыл свою славу в Фоссанова, сумев **спустить** мертвеца по винтовой лестнице.*

(67a) *Lui la prese e si incise con la punta la pelle del braccio, **facendone colare** un rivolo vermiglio.* (Manfredi)

(67b) (...) *а Александр надрезал лезвием кожу на руке, так что **потекла кровь**.*

Notiamo, tuttavia, che quando i soggetti coinvolti, causatore e causato, sono animati, la funzione causativa rimane preponderante (68a).

Osservando l'esempio (68) ricordiamo inoltre un aspetto molto importante per il confronto tra russo e italiano, menzionato nel par. 4.2., ossia la distinzione tra causativo

²⁰⁰ Fanno eccezione due casi con oggetto animato riscontrabili con il verbo *sembrare* e *assomigliare* (*far sembrare*, *far assomigliare*), poiché, seppur animato, il causato non solo è paziente, ma addirittura non prova alcuna sensazione o sentimento.

diretto e indiretto. Indipendentemente dalla presenza o meno di soggetti animati, tale distinzione in italiano non è sintatticamente marcata per verbi intransitivi: con i verbi transitivi il causativo diretto (12) e indiretto (13) sono invece espressi rispettivamente da un causativo sintetico (lessicale) e da uno analitico - cfr. gli esempi del par. 4.2: (12) *A ha rinchiuso B* vs (13) *A ha fatto rinchiusere B* -, mentre se il verbo è intransitivo, come in (68a), la frase resta semanticamente ambigua;

(68a) *Sì, ricordo una storia di re Marco che doveva condannare Isotta la bella e stava facendola salire sul rogo (...)* (Eco, *Il nome della rosa*)

A ha fatto salire sul rogo B può infatti significare sia che A ha direttamente condotto al rogo B, sia che A ha ordinato a un soggetto C di compiere tale azione (la stessa ambiguità è presente nel causativo lessicale russo, cfr. esempio di Padučeva: *Я сишла себе юбку*), di conseguenza, i problemi di interpretazione possono sorgere sia passando dal russo all'italiano sia dall'italiano al russo. Nell'esempio presente nel corpus il traduttore russo interpreta la costruzione causativa come indiretta, ma è solo una delle soluzioni possibili:

(68b) *Да-да, и я помню, в повести о короле Тристане...Когда он придумывал казнь для Изольды прекрасной и велел сжечь ее на костре (...)*

Diversamente da (68a), gli esempi con attanti inanimati presentano meno tale ambiguità, che è sciolta dalla semantica del verbo, ma soprattutto dal contesto, per via dello status inanimato del causatore. Cfr. (69a), dove l'interpretazione del causativo indiretto non è ammessa:

(69a) *Le radici avevano smantellato i gradini di pietra e fatto crollare il parapetto.*
(Ammaniti)

(69b) *Его корни раскрошили камень ступеней и обрушили parapet.*

Un'ultima possibile spiegazione dell'uso più frequente di F ci sembra legata al “principio di economia”²⁰¹. A parità di situazione extralinguistica, ossia in mancanza di situazioni

²⁰¹ Secondo André Martinet (1974: 197-200), “si può concepire l'evoluzione linguistica come governata da un'antinomia permanente fra la necessità di comunicazione dell'uomo e la tendenza che egli ha di ridurre al minimo la sua attività mentale e fisica”.

stilisticamente marcate in cui l'italiano richiede l'uso di una lingua più alta, la costruzione F + inf. risponde in un certo senso a questa “legge del minimo sforzo”, secondo la quale, “quando si parla per essere capiti, si spende energia nella misura in cui si può sperare che quel che si dice venga compreso” (Martinet: 1974: 199). Nel par. 4.3.1., abbiamo visto come anche in latino *facio* venisse talvolta preferito a verbi come *iubeo*, perché consentiva di evitare la ripetizione della congiunzione *ut*.

Più precisamente, la costruzione F + inf. concorre in italiano con altre costruzioni analitiche causative composte da verbi quali *costringere/obbligare*, *chiedere*, *aiutare* + infinito. Il loro uso al posto di F ovvierebbe, tra l'altro, ai problemi di ambiguità semantica che talvolta accompagnano il verbo F, che in diversi contesti sostituisce ciascuno di questi verbi, semanticamente diversi tra loro. Eppure, interrogando il nostro corpus, questi verbi sono molto meno frequenti (30 occorrenze per *costringere*; 16 occorrenze per *obbligare*; 22 occorrenze per *chiedere*; 9 occorrenze per *aiutare* rispetto alle 486 occorrenze per F + infinito). La preponderanza di F, che introducendo direttamente l'infinito semplifica le costruzioni degli altri verbi in cui una preposizione deve necessariamente precedere l'infinito (*costringere/obbligare a*, *chiedere di*, *aiutare a* + infinito), è dunque la variante più economica.

Inoltre, tornando agli esempi con verbo intransitivo, dove F, come abbiamo detto, introduce anche una trasformazione sintattica, notiamo che anche in questo caso sembra intervenire il principio di economia: rendendo la costruzione transitiva, il soggetto logico dell'infinito può essere espresso come un oggetto diretto. Altrimenti, si dovrebbe rendere questo soggetto logico anche soggetto grammaticale formando quindi una frase complessa, con due soggetti e due predicati come in (71), anziché semplice, quale invece è la costruzione con F + inf. (70a):

(70a) *Ancora una pausa, lunga abbastanza da **far sembrare** caduta la comunicazione.*

(Faletti)

(71) *Ancora una pausa, lunga abbastanza da **far sì che** la comunicazione **sembrasse** caduta (esempio nostro)*

Il russo non dispone dei mezzi dell'italiano e in traduzione opta per la seconda strategia:

(70b) *И снова молчание, настолько долгое, **что казалось**, будто связь прервалась.*

4.6. Osservazioni conclusive

In questo capitolo abbiamo innanzitutto voluto mostrare che l'analisi *corpus-based*, mettendo a disposizione molti esempi autentici, è utile non solo per confermare ipotesi o esemplificare temi già studiati, ma anche come spunto per nuove osservazioni. In particolare, confrontando originali italiani e traduzioni russe (e in parte viceversa), abbiamo fatto luce sia sulle funzioni della costruzione italiana F + infinito, sia su di un interessante aspetto che differenzia il russo dall'italiano, ossia la tendenza del russo ad esprimere un unico soggetto anche in presenza di due soggetti logici in italiano.

Per trarre conclusioni definitive sui fenomeni riscontrati, sono necessari ulteriori approfondimenti, sfruttando possibilmente corpora di maggiori dimensioni e disambiguati. In particolare, sarà necessario:

- esplorare il tema dell'enunciato a soggetto unico, anche in altre strutture a doppio soggetto italiane, diverse da F causativo;
- approfondire l'implicito del causativo lessicale [e di altri verbi non causativi. Cfr.: *otravit', lečit'sja*, negli esempi (49b) e (50b)], il quale può essere sia diretto che indiretto e quindi sottintendere un ulteriore attante. Tale aspetto ha particolari ricadute sulla traduzione dal russo all'italiano, dal momento che l'italiano impone di risolvere quest'ambiguità;
- verificare il comportamento di F causativo, sempre in ottica contrastiva con il russo, anche in testi non dialogici e stilisticamente differenziati (cosa che il CP italiano-russo al suo stato attuale ancora non permette in modo statisticamente rilevante) in modo da verificare altre eventuali regolarità dell'uso legate alla variazione tipologica dei testi.

Bibliografia

- Apresjan Ju. D. (1967). *Eksperimental'noe issledovanie semantiki russkogo glagola*. Moskva: Nauka.
- Apresjan Ju. D. (1995). *Izbrannye trudy, vol. 1, Leksičeskaja semantika (sinonimičeskie sredstva jazyka)*. Moskva: Jazyki Russkoj Kul'tury.
- Apresjan V. Ju. (2013). *Semantika emocional'nych kauzativov: status kauzativnogo komponenta*. In Komp'juternaja lingvistika i intellektual'nye tehnologii: po materialam ežegodnoj meždunarodnoj konferencii "Dialog", 12, Moskva: RGGU, pp. 43-57.
- Arkad'ev P. M., Letučij (2009). *Tranzitivnost' implikacii i estestvennyj jazyk: paradoksy kauzativnykh konstrukcij v tipologičeskom osveščeenii*. Seminar "Nekotorye primečanija matematičeskich metodov v jazykoznanii", MGU, Moscow, zasedanie № 135. (Tratto da: <http://www.inslav.ru/ob-institute/sotrudniki/279-peter-arkadiev>).
- Bajric S. (2008). *Le verbe faire en français contemporain: syntaxe et sémantique*. In Contemporary linguistics, Vol. 66 N. 2, Zagreb: Croatian Philological Society, pp. 143-197.
- Battaglia S. (1961). *Grande Dizionario della lingua italiana*. Torino: Unione tipografico-editrice torinese.
- Benigni V., Cotta Ramusino P. (2011). *Le costruzioni con verbo supporto in russo: il caso di delat'*. Studi Italiani di Linguistica Teorica e Applicata, anno XL, 1, pp. 7-26.
- Bertinetto P. M. (1991). *Il verbo*. In Renzi L., Salvi G. (a cura di), *Grande grammatica italiana di consultazione*. Il Mulino: Bologna, pp. 13-161.
- Bortolini U., Tagliavini C., Zampolli A. (1972). *Lessico di frequenza della lingua italiana contemporanea*. Milano: Garzanti.
- Brambilla Ageno F. (1964). *Il verbo nell'italiano antico: ricerche di sintassi*. Milano-Napoli: Ricciardi.
- Cantarini S. (2004). *Costrutti con verbo supporto. Italiano e tedesco a confronto*. Bologna: Patròn editore.

- Cerbasi (1998). *Le costruzioni causative in italiano, spagnolo e portoghese*. In Ramat P., Roma E. (a cura di), *Sintassi storica*. Atti del 30° Congresso Internazionale della Società di Linguistica Italiana, Pavia. Roma: Bulzoni, pp. 457-468.
- Cholodovič A. A. (ed.) (1969). *Tipologija kauzativnich konstrukcij. Morfoložičeskij kauzativ*. Leningrad: Nauka.
- Comrie B. (1989). *Language Universals and Linguistic Typology*. Chicago: The University of Chicago Press.
- Consales I. (2006). *Fra lessico e grammatica. Il problema dei verbi dà(re) e fà(re) nel Vocabolario del romanesco contemporaneo*. In Cresti, E. (a cura di), *Prospettive nello studio del lessico italiano*, Atti SILFI 2006. vol. I, Firenze: FUP, pp. 79-83.
- Cuj V., Bočina T. G. (2014). *Kauzativnaja konstrukcija v ruskom i kitajskom jazykach*. *Filologija i Kul'tura*, № 3(37), pp. 7-12.
- Dadueva A. E. (2011). *Obšaja charakteristika kauzativnyh glagolov*. *Vestnik SibTUTI, Sankt-Peterburg*, № 2, pp. 76-81.
- D'agostino E., Guglielmo D. (2012). *Varianti di verbi supporto in italiano*. In Ferreri S. (a cura di), *Lessico e lessicologia*. Atti del 44° congresso internazionale di studi della Società di linguistica italiana, Viterbo. Roma: Bulzoni, pp. 143-156.
- DeLancay S. (1984). Notes on agentivity and causation. In *Studies in language*, 8, 2, pp. 181-213.
- Delcorno C. (1970). *Fare*. In *Enciclopedia Dantesca*. Roma: Istituto dell'Enciclopedia italiana fondata da Giovanni Treccani, pp. 794-803.
- Dolinina I. B. (1991). *Refleksivnost' i kauzativnost' (kategorial'naja semantika refleksivnyh konstrukcij, sootnositel'nyh s kauzativnymi konstrukcijami)*. In Bondarko A. V. (ed.), *Teorija funkcional'noj grammatiki. Personal'nost'. Zalogovost'*. Sankt-Peterburg: Nauka, pp. 327-345.
- Gilquin G. (2010). *Corpus, cognition and causative constructions*. Amsterdam, Philadelphia: John Benjamins Publishing.
- Govorucho R. A. (2010a). *Kogezija v ital'janskom i ruskom tekstach: passivnye i kauzativnye modeli*. In *Lingvistika i metodika prepodavanija inostrannyh jazykov*. Institut jazykoznanija RAN: Moskva.
- Govorucho R. A. (2010b). *Qualche osservazione sull'uso dei verbi proposizionali in italiano e in russo*. In "L'Analisi Linguistica e Letteraria", Anno XVIII, 2/2010,

- Facoltà di Scienze Linguistiche e Letterature Straniere, Università Cattolica del Sacro Cuore, Milano, pp. 187-200.
- Gross M. (1968). *Grammaire transformationnelle du français. Syntaxe du verbe*. Paris: Larousse.
- Guasti M. T. (1996). *Semantic restrictions in Romance causatives and the incorporation approach*. In *Linguistic Inquiry*, Vol. 27. No. 2 pp. 294-313.
- La Fauci N., Mirto I. M. (2003). *Fare. Elementi di sintassi*. Pisa: Edizioni ETS.
- La Fauci N. (2009). *Compendio di sintassi italiana*. Bologna: Il Mulino.
- Lance LeBlanc (2010). *The polysemy of an "empty" prefix: a corpus-based cognitive semantic analysis of the Russian verbal prefix po-*. Chapel Hill: University of North Carolina.
- Letučij A. (2012). *Time reference in Russian causative constructions*. *Oslo studies in languages*, Vol. 4, No. 1, pp. 31-51.
- Martinet A. (1974). *Elementi di linguistica generale*. Roma: Laterza.
- McEnery T., Tono Y., Xiao R., (2006). *Corpus-based Language Studies: An Advanced Resource Book*. New York: Routledge.
- Mel'čuk I. A. (1999). *Opyt teorii lingvističeskich modelej "smysl – tekst"*. *Semantika, sintaksis*. Moskva: Jazyki russkoj kul'tury.
- Mirto I. M. (2003). *Che fare? Analisi di costrutti di un verbo critico in italiano*. In Marcellesi M. G., Rocchetti A. (a cura di), *Il verbo italiano: studi diacronici, sincronici, contrastivi, didattici*. Atti del 35° Congresso internazionale di studi della Società di linguistica italiana, Parigi, pp. 277-292.
- Moreno J. C. (1993). *"Make" and the semantic origin of causativity: a typological study*. In Comrie B., Polinskij M. (eds.), *Causative and Transitivity*. Amsterdam-Philadelphia: John Benjamins Publishing Company, pp. 153-164.
- Nedjalkov V. P., Sil'nickij G. G. (1969). *Tipologija kauzativnich konstrukcij*. In Cholodovič A. A. (pod red.), *Tipologija kauzativnich konstrukcij. Morfologičeskij kauzativ*. Leningrad: Nauka, pp. 5-19.
- Padučeva E. V. (1996). *Neopredelennost' kak semantičeskaja dominant russkoj jazykovoju kartiny mira*. In *Problemi di morphosintassi delle lingue slave*, v. 5. *Determinatezza e indeterminatezza nelle lingue slave*. Padova: Unipress, pp. 163-186.

- Padučeva E. V. (1997). *Verb Categorization and the Format of a Lexicographic Definition (Semantic Types of Causative Relations)*. In Wanner L. (ed.), *Recent trends in meaning-text theory*. Amsterdam/Philadelphia: John Benjamins, pp. 61-74.
- Padučeva E. V. (2001). *Kauzativnyj glagol i dekauzativ v ruskom jazyke*. Russkij jazyk v naučnom osveščenii, № 1, pp. 52-79.
- Padučeva E. V. (2003). *Is there an "Anticausative" component in the semantics of decausatives?* In *Journal of Slavic Linguistics*, 11, 1, pp. 173-198.
- Pelo A., Consales I. (2003). *Fare "vicario", "fare + N", "fare + V". Per un'analisi del verbo fare nell'italiano antico*. In Giacomo-Marcellesi M., Rocchetti A. (a cura di), *Il verbo italiano*. Atti del 35° Congresso Internazionale di Studi della Società di Linguistica Italiana, Parigi. Roma: Bulzoni, pp. 43-66.
- Podlesskaja V. I. (1993). *Causatives and causality: towards a semantic typology of causal relations*. In Comrie B., Polinskij M. (eds.), *Causative and Transitivity*. Amsterdam-Philadelphia: John Benjamins Publishing Company, pp. 165-176.
- Ponchon T. (1994). *Sémantique lexicale et sémantique grammaticale: le verbe faire en français médiéval*. Genève: Droz.
- Robustelli C. (1993). *Indagine diacronica sul costrutto latino facio + infinito*. In *Studi e saggi linguistici*, XXXIII, pp. 125-189.
- Robustelli C. (1994). *Il costrutto latino 'fare' e infinito nell'italiano dal 1400 al 1800*. In *Studi e saggi linguistici*, XXXIV, pp. 151-203.
- Robustelli C. (1995). *La costruzione di Fare con l'Infinito in Italiano moderno*. In *Studi e saggi linguistici*, XXXV, pp. 199-272.
- Rylov Ju. A. (2003). *O semantičeskich dominantach v jazykovej kartine mira. Na materiale ital'janskogo i ruskogo jazykov*. *Vesnik VGU, Serija Gumanitarnye Nauki*, 1, pp. 80-94.
- Rylov Ju. A. (2006). *Aspekty jazykovej kartiny mira: ital'janskij i ruskij jazyki*. Moskva: Gnozis.
- Salvi G. (1988). *La frase semplice*. In Renzi L. (a cura di), *Grande grammatica italiana di consultazione*, vol. I, Bologna: Il Mulino, pp. 29-113.
- Serianni L. (1991). *Grammatica italiana: italiano comune e lingua letteraria*. Torino: UTET.

- Shibatani, M. (1976). *The grammar of causative constructions: A conspectus*. In Shibatani, M. (ed.), *The Grammar of Causative Constructions*. New York: Academic Press, pp. 1–40.
- Simone R., Cerbasi D. (2001). Types and diachrony of Romance causatives. In *Romanische Forschungen*, 113, pp. 441-473.
- Skytte G. (1976). *I costrutti infinitivi con i verbi fattitivi e con i verbi di percezione*. In *Studi di Grammatica Italiana*, V, Firenze, pp. 355-400.
- Skytte G. (1978). *Il cosiddetto costruito dotto di accusativo con l'infinito in italiano moderno*. In *Studi di Grammatica Italiana*, VII, Firenze, pp. 281-315.
- Skytte G. (1983). *La sintassi dell'infinito in italiano moderno*. Vol. I. København: Munksgaards forlag.
- Song J. J. (1996). *Causatives and causation: a universal-typological perspective*. New York: Longman.
- Steksova T. I. (2008). *Polusub'ektnaja kauzativnaja konstrukcija s predikatom delat'/sdelat'*. In Mustajoki A., Kopoteva M. V., Birjulina L. A., Protasovoj E. Ju. (pod red.), *Instrumenty rusistiki: korpusnye podchody*, Slavica Helsingiensia, 34, pp. 332-342.
- Sumbatova N. (1993). *Causative constructions in Svan: further evidence for role domination*. In Comrie B., Polinsky M. (eds.), *Causatives and transitivity*. Amsterdam: Benjamins, pp. 253-270.
- Šorochova I. A. (2007). *Aspekty semantiki kauzativnych glagolov*. In *Vestnik Čeljabinskogo gosudarstvennogo universiteta*, 8, pp. 139-145.
- Talmy L. (2001). *Towards a cognitive semantics. Concept Structuring Systems*. Vol I. Cambridge, London: MIT Press.
- Titov V. T. (2004). *Kvantitativnaja leksikologija ital'janskogo jazyka*. In *Vestnik BGU, Serija "Lingvistika i mežkul'turnaja komunikacija"*, № 2, pp. 3-9.
- Wierzbicka A. (1988). *The semantics of grammar*. Amsterdam-Philadelphia: John Benjamins Publishing Company.
- Wierzbicka A. (1996). *Semantic primes and universals*. New York: Oxford University Press.
- Zolotova G. A. (1973). *Očerki funkcional'nogo sintaksisa russkogo jazyka*. Moskva: Nauka.
- Zolotova G. A. (1982). *Kommunikativnye aspekty russkogo sintaksisa*. Moskva: Nauka.

CAPITOLO 5

LA TRADUZIONE IN ITALIANO DEI DIALOGHI DI AKAKIJ AKAKIEVIČ NE "IL CAPPOTTO" DI N.V. GOGOL': UN CONFRONTO TRA 15 VERSIONI

“Нужно знать, что Акакий Акакиевич изъяснялся большею частью предложениями, наречиями и, наконец, такими частицами, которые решительно не имеют никакого значения. Если же дело было очень затруднительно, то он даже имел обыкновение совсем не оканчивать фразы, так что весьма часто, начавши речь словами: «Это, право, совершенно того...» — а потом уже и ничего не было, и сам он забывал, думая, что все уже выговорил.” (Gogol’ 1938: 149)²⁰².

Bisogna sapere che Akàkii Akàkievic si spiegava, nove volte su dieci, per via di preposizioni, avverbi, e in generale mediante particelle prive di senso. Se poi l’argomento era molto scabroso, allora il lasciare a mezzo la frase diveniva quasi una regola; così molto spesso, se iniziava il discorso con i suoi «ora, appunto, in fede mia, ecco, in quanto...» — ecco che si arenava là, e se ne scordava lui stesso, credendo invece di essersi già pronunziato esaurientemente (Traduzione di C. Rebora, 1922).

Il brano qui proposto, seguito dalla traduzione italiana di Clemente Rebora del 1922²⁰³, presenta uno dei tratti caratteristici della parlata di Akakij Akakievič, il povero impiegato di San Pietroburgo, protagonista di uno dei più celebri racconti di Nikolaj Gogol’, *Il Cappotto*²⁰⁴, “una macchietta trascurabile, un impersonale strumento di fatica congegnato

²⁰² L’edizione di riferimento è contenuta in Gogol’ N.V. (1937-1952) *Polnoe sobranie sočinenii v četyrnadcati tomach*. Leningrad: Izadel’stvo Akademii nauk SSSR, 1938, T. 3, *Povesti*, pp. 139-174.

²⁰³ Per la traduzione di brani esemplificativi del racconto abbiamo deciso di proporre la versione di C. Rebora, la prima presente nel nostro corpus e una delle più originali.

²⁰⁴ La vicenda di Akakij Akakevič è tra le più note della letteratura russa: questo anonimo e insignificante copista è costretto dal gelo di Pietroburgo a sostituire il suo vecchio e liso pastrano con un cappotto nuovo, che diventerà poi, fin dal momento della sua ideazione, la sua unica ragione di vita. Con la decisione di farsi confezionare un nuovo cappotto, iniziano per il misero impiegato mesi di sacrificio e digiuno per racimolare il denaro necessario; finalmente si materializza l’idea del cappotto, di cui - dice Rebora - Akakij s’innamora; ma il godimento dura poco e in una notte buia e gelata, di ritorno da una festa organizzata dai

apposta nell'ordigno sociale, per quella sua mansione chiusa e circolare" (Rebora 1990: 75-76)²⁰⁵. La vita di questo "piccolo uomo" non è triste o malinconica, potrebbe essere meglio definita "mesta" (Rebora 1990: 81), pacata, priva di eccessi, caratterizzata da pochi elementi essenziali: un lavoro, a cui si dedica con meticolosa precisione, un letto per dormire, del cibo e un logoro cappotto per difendersi dal gelo degli inverni del nord, "un aspro nemico di tutti coloro che riscuotono suppergiù quattrocento rubli l'anno"²⁰⁶.

Il linguaggio di Akakij Akakievič pare rifletterne la personalità: introverso e solitario, egli si distacca dal mondo esterno, con cui, probabilmente, è ormai incapace di comunicare, come mostrano le prime parole da lui pronunciate nel racconto:

(1) — *А я вот к тебе, Петрович, того...*

Le frasi di Akakij sono caratterizzate da due elementi: l'uso, quasi ossessivo, di particelle o "parole discorsive", da un lato, ed ellissi, dall'altro (si nota in particolare la tendenza ad omettere il verbo). Si tratta di fenomeni diffusi nella lingua parlata russa, ma le particelle sono qui impiegate in modo esasperato, rendendo spesso gli enunciati, come ammette anche la voce narrante, "privi di senso". La nostra analisi prenderà dunque in considerazione questi elementi e confronterà la loro resa da parte dei traduttori nelle versioni italiane presenti nella versione da noi ampliata del CP russo-italiano del NKRJa. La scelta del tema ha lo scopo, innanzitutto, di mostrare l'utilità di un'analisi *corpus-based* che metta a confronto molte traduzioni di un unico testo fonte e che possa avere implicazioni nell'analisi linguistica contrastiva e negli studi traduttivi. Non ci soffermeremo dunque sugli aspetti prettamente letterari legati allo stile gogoliano e alla sua poetica, sebbene l'analisi linguistica *corpus-based* di un testo letterario possa essere molto produttiva nel discernere gli elementi tipici dello stile di un autore²⁰⁷. Anche le

colleghi per celebrare il nuovo acquisto, Akakij viene derubato del suo cappotto. Vani saranno i tentativi di farsi giustizia attraverso autorità o "personaggi importanti", e dopo alcuni giorni di febbre delirante, dovuta sì al freddo, ma anche alle condizioni pietose del suo animo, Akakij Akievič si spegne. Il suo spettro però inizierà ad aggirarsi per le vie della città alla ricerca del suo cappotto, terrorizzando chi lo aveva tormentato in vita.

²⁰⁵ La traduzione di Rebora viene completata nel 1919 e nel 1920 scrive le "Annotazioni". Entrambe saranno poi pubblicate nel '22, ma l'edizione a cui faremo riferimento è quella contenuta in Gogol' (1990).

²⁰⁶ Così Rebora (1922) traduce l'originale gogoliano: "sil'nyj vrag vsech, polučajuščich četyresta rublej v god žalovan'ja ili okolo togo".

²⁰⁷ Come ricorda Biber (2011: 16) fornendo un'ottima rassegna, sono numerosi gli studi letterari degli ultimi anni - generalmente raggruppabili sotto il cappello comune di *corpus stylistics studies* - che fanno uso di metodologie *corpus-based* per analizzare lo stile di un autore, identificandone il lessico, le collocazioni e le strutture sintattiche ricorrenti. Tra questi ricordiamo, ad esempio, B. Fischer-Starcke (2010), che applica

tante interpretazioni²⁰⁸ dedicate a quest'opera non rientreranno nell'oggetto del nostro lavoro, mentre prenderemo in considerazione le sue traduzioni in lingua italiana.

Nel primo paragrafo, infatti, daremo alcune informazioni sulle traduzioni italiane del *Cappotto* e, in particolare, su quelle da noi scelte per il CP. Seguirà una breve introduzione al nostro corpus di esempi (5.2.) e agli elementi linguistici oggetto della nostra analisi: le parole discorsive (5.2.2.) e l'ellissi sintattica (5.2.3.). Nel paragrafo 5.3. presenteremo il confronto e l'analisi delle quindici traduzioni italiane: essendo il nostro fine descrittivo, ci asterremo dall'esprimere giudizi sulla buona o cattiva riuscita delle traduzioni analizzate.

5.1. Le traduzioni italiane del *Cappotto*

Il Cappotto ha suscitato evidente interesse a partire dalla sua pubblicazione negli anni Quaranta dell'Ottocento fino ai giorni nostri, e non solo in Russia. In una recensione di Ettore Lo Gatto alla traduzione di C. Reborà del 1922 pubblicata sulla rivista «L'Italia che scrive»²⁰⁹, si legge:

«Che dal “Cappotto” di Gogol sia uscita tutta la letteratura russa moderna, è ormai affermazione comune nella critica letteraria russa e sta ad indicare non solo che col “Cappotto” in un certo senso comincia la letteratura russa moderna, ma che lo spirito che animò Gogol nella creazione del suo immortale Akakij Akakievic è lo spirito che ha immortalata tutta la letteratura russa. Basti dire di Dostojevskij, a noi più noto, il quale forse senza l'impulso dato alla sua “Povera

il metodo *corpus-based* in un'indagine sui romanzi di Jane Austin, e M. Hory (2004) e M. Mahlberg (2007) che allo stesso modo analizzano lo stile di Charles Dickens.

²⁰⁸ Secondo l'analisi esposta da B. Ejchenbaum in *Kak sdelana “Šinel” Gogolja*, una delle più note, (Ejchenbaum 1919, tratto da “*Poetika*” (<http://www.opojaz.ru/manifests/kaksdelana.html>) (dicembre 2016) il *Cappotto* darebbe adito a due linee interpretative: una prima, che vede Akakij Akakievic come il prototipo dell'uomo piccolo (*malen'kij čelovek*) (Zyrjanov 2004: 124) e che si sarebbe riflessa anche in opere dostoevskiane come “*Bednye ljudi*”, “*Dvojniki*”, “*Unižennye i oskorblennye*”; la seconda linea, secondo Ejchenbaum preponderante, metterebbe invece in evidenza l'elemento “grottesco e umoristico”, espresso dalla voce narrante che, di fatto, si prende gioco del povero impiegato (cfr. anche Padučeva 1997).

²⁰⁹ La rivista, fondata nel 1918 da A. F. Formiggini, ebbe fin dal suo esordio una rubrica intitolata *Letterature straniere in Italia* in cui trovava spazio anche la rassegna delle traduzioni dal russo pubblicate in Italia in quegli anni – anni Venti e Trenta (la rivista cessò il suo operato nel 1938). L'inizio della collaborazione di Lo Gatto, a partire dal 1920, segnò un momento importante per la testata e le pagine dedicate alla letteratura russa divennero tra le più curate (Tortorelli 1996: 91).

gente” dalla piccola tragedia di Akakij Akakievič, non sarebbe assurdo così rapidamente alla gloria» (Lo Gatto 1922: 151).²¹⁰

In Italia, sebbene l’interesse nei confronti di Gogol’ (e in generale della letteratura russa) sia emerso relativamente tardi rispetto ad altri Paesi europei²¹¹, *Il Cappotto* è l’opera russa più tradotta: nell’arco di un intero secolo - dal 1903 al 2012 - si sono susseguite almeno 31 versioni. Riportiamo l’elenco più completo di cui siamo a conoscenza²¹² (Inkova 2014: 55-56):

1. *Il mantello*, traduzione di G. Loschi. Udine, Tipografia del patronato, 1903 (ristampa Firenze, Rassegna Nazionale, 1918).
2. *L’uniforme*, in Nikolaj Vasilevič Gogol, *Novelle* traduzione e prefazione di D. Ciampoli. Milano, Istituto Editoriale Italiano (Gli immortali e altri massimi scrittori. Ser. 2), 1916.
3. *Il cappotto*, traduttore anonimo. Milano, Casa Editrice Sonzogno, 1919.
4. *Il mantello*, in *Novelle russe*, v. 1, a cura di C. Alvaro. Milano, R. Quintieri, 1920.
5. *Il cappotto*, traduzione di C. Reborà. Milano, Il Convegno editoriale, 1922 (ristampa con una nota di P. Giovannetti 1992, 1996, 2010, Milano, Feltrinelli).
6. *Il cappotto*, in *Taras Bulba. Il cappotto*, traduzione di G. Bergamino. Torino, Casa editrice A.B.C., 1932.
7. *Il pastrano*, traduzione di E. Carafa D’Andria. Torino, UTET, 1937.
8. *Il mantello*, in: *Racconti di Pietroburgo*, traduzione di T. Landolfi. Torino, Rizzoli Editore, 1941. Torino: Einaudi, 1995
9. *Il cappotto*, in: *Opere*, vol. 1, traduzione di N. Bavastro. Milano, A. Corticelli, 1944.

²¹⁰ Gogol’ N. V., *Il cappotto*, traduzione di Reborà C. Recensione di: Lo Gatto E., in “L’Italia che scrive”, V, 8 (agosto) 1922, p.151. Relativamente all’influsso di Gogol’ sulla letteratura russa, Lo Gatto fa riferimento alla celebre frase attribuita ora a Dostoevskij ora a Turgenev “Noi tutti siamo usciti dal cappotto di Gogol’” (cfr. Zyryanov 2004: 123 e De Michelis nell’introduzione da lui curata de *La Mantella*, tradotta da N. Marcialis del 1991).

²¹¹ Come riporta Inkova (2014: 47), la prima traduzione italiana di Gogol’ (dell’opera *Taras Bulba*) compare nel 1877, mentre in Francia la prima versione risale già al 1844. Pala ricorda invece che le prime traduzioni integrali di opere di rilievo come *I demoni*, *I fratelli Karamazov* e *Anna Karenina* sono giunte in Italia solo a metà degli anni ’20, più di quarant’anni dopo rispetto alle versioni francesi (Pala 2009: 34).

²¹² Inkova aggiunge una versione di E. Carafa D’Andria del 1960, ripubblicata con il titolo “*Il Cappotto*”. È proprio questa del 1960 la versione contenuta nel nostro corpus, ma confrontandola con alcuni brani riportati da Pala tratti dalla versione del ’37, deduciamo che la ristampa non abbia subito variazioni se non nel titolo. Per questo quando parleremo della traduzione della Duchessa d’Andria terremo il 1937 come anno di riferimento. Inoltre segnaliamo che nelle note del traduttore contenute nella versione di N. Marcialis del 1991 si attestano altre due traduzioni non presenti in questo o in altri elenchi: di Nerina Martini Bernardi Carrescia (1975) e di Anjuta Gančikov (1981).

10. *Il cappotto*, traduzione di O. Del Buono. Milano, Rizzoli, 1949.
11. *Il cappotto*, in: *Tutti i racconti*, traduzione di L. Pacini Savoj. Firenze, G. Casini, 1957.
12. *Il cappotto*, in: *Il cappotto e altri racconti*, traduzione di P. Cazzola. Torino, Paravia, 1958.
13. *Il cappotto*, in: *Il Cappotto e altri racconti*, traduzione di M. Monti. Roma, Editori riuniti, 1961.
14. *Il cappotto*, traduzione di G. De Dominicis Jorio, Pescara, Edizioni paoline, 1962.
15. *Il cappotto*, in: *Taras Bulba. I racconti di Pietroburgo*, traduzione di G. Pacini. Roma, Istituto geografico de Agostini, 1963.
16. *Il cappotto*, in: *Il cappotto. Il naso*, traduzione di A. Julovic. Firenze, Sansoni, 1964.
17. *Il cappotto*, in: *I racconti di Pietroburgo*, traduzione di P. Zveteremich. Milano, Garzanti, 1967.
18. *Il cappotto*, in: *Racconti di Pietroburgo*, traduzione di G. Raspi. Milano, Fabbri, 1968.
19. *Il cappotto*, in: *Il cappotto e altri racconti*, traduzione di R. Abbate. Napoli, IEM, 1970.
20. *Il cappotto: da i Racconti di Pietroburgo*, traduzione di N. Martini Bernardi. Verona, Officina Bondoni, 1975.
21. *Il cappotto*, traduzione di E. Bazzarelli. Milano, Rizzoli, 1980.
22. *Il cappotto*, in: *I racconti di Pietroburgo*, a cura di F. Mariano. Milano, Mondadori, 1986.
23. *Il cappotto*, in: *Il cappotto e altri racconti*, traduzione di S. Beffa. Sesto san Giovanni, A. Peruzzo, 1986.
24. *La mantella*, prefazione di C. De Michelis, traduzione di N. Marcialis. Roma, Salerno Edizioni, 1991.
25. *Il cappotto*, in: *Il cappotto e Il naso*, traduzione di L. De Nardis. Roma, Newton Compton, 1993.
26. *La mantella*, in: *Opere*, vol. 1, traduzione di S. Prina. Milano, I meridiani Mondadori, 1994.
27. *Il cappotto*, in: *I racconti di Pietroburgo*, traduzione di E. Guercetti. Milano, Rizzoli, 1995.
28. *Il cappotto*, a cura di O. Gnerre. Napoli, F.lli Conte, 1995.
29. *Il cappotto*, in: *Racconti di Pietroburgo*, traduzione di F. Legittimo. Venezia, Marsilio Editori, 2001.

30. *Il cappotto*, traduzione di E. Bollardi. Roma, Fermento, 2005.

31. *Il cappotto*, in: *Racconti di Pietroburgo*, traduzione di F. Pizzi. Milano, Dalai, 2012.

Per quanto riguarda il titolo, oltre al più diffuso, *Il Cappotto*, abbiamo quattro varianti traduttive: *Il mantello* (di Loschi 1903, poi ripreso da Alvaro 1920 e Landolfi 1941), *L'uniforme* (Ciampoli 1916), *Il pastrano* (Carafa D'Andria 1937) e *La mantella* (Marcialis 1991, Prina 1994). Ciò deriva probabilmente dal duplice significato della parola russa *šinel'* che, come nota Inkova, indicava all'epoca di Gogol' due tipi di soprabito: “quello militare di taglio speciale, con una piega sulla parte posteriore e il cappotto da uomo di taglio ampio con un piccolo collo di pelliccia e una mantellina”, indossato dagli impiegati statali (Inkova 2014: 44). In entrambi i casi, non si trattava del capo di abbigliamento che oggi conosciamo. Alcuni traduttori avrebbero scelto dunque di tradurre *šinel'* con un termine che ne ricordasse maggiormente le caratteristiche: il termine “uniforme” indica una divisa indossata da un particolare rango. “Mantello” richiama le caratteristiche estetiche dell'indumento (la parte superiore ricorda effettivamente un mantello). “Pastrano” si riferisce a un tipo di uniforme prettamente maschile, quale era appunto la *šinel'* del racconto. Infine, la scelta di Marcialis e Prina di mantenere il genere femminile della parola russa *šinel'* consente di non alterare il simbolismo del racconto, per cui il cappotto nuovo di Akakij è allegoria della donna desiderata (Inkova 2014: 50).

Per quanto riguarda le numerose traduzioni, quella di Giuseppe Loschi del 1903 rimane a molti sconosciuta, o meglio si ritiene che risalga al 1918 (data della ristampa), attribuendo più comunemente la prima versione italiana dell'opera di Gogol' a Domenico Ciampoli (1916). Le due traduzioni, tuttavia, potrebbero essersi ispirate alla versione francese già esistente, come precisano Inkova (2016: 96-97) e Pala (2009: 35). Pala testimonia anche l'insoddisfazione generale diffusa tra i maggiori slavisti dell'epoca – anni Venti – per la comparsa in Italia di traduzioni superficiali e approssimative, aventi come testo fonte non l'originale russo, bensì il francese, e di conseguenza “ridicole”, secondo una dichiarazione di Lo Gatto comparsa in una recensione del 1921 a *Novelle russe* tradotte da C. Alvaro²¹³. La versione di Rebora, pubblicata nel 1922, venne dunque

²¹³ Pala riporta per intero le parole di Lo Gatto: “Ed è veramente ridicolo, in un paese come il nostro, in cui si sono trovati traduttori magnifici di opere sanscrite, persiane e da tutte le lingue più difficili dell'universo globo, sentir della gente ripetere, come in cantilena, che il russo è troppo estraneo a noi perché si possa affrontare direttamente” (Lo Gatto, Recensione alle *Novelle Russe*, “Russia”, I, 4-5, 1921, p.114).

riconosciuta, a quell'epoca, come la prima traduzione valida: riportiamo le parole di E. Lo Gatto tratte dalla sopracitata recensione comparsa ne «L'Italia che scrive»:

“Il Cappotto è uno dei primi capolavori della letteratura russa, ed era una lacuna profonda per noi non averne una buona traduzione. Questa del Rebora, nonostante l'impronta personale che hanno tutte le traduzioni del Rebora, colma assai lodevolmente questa lacuna, che la traduzione mediocre di Ciampoli lasciava ancora risentire” (Lo Gatto 1922: 151).

In ordine cronologico, dunque, la prima traduzione che abbiamo deciso di inserire nel CP ru-ita è proprio quella di Rebora, probabilmente la più originale, grazie all'influsso dello stile tipico del Rebora poeta, ma non per questo infedele al testo russo (Scandura 1997: 290).

Per quanto riguarda le successive versioni, abbiamo scelto innanzitutto quelle firmate da traduttori noti nella slavistica italiana e appartenenti a periodi diversi del Novecento, conosciuti per aver curato o tradotto, oltre al *Cappotto*, molte altre opere di rilievo della letteratura russa:

- E. Carafa D'Andria (1937)
- T. Landolfi (1941)
- L. Pacini-Savoj (1957)
- G. Pacini (1963)
- P. Zveteremich (1967)
- E. Bazzarelli (1980)
- N. Marcialis (1991)
- S. Prina (1994)
- E. Guercetti (1995)

Enrichetta Carafa D'Andria, nota con lo pseudonimo di Duchessa d'Andria, assidua frequentatrice degli ambienti editoriali italiani della prima metà del XX secolo, era molto conosciuta e stimata alla sua epoca (come testimoniano gli elogi e le osservazioni critiche alle sue traduzioni di Lo Gatto²¹⁴). Collaborava alla casa editrice «Slavia» fondata da A.

²¹⁴ Cfr., ad esempio, «L'Italia che scrive», Anno IV, n. 8, agosto 1921, p 167-168.

Polledro e si può dire che fu tra i primi traduttori italiani che diedero vita a una scuola di traduzione dal russo (Mazzucchelli 2005: 90-91). Alla traduttrice si deve, ad esempio, la prima versione integrale di *Guerra e Pace*, pubblicata nel 1928 a Torino da «Slavia». La sua versione del *Cappotto* e le scelte traduttive tradiscono, più delle altre versioni qui proposte, il segno del passato, ciononostante, riteniamo che l'inserimento nel CP di traduzioni più datate sia fondamentale per consentire un'analisi in chiave diacronica.

Tommaso Landolfi, particolarmente dedito allo studio di autori russi, primo fra tutti Gogol', iniziò la sua carriera di traduttore proprio con *I Racconti di Pietroburgo*. Anch'egli insofferente nei confronti di alcune delle traduzioni approssimative e poco fedeli al testo fonte che lo avevano preceduto²¹⁵, sembra essere il primo, come afferma Pala, a voler proporre una versione realmente fedele del racconto anche “a costo di affaticare il lettore” (Pala 2009: 37).

Leone Pacini-Savoj, noto slavista e critico letterario, cura l'antologia *Teatro russo*²¹⁶, alcune opere di L. Tolstoj; traduce, oltre al *Cappotto*, altre opere di N. Gogol' e le poesie di V. Solov'ev

Gianlorenzo Pacini, traduttore, saggista e docente di letteratura russa, è curatore di alcune opere di L. Andreev e V. Korolenko; ha tradotto I. Babel', F. Dostoevskij, Ju. Nagibin, A. Platonov, L. Tolstoj.

A Pietro Zveteremich si devono numerose traduzioni di M. Cvetaeva, Ju. Dombrovskij, F. Dostoevskij, I. Ehremburg, V. Erofeev, E. Evtušenko, V. Grossman, V. Majakovskij, B. Pasternak, K. Paustovskij, B. Pilniak, F. Sologub, A. Solženicyn, L. Tolstoj, M. Zoščenko.

Di Bazzarelli si ricordano le innumerevoli traduzioni dei maggiori prosatori russi (L. Tolstoj, F. Dostoevskij, A. Čechov, M. Bulgakov) e di poeti quali A. Achmatova, A. Blok, A. Puškin e F. Tjutčev.

Nicoletta Marcialis, docente di filologia slava, ha curato e tradotto le opere di numerosi autori russi, tra cui M. Bachtin, D. Fonvizin, L. Ginzburg, V. Ivanov, Ju. Lotman, B. Uspenskij, P. Zachar. Come emerge da (Scandura 1997: 292), la versione di Marcialis si differenzia molto dalle precedenti per via dell'interpretazione marcatamente

²¹⁵ Cfr. recensione di Lo Gatto, nota 212.

²¹⁶ *Teatro russo. Dalle origini a Čechov*. Milano: Nuova Accademia, 1960.

filologica²¹⁷. Inoltre, come detto, è la prima a proporre per il racconto gogoliano *Šinel'* il titolo *La Mantella*²¹⁸.

Serena Prina ha tradotto M. Bulgakov, F. Dostoevskij, N. Dubovickij, *Opere* di N. Gogol' oltre al *Cappotto* – di cui come Marcialis propone come titolo *La Mantella* – alcune lettere di V. Majakovskij, Ju. Nagibin, B. Pasternak, G. Vladimov.

Sono, infine, di Emanuela Guercetti una versione di *Delitto e Castigo* di F. Dostoevskij e una di *Resurrezione* di L. Tolstoj. Traduce, inoltre, A. Blok, M. Bulgakov, A. Čechov, I. Fazil. D. Savickij, e L. Ulickaja.

Questa rassegna, che non ha alcuna pretesa di essere esaustiva, intende testimoniare la presenza importante dei traduttori prescelti nel panorama della letteratura e della cultura russa in Italia nell'arco del Novecento. Ciononostante, sebbene abbiamo ritenuto fondamentale inserire le loro versioni del *Cappotto* nel CP ru-ita, crediamo anche che un corpus rappresentativo di una data lingua debba contenere il maggior numero di porzioni della stessa indipendentemente dalla buona riuscita o meno del testo prodotto, che sia esso un originale o una traduzione. Se è vero che la traduzione poggia su di un testo fonte da cui viene indubbiamente influenzata, è anche vero che costituisce un prodotto linguistico che può essere studiato come fenomeno a sé stante. Questa prospettiva, che predilige un atteggiamento puramente descrittivo e non valutativo, è adottata dai *Descriptive Translation Studies*, che analizzano il fenomeno della traduzione partendo dal testo d'arrivo, anziché dal testo fonte²¹⁹. Sulla base di ciò, nessuna traduzione dovrebbe essere esclusa a priori da un CP.

Pertanto, ad eccezione delle prime traduzioni di Loschi, Ciampoli e Alvaro, per cui non si è certi se abbiano tradotto il racconto di Gogol' direttamente dal russo o dal francese²²⁰, il nostro fine è riuscire ad inserire tutte le versioni italiane del *Cappotto* reperibili. Presentiamo quindi l'elenco delle traduzioni fino ad ora raccolte e già disponibili sul CP ru-ita, che vanno ad aggiungersi alle dieci sopracitate:

²¹⁷ Scandura (1997: 292) cita ad esempio la volontà della studiosa di trovare equivalenti italiani per i nomi di battesimo del protagonista.

²¹⁸ Purtroppo, a causa di un problema verificatosi durante l'allineamento, questa versione non è potuta rientrare nel CP in occasione del suo primo ampliamento (dicembre 2015), ma sarà aggiunta successivamente.

²¹⁹ Cfr. Toury G. (2012). *Descriptive Translation Studies – and beyond*.

²²⁰ La questione relativa al testo di partenza rimane poco chiara. Nonostante le affermazioni di Lo Gatto (cfr. nota 212), Sandura, ad esempio, sostiene che la traduzione dal francese, anziché dal russo, non sarebbe avvenuta per le versioni di Alvaro e di Ciampoli (Scandura 1997: 288).

- O. Del Buono (1949)
- A Julovic (1964)
- F. Mariano (1986)
- S. Beffa (1986)
- L. De Nardis (1993)
- F. Legittimo (2001)

Alcuni di questi traduttori, pur meno presenti nel mondo della slavistica rispetto ai precedenti, hanno avuto modo di curare e tradurre altre opere (non solo dal russo) oltre al *Cappotto*. Oreste Del Buono, noto giornalista, scrittore e critico letterario, ha tradotto molte opere, soprattutto dal francese. Luisa De Nardis è curatrice di varie raccolte di poesie di M. Cvetaeva e traduttrice di autori cechi e russi tra cui Gogol' e A. Afanas'ev. Silvana Beffa è co-traduttrice di una versione di *Guerra e Pace* di L. Tolstoj. Francesca Legittimo insegna lingua russa e sembrerebbe aver esordito come traduttrice letteraria dal russo proprio con *il Cappotto*. Nulla si sa invece di Francesco Mariano e A. Julovič, di cui non è reso noto nemmeno il nome. Inoltre, Mariano è indicato più precisamente come curatore della versione del 1986, ma non è chiaro – perché non specificato sull'edizione - se sia stato anche il traduttore. Ciononostante, riteniamo che anche queste versioni possano essere utili per un'analisi di tipo prettamente linguistico.

5.2. I dialoghi di Akakij Akakevič: analisi linguistica

Dal momento che la nostra analisi verte su due caratteristiche del parlato russo presente nei dialoghi di Akakij Akakievič, ci sembra necessario chiarire in che senso intendiamo il rapporto fra la lingua parlata e la sua imitazione nei testi letterari.

Come spiega Ejchembaum (1919), Gogol' fa parlare poco i personaggi del *Cappotto*, ma in quei rari casi il discorso ha una forza e un valore particolare e, soprattutto, è stilizzato e non ricrea mai del tutto dialoghi di carattere quotidiano. Ogni personaggio ha dunque il suo modo di esprimersi: secondo il critico, il linguaggio frammentario ed ellittico di Akakij Akakievič rientra, per esempio, “*v obščuju sistemu gogolevskoj zvukoreči i mimičeskoj artikuljacii*”²²¹, mentre il linguaggio di Petrovič, sicuro di sé e severo, è creato appositamente per contrapporsi al primo.

²²¹ “nel sistema generale della fonetica del linguaggio e dell'articolazione mimica di Gogol'”.

Tuttavia, dal punto di vista delle strutture linguistiche, le frasi pronunciate da Akakij - con i numerosi segnali discorsivi, che di norma servono a costruire il discorso, stabilire collegamenti logici ed esprimere giudizi (cfr. par. 5.2.1.), e con le ellissi - riproducono il russo parlato quotidiano (*rususkaja razgovornaja reč'*)²²², che Zemskaja, Kitajgorodskaja e Širjaev (1981: 49) definiscono come una varietà della lingua standard, riconoscibile grazie a caratteristiche specifiche e a unità linguistiche particolari presenti in tutti i livelli della lingua (sintassi, fonetica, morfologia ecc.). Nella descrizione di (Zemskaja et al. 1981: 5-7) si evidenzia innanzitutto lo stretto rapporto tra la lingua parlata e la situazione extralinguistica, definita dagli autori “*konsituacija*” (1981: 25), che è parte integrante dell’atto comunicativo e racchiude in sé una spiccata componente non-verbale; il senso di un enunciato non è infatti mai interamente espresso a parole, ma viene in parte inferito dal contesto, dando vita a svariati fenomeni compresi quelli che vedremo nei prossimi paragrafi. Altre caratteristiche fondamentali della lingua parlata sono: 1) l’assenza di previa preparazione, dunque l’improvvisazione²²³; 2) La naturalezza dell’atto comunicativo; 3) la partecipazione diretta dei parlanti all’atto comunicativo; 4) la tendenza a costruire il discorso liberamente, dando vita anche ad unità linguistiche innovative o non standard; 5) la tendenza ad usare formule fisse (benché contrastanti, queste ultime due caratteristiche convivono nel sistema della lingua parlata); 6) l’oralità come mezzo di espressione. Infine, per quanto riguarda la sintassi del parlato, essa è caratterizzata da numerosi fenomeni di ellissi (Zemskaja et al. 1981: 201-206), tali per cui spesso gli attanti sono espressi non all’interno dell’enunciato, ma nella *konsituacija*²²⁴.

²²² In (Zemskaja et al. 1981: 5) si precisa che la parola *reč'* (*discorso*) non si riscontra nelle altre lingue, che al contrario preferiscono la dicitura “lingua parlata”. L’uso del termine “discorso” sarebbe improprio, poiché evidenzia una presunta opposizione tra la lingua standard codificata (*kodificirovannyj literaturnyj jazyk*) e il parlato. Secondo gli autori (Zemskaja et al.) infatti, il parlato è a tutti gli effetti un sistema linguistico.

²²³ L’improvvisazione può variare da un massimo, come avviene in un dialogo inatteso fra due persone incontratesi casualmente, a un minimo, per esempio in un discorso preparato e studiato prima di essere proferito in pubblico (Zemskaja: 2004: 249).

²²⁴ Anche nei testi letterari l’esistenza di un non-detto al di là dell’atto comunicativo vero e proprio può portare a diverse forme di ellissi sintattica che tratteremo più nel dettaglio nel par. 5.2.3. Quando il non detto viene realizzato in un testo letterario che imita il parlato può essere ricostruito in due modi: 1) mediante il riferimento al background di informazioni che autore e lettori condividono a priori, in quanto parlanti di una stessa lingua o appartenenti a una medesima cultura; 2) mediante le esplicitazioni della voce narrante al di fuori dei dialoghi, sotto forma di narrazione di fatti o attraverso la descrizione di luoghi, situazioni, personaggi.

I dialoghi di Akakij Akakevič che prenderemo in considerazione in questo capitolo appartengano dunque a una varietà parlata e in parte ne riproducono le caratteristiche, ma non corrispondono a campioni di lingua parlata autentica, bensì a brani di un testo scritto letterario che intende semplicemente riprodurla²²⁵. Infatti, i dialoghi di un romanzo, per quanto vogliano riprodurre lo stile del parlato, e per quanto naturali possano risultare, presuppongono un lavoro di riflessione e preparazione non indifferente²²⁶. Nel nostro caso, come detto, si aggiunge la volontà di Gogol' di rendere il linguaggio di Akakij Akakievič una caricatura.

La nostra ricerca, dunque, non aspira ad essere uno studio della lingua parlata, ma ha come punto di partenza un linguaggio molto specifico contenuto nei dialoghi di una singola opera, *Il Cappotto*. Il nostro scopo è più modesto: studiare come vengono resi in lingua italiana gli elementi tipici del parlato presenti nei dialoghi di Akakij, dal momento che il CP ru-ita dispone di ben quindici traduzioni italiane di quest'opera e offre la possibilità di consultarle simultaneamente. Resta chiaro che per uno studio approfondito del parlato i nostri esempi non possono fornire una base empirica sufficiente, ma senz'altro costituiscono un valido punto di partenza. D'altra parte, anche L. Serianni, nella sua *Grammatica Italiana*, commentando i segnali discorsivi puntualizza il loro largo impiego nel discorso orale o nella prosa letteraria che tende a riprodurlo più da vicino (Serianni 1991: 363).

5.2.1. Aspetti linguistici, metodologia e limiti

Per la ricerca presentata in questo capitolo abbiamo adottato un approccio diverso rispetto ai capitoli 3 e 4, dove l'analisi degli esempi era preceduta da un'introduzione sullo status quaestionis degli studi relativi agli elementi che ci accingevamo ad analizzare (prefissi verbali russi e causativi). In quei casi il corpus era stato interrogato in modo tradizionale, ricercando i prefissi verbali e i causativi all'interno dei due macro sotto-corpora disponibili (russo-italiano e italiano-russo). Ora, al contrario, il punto di partenza è una

²²⁵ Per quanto riguarda la differenza tra lingua parlata e testi letterari, infatti, un'importante distinzione coinvolge proprio il rapporto tra discorso e realtà: se la lingua parlata si lega strettamente agli elementi della realtà extralinguistica e alla *konsituacija*, i testi letterari, al contrario, riflettono una realtà immaginata, fittizia, meno legata al mondo concretamente esistente in cui viviamo (Onipenko, Sidorova, Zolotova 2004: 12).

²²⁶ Širiaev (2001: 132) assume una posizione rigorosa al riguardo, ribadendo la differenza tra lingua parlata e lingua standard codificata e affermando che i dialoghi del parlato hanno caratteristiche molto specifiche e quasi totalmente assenti in altri testi, compresi i dialoghi della narrativa, per quanto simili.

singola opera. Il sotto-corpus preso in considerazione contiene dunque esclusivamente il racconto *il Cappotto* e le sue quindici versioni allineate. L'estrazione delle parti dialogate è avvenuta manualmente all'interno dell'originale russo e solo in un secondo momento è stato interrogato il motore di ricerca per ottenere le traduzioni allineate di ogni singolo estratto. I fenomeni da indagare sono stati scelti solo dopo una prima fase euristica, in cui abbiamo osservato i dialoghi dal punto di vista linguistico. Qui vogliamo dunque riproporre, nell'ordine, le fasi del nostro lavoro, presentando in primo luogo gli estratti da esaminare. In essi evidenzieremo gli elementi linguistici più ricorrenti e solo al termine li collocheremo all'interno di un quadro teorico di riferimento.

Come già detto, i dialoghi non sono la parte preponderante del racconto e le battute pronunciate da Akakij Akakievič si limitano a una quindicina. Ne riportiamo qui la maggior parte, a cui se ne aggiungerebbero poche altre, troppo brevi e poco significative per essere prese in considerazione:

(1) — А я **вот** к тебе, Петрович, **того**...

(2) — А я вот того, Петрович... шинель-то, сукно... вот **видишь**, (...) да **вот** только в одном месте немного того... на спине, да еще **вот** на плече одном немного попортерлось, да **вот** на этом плече немножко — **видишь, вот** и все.

(3) — Пусть ползет, а ты тотчас заплаточку.

(4) — Ну, да уж прикрепи. Как **же** этак, **право, того!**..

(5) — Ну, а если бы пришлось новую, как бы она того...

(6) «**Этакое-то** дело этакое, (...) я, **право**, и не думал, чтобы оно вышло **того**...

(7) Так **вот** как! наконец **вот** что вышло, а я, **право**, совсем и предполагать не мог, чтобы оно было этак».

(8) «Так **этак-то!** **вот** какое уж, точно, никак неожиданное, **того**... этого бы никак... этакое-то обстоятельство!»

(9) «Ну нет, — сказал Акакий Акакиевич, — теперь с Петровичем нельзя толковать: он теперь **того...** (...) а в это время я ему гривенничек и **того**, в руку, он и будет сговорчивее и шинель тогда и того...»

(10) «Ну, уж эти французы! что и говорить, уж **ежели** захотят что-нибудь **того**, так уж точно **того...**»

(11)— я ваше превосходительство осмелился утрудить потому, что секретари **того...** ненадежный народ...

(12) «А! так **вот** ты наконец! наконец я тебя **того**, поймал за воротник! твоей-то шинели мне и нужно!

In questi esempi abbiamo evidenziato, da un lato, le particelle o i segnali discorsivi (per una precisazione sulla differenza fra questi due termini cfr. par. 5.2.2) e, dall'altro, i fenomeni di ellissi sintattica. Riportiamo nella tabella 17, in ordine di frequenza, gli elementi estrapolati appartenenti a queste due categorie.

Tab. 17. Elenco dei segnali discorsivi e degli esempi di ellissi riscontrati nei dialoghi del *Cappotto*

ELEMENTO	CATEGORIA	FREQUENZA
ТОГО	segnale discorsivo	11
ВОТ	segnale discorsivo	9
ТО	segnale discorsivo	5
ПРАВО	segnale discorsivo	4
НУ	segnale discorsivo	4
ВИДИШЬ	segnale discorsivo	2
ЖЕ	segnale discorsivo	1
ЕЖЕЛИ	segnale discorsivo	1

А я вот к тебе	ellissi verbale	1
А я вот того, Петрович... шинель-то	ellissi verbale	1
в одном месте немного того ... на спине,	ellissi verbale	1
а ты тотчас заплаточку	ellissi verbale	1
а если бы пришлось новую,	ellissi verbale	1
как бы она того...	ellissi verbale	1
этого бы никак	ellissi verbale	1
я ему гривенничек	ellissi verbale	1
в руку	ellissi verbale	1
и шинель тогда и того	ellissi verbale	1

Questa analisi ci ha permesso di considerare due fenomeni linguistici altrimenti difficili da estrarre da un corpus. Molti dei segnali discorsivi elencati in tabella 17, infatti, presentano omonimi appartenenti a diverse parti del discorso. Prendiamo, ad esempio, *togo*, largamente usato come pronome dimostrativo al genitivo e ricorrente per giunta in svariate costruzioni, quali *krome togo*; *posle togo*, *kak*, *dlja togo*, *čtoby* e così via; o ancora *vidiš'*, che si presuppone ricorra molto frequentemente anche come predicato, oltre che con funzione discorsiva. La ricerca lessico-grammaticale del NKRJa consente di selezionare la parte del discorso da indagare. Nel caso di *togo*, è possibile limitare la ricerca agli estratti in cui esso compare come particella e non come pronome. Tuttavia, ciò funziona solo per gli estratti previamente sottoposti a disambiguazione, in numero molto esiguo rispetto al volume totale del corpus e al momento presenti solo nel corpus principale.

Un'altra difficoltà nell'interrogare il corpus si presenta per gli esempi di ellissi sintattica: per via della loro eterogeneità sia formale sia funzionale, la loro identificazione richiede infatti diverse operazioni. Innanzitutto è necessario individuare le strutture sintattiche più ricorrenti e limitare l'analisi a questi casi (es.: pronome personale al nominativo + preposizione + sostantivo all'accusativo), in secondo luogo è necessario un

intervendo manuale per estrapolare dai risultati solo i casi realmente corrispondenti a fenomeni di ellissi.

Chiarito il procedimento utilizzato per identificare nel nostro corpus gli elementi da analizzare, nei due paragrafi che seguiranno descriviamo i principali risultati della riflessione linguistica sui segnali discorsivi e l'ellissi sintattica.

5.2.2. I segnali discorsivi

La nostra disamina degli studi su questa categoria linguistica prenderà in considerazione le tradizioni italiana e russa, pertinenti alle due lingue del nostro CP. Siamo ben consapevoli dei numerosi studi sui segnali discorsivi prodotti, soprattutto negli ultimi anni, in ambito anglosassone e germanico (ad esempio, l'utile *Approaches to discourse particles*, a cura di K. Fischer, 2006), tuttavia la nostra scelta è dettata dalla funzione di questo capitolo, che vuole essere più un esempio di come utilizzare il corpus parallelo per la ricerca linguistica che un'indagine esaustiva.

Serianni (1991: 361) definisce i “segnali discorsivi” (ingl. “*discourse markers*”) come elementi con “la funzione di organizzare la presentazione del testo comunicativo secondo criteri dimensionali e logico-narrativi”. Essi contribuirebbero dunque a garantire in un testo due proprietà fondamentali: coesione (“il rispetto delle relazioni formali tra le varie parti del testo”) e coerenza (l’insieme dei procedimenti che consentono di produrre senso), senza però influire sul piano contenutistico della frase. Come precisano Aijmer, Foolen e Simon-Vandenberg (2006: 101), essendo *pragmatic markers* svolgono una pura funzione discorsiva.

Non esiste una classificazione universale dei segnali discorsivi. Come afferma Fischer, gli studi ad essi dedicati sono eterogenei e poco comparabili, per via delle diversità per quanto concerne terminologia, lingue considerate e metodo di analisi (Fischer 2006: 1). In italiano, oltre al termine “particelle”, che, a seconda dell’approccio, può riferirsi a tutti i segnali discorsivi o solo ad un gruppo specifico, Serianni (1991: 362) individua, già alla fine degli anni Ottanta²²⁷, numerose altre denominazioni, tra cui “connettivi pragmatici”, “riempitivi”, “elementi di articolazione”.

Oltre alla discussione sulla terminologia si registrano discordanze per quanto riguarda l’oggetto di studio: la tradizione russa e poi sovietica, risalente agli inizi del XX

²²⁷ La prima edizione della Grammatica risale al 1989.

secolo per lungo tempo ha preso in considerazione solo le particelle in senso stretto, ovvero quelle unità non flesse e contrapposte alle parti del discorso tradizionali; diversamente, gli studi occidentali, più propensi ad indagare il fenomeno da un punto di vista pragmatico, fin dall'inizio inclusero fra i segnali discorsivi unità linguistiche varie, tra cui avverbi, verbi o intere frasi (cfr. l'italiano “*va bene?*”)²²⁸.

Infine, molti segnali discorsivi possono essere considerati come usi pragmatici di parti del discorso tradizionali, di cui potrebbero essere considerati omonimi (cfr. Fernandez 1994). Per esempio, *diciamo*, in quanto segnale discorsivo (13) perderebbe la sua funzione verbale per diventare una particella, omonima della prima persona plurale del presente indicativo di *dire*.

(13) *Perciò se Alik andava, **diciamo**, a Washington per una settimana, sapeva che Nina non sarebbe andata a fare la spesa (...)* (Ulickaja)²²⁹

In (1995: 225) Bazzanella conferma che i segnali discorsivi sono quegli elementi che, “svuotandosi in parte del loro significato originario, assumono dei valori aggiuntivi”. Tuttavia, la studiosa non tratta espressioni simili a “*diciamo*” come omonimi della parola originaria. Si limita, piuttosto, ad affermare che, appartenendo a svariate parti del discorso e abbandonandone le funzioni originarie, i segnali discorsivi non possono essere analizzati né dal punto di vista morfologico, né da quello lessicale, ma solo funzionale.

In (2008: 18-19), Bonola individua alcuni criteri formali e semantico-funzionali per distinguere all'interno della lingua russa le particelle (*časticy*) dal più vasto insieme dei segnali discorsivi. Per definirsi tali, queste ultime devono essere sincategorematiche e indeclinabili - dal punto di vista formale - ed essere dedicate alle funzioni modale e connettiva. Esistono evidentemente anche avverbi, locuzioni avverbiali o, come abbiamo visto, verbi con funzione modale o connettiva, ma in essi la funzione modale non può essere considerata preferenziale. Unità linguistiche di questo tipo possono dunque essere impiegate in veste di segnali discorsivi, ma non nascono come tali, a differenza delle particelle, che invece presentano i tratti sopra descritti come primari.

Per quanto riguarda la russistica, abbiamo già accennato al fatto che fino all'inizio del XX secolo prevale la distinzione formale fra parti del discorso (*časti reči*) e particelle

²²⁸ Per la precisione anche in russo, come nota Gobber in (2004: 340), vi sono parole quali *konečno* che esistono come enunciati autonomi.

²²⁹ *Veselye pochorony* (1997).

(*časticy*); Peškovskij, nel suo *Russkij sintaksis* (la prima edizione è del 1914), parla delle particelle come di parole “senza forma”, ossia indeclinabili. Inoltre, per distinguerle da alcuni gruppi di avverbi (cfr. *včera – ieri*), sottolinea la loro sinsemanticità come criterio fondamentale per distinguerle dalle parole categorematiche (con significato autonomo). Sulla base di questo approccio le particelle, seppur suddivise in diverse classi (Peškovskij ne conta otto²³⁰) sono accomunate a altre “parole parziali”, come congiunzioni e preposizioni, trattandosi in tutti i casi di unità “ausiliarie” (*služebnye*) o “di raccordo” (*svyazočnye*), indeclinabili e sincategorematiche²³¹ (Peškovskij 2001: 39-40).

Vinogradov (2001²³²: 547-554), seguendo la tradizione grammaticale russa, suddivide le particelle russe in otto gruppi, ma ne amplia lo spettro, comprendendo anche avverbi la cui sinsemanticità è meno evidente rispetto a particelle come *že* o *ved’*: *edinstvenno* (naturalmente), *izključitel’no* (esclusivamente), *rešitel’no* (in maniera decisiva), *položitel’no* (positivamente), *prосто* (semplicemente) e così via. Inoltre, sottolinea con forza che le particelle russe si distinguono da altre parti del discorso o da ulteriori parole “parziali” non per l’aspetto formale, ma per la loro funzione, che in *O kategorii modal’nosti i modal’nych slovač v russkom jazyke* (redatto per la prima volta nel 1950²³³) definirà “modale”. Nella lingua russa, infatti, intonazione e “parole ausiliarie speciali (come le particelle: *ne, li, ved’, razve* ecc.)” sono preposte ad esprimere il valore modale dell’enunciato (1975: 56). Anche in russistica, dunque, dagli anni Cinquanta in poi comincia a prevalere l’approccio semantico e, soprattutto, funzionale. Per la prospettiva pragmatica bisognerà invece attendere gli anni Ottanta (Bonola 2008: 15-17), che vedono la pubblicazione dello studio di T. M. Nikolaeva sulle particelle (Nikolaeva 1985). Sarà poi con D. Paillard, alcuni anni più tardi, che la tradizione russa e quella occidentale convergeranno e le particelle russe verranno raggruppate sotto la più ampia categoria dei “segnali discorsivi” (*diskursivnye slova*). Kiseleva e Paillard (2003) inseriscono lo studio dei segnali discorsivi russi nell’ambito della semantica discorsiva e li definiscono “parole relazionali” (*slova-reljatory*) in grado di conferire un determinato

²³⁰ 1. Preposizioni (es.: *v, iz*). 2. Congiunzioni (es.: *i, čtoby*). 3. Verbi copulativi (es.: *est’*). 4. Particelle rafforzativi o enfaticizzanti (es.: *daže, to*). 5. Interiezioni (es.: *konečno, net*). 6. Particelle esortative (es.: *pust’, da*). 7. Negazioni (es.: *ne, ni*). 8. Particelle interrogative (es.: *razve, neuželi*) (Peškovskij 2001: 41-42).

²³¹ Notiamo che Vinogradov, nel suo *Russkij jazyk*, cita una teoria già presente all’epoca e presentata nella grammatica di Sidorov e Avanesov (1945), che distinguevano tra particelle in senso stretto, da un lato, e preposizioni e congiunzioni, dall’altro. Oggi, come afferma Jurij Apresjan nell’introduzione all’ottava edizione di *Russkij sintaksis* di Peškovskij, questa distinzione è definitivamente accettata in linguistica.

²³² Prima edizione del 1947.

²³³ La versione a cui facciamo qui riferimento risale al 1975.

“status discorsivo” al proprio *scope*, ovvero ai frammenti di discorso da essi collegati (Kiseleva, Pajar 2003: 9). È questo l’elemento che li differenzia da altre unità linguistiche come i connettori, che, come suggerisce il termine, svolgono allo stesso modo una funzione di legame o connessione. Un segnale discorsivo è dunque classificato in base allo “status discorsivo” che realizza. Si evidenziano, in particolare, due tipi di status discorsivo: a) il punto di vista sul mondo (il segnale discorsivo indica un punto di vista distinto da altri); b) il rapporto con il mondo, per cui il segnale discorsivo indica il grado di adeguatezza nel rapporto Z-p, dove p è l’enunciato e Z il mondo (2003: 13). Citiamo uno degli esempi riportati dagli autori (2003: 15): in (14) *pravda* esprime un punto di vista, mentre in (15) indica che l’enunciato p è veritiero, quindi adeguato a Z:

(14) *Позвольте поднести Вам колокольчики. Правда, они звенят немного печально, но это ничего. Утром они завянут и успокоятся.* (Е. Шварц)

(15) - *А вы закройте глаза, - посоветовал Дробаха, - иногда это помогает. Женщина улыбнулась и правда закрыла глаза.* (Р. Самбук)

Secondo la *Relevance Theory* (cfr. Gobber 2002: 46) alcuni segnali discorsivi possono essere interpretati come predicati sovraordinati impliciti che indicano l’atteggiamento del parlante nei confronti dell’enunciato²³⁴.

Come aggiunge Khachaturyan (2011), che sulla base della classificazione di Paillard per il francese e il russo elabora una serie di criteri formali che permettono di considerare la classe dei segnali discorsivi alla stregua di una parte del discorso a sé stante, il segnale discorsivo, in sostanza, riflette “un modo particolare di dire Z” (Khachaturyan 2011: 97).

Infine, dal momento che la nostra analisi si basa su un corpus parallelo composto in parte da traduzioni, riportiamo l’opinione di Aijmer et al. (2006: 101), secondo i quali il cosiddetto “*translation method*” sarebbe uno strumento utilissimo per lo studio comparato delle particelle, poiché offre i dati empirici su come le diverse lingue svolgono le funzioni modali e pragmatiche tipiche delle particelle discorsive²³⁵.

Vista l’impossibilità di elaborare una lista completa di tutti i segnali discorsivi esistenti e delle loro funzioni, soprattutto per via dell’eterogeneità degli approcci proposti

²³⁴ Gobber propone l’esempio del francese *heureusement* nella frase: *Il s’en est bien tiré, heureusement.*

²³⁵ Il termine usato da Aijmer et al. è *discourse particles*.

negli ultimi decenni, riproponiamo l'utile sintesi di Carla Bazzanella (2006: 449) riguardo a quanto accettato dai linguisti su questo tema: i segnali discorsivi 1) non incidono sul contenuto proposizionale di un enunciato; 2) si legano alla situazione in cui si inserisce l'atto comunicativo; 3) hanno valore modale ed esprimono attitudini e emozioni; 4) sono multifunzionali e operano contemporaneamente su diversi livelli.

A ciò aggiungeremo che:

1. In termini generali, i segnali discorsivi sono uno strumento utile al parlante per collegare testo e realtà e per facilitare l'esposizione di un pensiero, soprattutto nella lingua parlata.
2. Ogni segnale discorsivo ha delle funzioni particolari legate alla sua semantica.
3. Molti segnali discorsivi si differenziano da lingua a lingua e sono pertanto poco comparabili; tuttavia, se il segnale discorsivo riflette "un modo particolare di parlare del mondo" (Khachaturyan 2011: 97), il loro studio attraverso le traduzioni è utile per identificare in che modo due lingue "parlano" della stessa porzione di mondo (Aijmer et al. 2006).
4. Un gruppo particolare di segnali discorsivi è costituito dalle particelle (cfr. anche Aijmer et al. 2006: 103), molto frequenti nella lingua russa e più rare in altre lingue, come per esempio l'italiano; pertanto, un'indagine in chiave comparativa può risultare interessante sia per identificare le strategie di resa, sia per comprenderne meglio le funzioni.

5.2.3. L'ellissi sintattica

In linguistica l'ellissi è l'omissione di uno o più elementi del discorso di norma ritenuti indispensabili. Essendo un fenomeno significativamente frequente nella lingua russa, molti sono i russisti che lo hanno reso oggetto dei loro studi, sebbene diverse questioni rimangano irrisolte²³⁶.

Quando si parla di ellissi si è soliti prendere in considerazione la singola proposizione semplice, considerata unità fondamentale della sintassi e unità minima del discorso dotata di predicatività (Vjatkina 2009: 67-68). Come nota Dibrova (2008), nello

²³⁶ Già in Skovorodnikov 1973 si parlava di "crisi dell'ellissi", ossia della difficoltà a trovare un unico punto di vista per definire i criteri che indicano incompletezza e, dunque, la presenza di un'omissione (Skovorodnikov 1973: 114).

studio delle proposizioni grande attenzione viene rivolta alla struttura sintattica; in particolare emerge il concetto di schema strutturale della frase²³⁷ sviluppato da N. Ju. Švedova nella “*Akademičeskaja Russkaja Grammatika*” – 80. Švedova definisce lo schema strutturale come la base astratta sulla quale può essere costruita una proposizione minima, relativamente finita e con significato predicativo. Lo schema strutturale è costituito da due elementi principali, soggetto e predicato. I restanti elementi della frase, definiti secondari, dipendono dai primi e non concorrono a comporre lo schema strutturale (Švedova 1980: § 1904)²³⁸. Ogni proposizione può presentare una realizzazione completa o incompleta di questa sua forma grammaticale minima, e può quindi dar luogo a proposizioni complete o incomplete.

Non vi è però unanimità nel definire gli elementi strutturalmente indispensabili, che dovrebbero essere tali in quanto portatori di predicatività. Il concetto stesso di “predicato”, infatti, non gode di una definizione universalmente valida in linguistica. Il verbo è considerato espressione del predicato per eccellenza, ma anche sostantivi, aggettivi o interi sintagmi verbali possono svolgere tale funzione in un enunciato.

Nella *Grande grammatica italiana di consultazione*, Salvi parla di elementi nucleari e di elementi circostanziali, facendo rientrare nel novero dei primi anche gli argomenti obbligatori del verbo: il verbo “incontrare”, ad esempio, presuppone due argomenti, soggetto e complemento oggetto indicante la persona incontrata; pertanto, una frase del tipo “*Maria ha incontrato*” è da considerarsi scorretta (Salvi 1988: 30). In questo caso, tuttavia, non ha senso parlare di incompletezza sintattica proprio perché la frase non ha ragion d’essere, è priva di senso. Le frasi incomplete russe di cui si è occupata la russistica, al contrario, sono sempre grammaticalmente corrette e sensate, e di queste ci occuperemo anche noi.

Passando in rassegna gli studi più significativi (Dibrova 2008, Lekant 2004, Mel’cuk 1974, Peškovskij 2001, Skovorodnikov 1973, Valgina 2000, Vjatkina 2009, Zemskaja et al. 1981) possiamo evidenziare tre questioni principali relative alla definizione di ellissi:

1) l’identificazione di criteri chiari e univoci per definire l’incompletezza di una frase;

²³⁷ Il concetto fu introdotto in V. Mateziuz (1967) *O sistemnom grammatičeskom analize*.

²³⁸ La grammatica di Švedova è stata consultata dal sito <http://www.rusgram.narod.ru/> (novembre 2016).

2) la terminologia: si parla ora di ellissi, ora di *nepolnye predloženiya* (frasi incomplete, d'ora in poi NP) e, in altri casi ancora, di zero sintattico;

3) la distinzione tra incompletezza sintattica e semantica.

In relazione al primo problema, ossia l'identificazione di criteri validi per definire una frase incompleta, è indispensabile prima di tutto menzionare la classificazione tradizionale delle proposizioni nella lingua russa, che distingue tra proposizioni monosintagmatiche (*odnosostavnye*) e bisintagmatiche (*dvuchsostavnye*). Sono definite bisintagmatiche le proposizioni in cui il nucleo predicativo è espresso da due posizioni, soggetto e predicato (Vjatkina 2009: 83). Al contrario, sono monosintagmatiche le proposizioni nelle quali è presente un solo elemento principale con funzione predicativa (generalmente soggetto o verbo) (Vjatkina 2009: 105).

Esistono vari tipi di proposizioni monosintagmatiche. Generalmente si distinguono sei gruppi. Nei gruppi 1, 2, 3 e 5 l'elemento principale è il verbo; il gruppo 4 è costituito da frasi impersonali in cui vari elementi possono svolgere la funzione di predicato (non necessariamente un verbo, come nell'esempio presentato); infine, nell'ultimo gruppo l'elemento principale è il soggetto:

1) determinatamente personali (*opredelenno-ličnye*), del tipo *em jabloko* (*mangio una mela*), dove il pronome personale è omissivo, ma il riferimento è chiaro;

2) indeterminatamente personali (*neopredelenno-ličnye*), in cui il soggetto non è espresso perché ignoto, indeterminato o irrilevante, come in *tebja sprašivajut* (*chiedono di te* – ovvero una persona precisa, di cui però si ignora l'identità);

3) personali generiche (*obobščenno-ličnye*), dove il verbo indica un'azione atemporale e generica, spesso riferita a tutti gli individui indistintamente. Es.: *u nas ne kurjat*;

4) impersonali (*bezličnye*), del tipo *mne choldno*;

5) infinitive (*infinitivnye*), dove il predicato è espresso da un verbo all'infinito indipendente. Es.: *otkuda vzyat' den'gi?* (*dove prendere i soldi?*);

6) nominali (*nominativnye*), del tipo *vot luna* (*ecco la luna*).

In particolare, ci interessa qui sottolineare che la tradizione grammaticale russa considera questo tipo di proposizioni complete (*polnye*)²³⁹. Per dimostrarlo, Valgina (2000) analizza alcune frasi monosintagmatiche, appartenenti rispettivamente al primo, secondo e quarto gruppo:

(16) *Пишу тебе письмо.*

(17) *В газетах пишут о новых научных открытиях.*

(18) *Мне не спится.*

In (16) l'omissione del soggetto non può considerarsi come indizio di incompletezza perché il significato grammaticale che il pronome personale *ja* esprimerebbe è già compreso nella forma del verbo-predicato. In (17) non c'è riferimento a una persona agente, la fonte non è importante, anzi è sconosciuta, e l'attenzione è incentrata sull'azione. Anche in questo caso, non si può parlare di ommissione del soggetto, perché quest'ultimo è inesistente. L'inserimento di un ipotetico *oni* (loro) sarebbe considerato un errore. Infine, in (18), la forma del predicato non richiede l'espressione di un soggetto al caso nominativo e il riferimento alla persona che si trova nella condizione di "non riuscire a dormire" (*ne spitsja*) è espresso al dativo. In tutti e tre i casi la predicatività della frase è racchiusa nell'unico elemento principale presente (il verbo) (Valgina 2000: 153-179).

Per stabilire la presenza o meno di incompletezza è fondamentale tener conto inoltre del concetto di "zero sintattico" introdotto da I. Mel'cuk (1974). In (1974: 350), il linguista prende come esempio le frasi (19) *Улицу засыпали песком* e (20) *Улицу засыпало песком*. In entrambi i casi il soggetto rimane inespresso, ma è certo che in (19) la strada è stata cosparsa di sabbia da esseri umani (*ljudi*), mentre in (20) da forze della natura non meglio identificate (*stichii*). È dunque lecito postulare la presenza di due "lessemi-zero" (appunto zero sintattico) nella struttura sintattica delle due proposizioni, rispettivamente $\emptyset ljudi$ e $\emptyset stichii$ e in questo caso non siamo in presenza di proposizioni incomplete (*nepolnye*): nel sistema della lingua non esiste infatti un'espressione sinonimica che possa essere enunciata attraverso una forma "non-zero" e dunque in questi casi si parla di zero sintattico. In conclusione, sia che si parli di proposizioni

²³⁹ Alcuni linguisti, come ad esempio Šmelev (2002: 158-160) in questi casi, soprattutto per i primi due gruppi, parlano di pronomi zero, in riferimento allo zero sintattico di Mel'cuk, di cui parleremo nelle prossime pagine.

monosintagmatiche con “pronomi personale-zero”, sia che si adotti il termine "zero sintattico" di Mel'čuk, si tratta in ogni caso di frasi complete, che nulla hanno a che fare con l'incompletezza sintattica.

Aggiungiamo che frasi di questo tipo – ossia impersonali o con omissione del pronome soggetto – si realizzano anche in italiano e, pertanto, non suscitano particolare interesse dal punto di vista comparativo. Diversa è invece la situazione per le proposizioni definite bisintagmatiche, ovvero originariamente costituite da soggetto e predicato, le quali possono realizzarsi in maniera completa o incompleta. Come accennato all'inizio del paragrafo, esistono varie classificazioni delle proposizioni incomplete o “ellittiche”.

Una delle più ampie è di Peškovskij (2001: 396-403), che suddivide le NP, in numerose categorie:

1. Proposizioni senza soggetto
2. Proposizioni senza predicato
3. Proposizioni senza copula al passato o futuro
4. Proposizioni senza complemento diretto
5. Proposizioni con sintagma nominale incompleto, in cui un aggettivo non è affiancato dal sostantivo a cui si riferisce.

A queste si aggiungono NP la cui incompletezza dipende da altri fattori:

6. dal legame con una frase precedente, specialmente nelle repliche nei dialoghi:

(21) *Когда ты вернешься? — Послезавтра <вернусь>;*

7. dalla situazione extralinguistica:

(22) *Идет!* (vedendo l'autobus che si avvicina alla fermata);

8. da fattori di carattere emozionale:

(23) *Нет, уж терпенья не хватает... Женю я его... суди меня бог, а я его женю...* (Ostrovskij).

Infine, Peškovskij menziona un tipo particolare di NP, definite “stazionarie” (*stacionarnye*), comprendenti espressioni ormai fissate nel linguaggio quotidiano e il cui uso è talmente diffuso e naturale che soltanto un'analisi grammaticale ne rivelerebbe l'incompletezza. Rientrano in questo gruppo espressioni come “*что с ней?*”, nonché le

formule di auguri e felicitazioni presenti nella lingua russa, quali *dobrogo zdorov'ja* o *s Novym Godom!*²⁴⁰. Peškovskij propone di distinguere queste proposizioni dalle classiche NP, considerando che il loro uso nella forma completa è quasi scomparso dal linguaggio comune, salvo qualche rara eccezione²⁴¹.

In generale Peškovskij non giudica solo la mancanza di soggetto o verbo come cruciale per determinare l'incompletezza di una frase. La questione degli elementi strutturalmente indispensabili resta dunque irrisolta, anche se, in questo lavoro, ci occuperemo solo delle proposizioni senza verbo, per via della loro importanza dal punto di vista contrastivo con l'italiano.

Resta poi da chiarire la questione terminologica, ossia se vi sia una differenza tra *nepolnye predloženiya* ed ellissi. Peškovskij, ad esempio, non distingue i due casi e non impiega mai il termine "ellissi" nella sua descrizione. Altri linguisti, più tardi, hanno invece chiarito la differenza tra i due termini attribuendo a ciascuno caratteristiche ben precise. Le frasi ellittiche vengono dunque definite come un tipo particolare di costruzioni "incomplete", la cui specificità strutturale consiste nell'omissione del predicato; quest'ultimo, a differenza di quanto accade nelle restanti NP (cfr. la classificazione di Peškovskij), non è mai menzionato nel contesto. Quando le frasi ellittiche vengono messe a confronto con ipotetiche realizzazioni complete, è solo per comprenderne la specificità strutturale. Per queste ragioni, a parere di alcuni studiosi, l'ellissi non sarebbe da considerarsi affatto come un fenomeno d'incompletezza. Mentre le NP non sono comprensibili al di fuori di un contesto, per comprendere le frasi ellittiche non è necessario reintegrare gli elementi mancanti. Dal punto di vista comunicativo esse sono assolutamente complete, anche quando prese isolatamente. Lekant (2004: 232), ad esempio, elenca tre motivi per cui l'ellissi sintattica sarebbe un fenomeno distinto dall'incompletezza:

1. se considerate singolarmente, sono complete dal punto di vista semantico e, quindi, indipendenti dal contesto;
2. gli elementi di tali proposizioni non presentano legami grammaticali con lessemi contenuti in frasi contigue (come nei casi di parallelismo sintattico).

²⁴⁰ Nella forma completa, rispettivamente, *želat' dobrogo zdorov'ja* e *pozdravljat' s Novym Godom*.

²⁴¹ La descrizione di tali proposizioni è fatta rientrare più spesso nell'ambito della fraseologia piuttosto che nella sintassi.

3. non possono essere sostituite da versioni complete sinonimiche: la frase “*Ja domoj*” non può essere completata inserendo il verbo mancante, perché non si conosce con certezza l’alternativa corretta. Eppure il suo senso è chiaro.

Skovorodnikov (1973: 118) ricorda che alcuni linguisti, riferendosi a Melčuk, ritengono che si possa parlare in questi casi di “predicato-zero”, per la loro analogia con lo zero sintattico. Tuttavia proprio Melčuk in (1974: 355-356) esclude la possibilità di immaginare un lessema-zero in questi tipi di frasi: data la larga diffusione di simili proposizioni, bisognerebbe ogni volta, a seconda del caso, specificare il tipo di “predicato-zero” mancante. Si dovrebbe così distinguere verbi-zero di movimento (*nulevye glagoly dviženija*), verbi-zero di discorso (*nulevye glagoly soobščeniija*), verbi-zero di gioco (*nulevye glagoly igry*), e così via.

In questa sede accettiamo la distinzione tra NP ed ellissi e prendiamo in considerazione solo quest’ultima, sia perché significativa, se contrapposta all’italiano, sia perché i dodici esempi da noi analizzati e presentati nel paragrafo 5.2.1., contengono quasi esclusivamente casi di questo genere.

L’ultima questione che vogliamo toccare riguarda il rapporto tra incompletezza sintattica e semantica. In (2001: 403), Peškovskij afferma: “*Неполные предложения, хотя их в разговорной речи, несомненно, гораздо больше, чем полных, в сознании нашем всегда равняются по полным*”²⁴². In altre parole, l’incompletezza è studiata esclusivamente da un punto di vista strutturale, poiché qualsiasi NP nella mente degli interlocutori è sempre completa, grazie a segni extralinguistici, al contesto e al cotesto che permettono di inferire gli elementi mancanti. Ed essendo l’incompletezza un fenomeno strutturale, essa si presenta come un’alternativa alla possibile realizzazione completa: data una proposizione strutturalmente completa e finita, possono esistere realizzazioni incomplete che ne costituiscono una variante strutturale, ma non semantica.

Al pari di Peškovskij, Valgina ribadisce che, se si guarda al valore comunicativo delle NP, esse non si distinguono in alcun modo dalle proposizioni complete, essendo perfettamente comprensibili (Valgina 2000: 194). Altri autori, al contrario, classificano i diversi tipi di proposizioni incomplete distinguendo tra incompletezza semantica e incompletezza strutturale. In Dibrova (2008: 323), per esempio, viene presentata una

²⁴² Le frasi incomplete, sebbene nella lingua parlata siano molto più frequenti di quelle complete, nella nostra mente vengono sempre percepite come queste ultime”.

classificazione che porta a considerare anche l'incompletezza semantica. La studiosa parla infatti di tre tipi di *nepolnye predloženiya*:

- strutturalmente complete, ma semanticamente incomplete; vale a dire proposizioni che pur avendo tutti gli elementi strutturali al loro posto, risultano per qualche motivo ambigue, implicite. (Es: *Potom vdrug kak-budto čto-to vspomnilos' ej* – L. Tolsoj)
- strutturalmente e semanticamente incomplete, incomprensibili al di fuori di un contesto (Es.: *mne iz klassikov*);
- strutturalmente incomplete, ma semanticamente complete (Es.: *edu domoj*);

Riteniamo che il dibattito passi qui in secondo piano dal momento che ogni esempio del corpus è analizzato all'interno di un contesto. Nella nostra analisi prenderemo dunque in considerazione solo l'incompletezza sintattica (frasi tipo *edu domoj*), quella definita da Lekant come vera e propria ellissi e che Zemskaja et al. (1981: 201-206) evidenziano come tipica della lingua parlata. Gli autori suddividono le frasi ellittiche in vari gruppi, in base al tipo di verbo omesso e alla struttura sintattica. Le situazioni più frequenti sono:

1) Omissione di un verbo con significato di moto, che può avere diverse realizzazioni:

- soggetto + sostantivo in caso obliquo (preceduto spesso da preposizione):
(24) *Я за продуктами.*
- Soggetto + infinito con significato di scopo:
(25) *Я обедать.*
- Soggetto + avverbio:
(26) *Я направо.*

2) Omissione di un verbum dicendi, che, anche in questo caso, può presentare diverse realizzazioni:

- soggetto + sostantivo in caso obliquo:
(27) *Я вот про эту скульптуру в Волгограде.*
- Soggetto + avverbio:
(28) *Ты это серьезно?*

3) Omissione di un verbo con significato simile a “colpire” (anche se stessi):

- (29) *Он его не очень сильно.*

Segue poi un elenco aperto di casi ugualmente raggruppabili in base al campo semantico del verbo omesso: omissione di un verbo con significato simile a “chiamare”:

(30) *Была бы у меня дочь / я бы ее только Анной.*

Omissione di un verbo con significato simile a “dare”:

(31) *Я вам сегодня рубль / а остальные завтра / ладно?*

E così via.

Completano il repertorio le frasi con omissione di un verbo all’infinito, il quale può dipendere da verbi modali in forma personale o impersonale (cfr. *mne chotelos’*), da aggettivi in forma breve (ad es. *dolžen*), da un altro infinito o da un avverbio, in entrambi i casi in costrutti impersonali.

Quanto alle funzioni dell’ellissi, nessuno degli autori citati vi rivolge particolare attenzione. Si può tuttavia intuire che essa sia prevalentemente associata alle esigenze di sintesi e concisione tipiche della lingua parlata.

Per quanto riguarda la nostra indagine, come già accennato, questo tipo di frasi ellittiche “a-verbali” sono le più interessanti dal punto di vista contrastivo: dal momento che in italiano non sono ammesse omissioni del verbo di questo genere, analizzeremo in che modo i traduttori colmano il vuoto lasciato dal russo nei casi sopra citati (cfr. par. 5.3.2.).

5.3. L’analisi

In questo paragrafo confronteremo le quindici traduzioni del racconto a nostra disposizione concentrandoci sui dialoghi del protagonista e, in particolare, sui due temi introdotti nel paragrafo 5.2. Tra i segnali discorsivi, abbiamo scelto di analizzare la particella *togo* per via della sua alta frequenza nei nostri esempi e per l’interesse che suscita nel confronto con l’italiano, che sembrerebbe non disporre di un perfetto equivalente. La particella *tot*, ad esempio, pur essendo molto ricorrente, riveste in russo funzioni analoghe all’italiano “*ecco*”, che infatti è impiegato nel 100 per cento dei casi proprio per rendere *tot*. Per quanto riguarda le frasi ellittiche, invece, ne prenderemo in considerazione solo alcune a titolo esemplificativo, dal momento che gli esempi presenti nel corpus analizzato hanno tutti una struttura sintattica molto simile, seppur inserita in contesti diversi.

5.3.1. La particella *togo*

Presentiamo innanzitutto le definizioni lessicografiche della particella *togo* tratte dal dizionario di Ušakov e dal *Malyj Akademičeskij Slovar* (MAS). In entrambi i dizionari *togo* è segnalato nel suo primo significato proprio come *častica* con la funzione di riempire una pausa dettata da una difficoltà di espressione²⁴³. Come si legge nel MAS prima del secondo esempio, la pausa può essere anche attribuibile a un'esitazione dovuta alla riluttanza a comunicare qualcosa di spiacevole.

Ušakov:

ТОГО́

[тово] (или тово), частица [из род. п. от то (2)] (простореч.).

1. Заполняет паузу при затруднении говорящего выразить свою мысль, подобрать нужное слово, или даже как бы заменяет собой всё недосказанное. «Уж ты, того, постарайся, Алексей Алексеич... Сердечнопрошу...» Чехов.

MAS

Того и тово, частица. Прост.

1. Служит для заполнения паузы при заминке в речи, затруднении в выборе слов, иногда взамен какого-л. слова или словосочетания.

— *Товарищ Попов, орудуй... тово... потише.* Маяковский, Столп.

|| при сообщении о каком-л. неблагополучии, неприятности или о чем-л. не совсем благовидном.

— *Хе, — сказал он, тыкнув пальцем на один холст, где была изображена нагая женщина, — предмет, того... игривый.* Гоголь, Портрет.

Riprendendo le frasi del corpus in cui compare – tutte ad eccezione di (3) e (7) - si nota che *togo* è usato appunto dal protagonista per colmare un vuoto lasciato dall'incapacità di esprimere in modo chiaro e lineare il proprio pensiero. Il narratore stesso in un punto

²⁴³ Come riportato da entrambi i dizionari qui citati, nel suo secondo significato *togo* può essere impiegato anche come sostitutivo di un predicato soprattutto per indicare una qualche anomalia attribuibile a cose o persone. Cfr.: *Он nemnogo togo.*

del racconto chiarisce che l'uso di *togo* da parte di Akakij è direttamente proporzionale al suo stato di agitazione:

(32) *Акакий Акакиевич уже заблаговременно почувствовал надлежащую робость, несколько смутился и, как мог, сколько могла позволить ему свобода языка, изъяснил с прибавлением даже чаще, чем в другое время, **частиц «того»**, что была-де шинель совершенно новая, и теперь ограблен бесчеловечным образом (...)*²⁴⁴

In alcuni casi *togo* compare alla fine di una frase che rimane poi incompleta, lasciando molto sottinteso, come in (1), (4), (5), (6), (10); in altri, come in (2), (8), (9), (11) e (12), sembra piuttosto un intercalare, un elemento di raccordo tra due frasi, pur sempre incomplete e spesso difficilmente comprensibili.

(1) — А я вот к тебе, Петрович, **того**...

(2) — А я вот **того**, Петрович... шинель-то, сукно... вот видишь, (...) да вот только в одном месте немного **того**... на спине, да еще вот на плече одном немного попротерлось, да вот на этом плече немножко — видишь, вот и все.

(4) — Ну, да уж прикрепи. Как же этак, право, **того**!..

(5) — Ну, а если бы пришлось новую, как бы она **того**...

(6) «Этаково-то дело этакое, (...) я, право, и не думал, чтобы оно вышло **того**...

(8) «Так этак-то! вот какое уж, точно, никак неожиданное, **того**... этого бы никак... этакое-то обстоятельство!»

²⁴⁴*Akàkii Akàkievic, che a tempo debito aveva già sentito in sé una conveniente timidezza, si turbò un poco; e come poté, quando poté sciogliere la lingua, si diede a spiegare, con una prodigiosa profusione di in quanto – che, diceva, il cappotto era perfettamente nuovo, e ora glielo avevano strappato via in un modo inumano (...)* (Rebora, 1922).

(9) «Ну нет», — сказал Акакий Акакиевич, — теперь с Петровичем нельзя толковать: он теперь **того**... (...) а в это время я ему гривенничек и того, в руку, он и будет сговорчивее и шинель тогда и того...»

(10) «Ну, уж эти французы! что и говорить, уж ежели захотят что-нибудь того, так уж точно **того**...»

(11)— я ваше превосходительство осмелился утрудить потому, что секретари **того**... ненадежный народ...

(12) «А! так вот ты наконец! наконец я тебя **того**, поймал за воротник! твоей-то шинели мне и нужно!»

Grazie al contesto, gli interlocutori riescono sempre a inferire i messaggi di Akakij, sebbene il suo scopo comunicativo rimanga irrealizzato: durante i dialoghi con il sarto Petrovič (1) – (5) il tentativo di convincerlo a ricucire il vecchio cappotto fallisce. Allo stesso modo, il protagonista fallisce nel suo misero sforzo di rivolgersi alle autorità per ottenere giustizia in seguito al furto del cappotto. Akakij sembra quindi usare gli strumenti linguistici a sua disposizione in modo inefficace, generando spesso un non-senso che conduce al fallimento dell'aspetto perlocutivo del suo messaggio.

Per quanto riguarda la resa di *togo* nelle traduzioni italiane, i traduttori cercano, talvolta, di rendere proprio questo non-senso²⁴⁵, mentre, in altri casi, traducono la particella con parole o locuzioni che possano svolgere una funzione analoga in italiano, ovvero di ausilio nell'esposizione di un pensiero.

Come afferma Bazzanella (1995: 227), l'analisi dei segnali discorsivi può avvenire soltanto “sul piano dell'enunciazione e <deve> tenere conto delle variabili della comunicazione in corso”. Ognuna delle frasi da noi esaminate, infatti, si colloca in un determinato contesto e alcuni traduttori adattano la resa di *togo* all'enunciato in cui sono inseriti [un esempio molto chiaro emergerà nelle traduzioni dell'esempio (6)]. Tuttavia, ciò non costituisce la regola e, analizzando le versioni a nostra disposizione, abbiamo registrato due tendenze: alcuni traduttori impiegano un unico termine (o locuzione) valido

²⁴⁵ Per “non-senso” si intende qui il risultato di una porzione di discorso in cui vengono violate le norme logico-comunicative indispensabili alla realizzazione dello scopo comunicativo (Rigotti, Cigada 2004: 24).

per tutti i casi in cui la particella compare; altri, invece, non scelgono una parola fissa per rendere *togo* in italiano, ma variano di volta in volta in base al senso generale della frase. La prima tendenza genera più spesso non-senso rispetto alla seconda.

Queste scelte differenti producono due effetti distinti nel testo di arrivo: nel primo caso, il lettore percepisce, da un lato, l'insensatezza già sottolineata dalla voce narrante - "[...] *такими частицами, которые решительно не имеют никакого значения*" (Gogol', 1938: 149) -, dall'altro, l'esistenza di una parola – o segnale discorsivo - tipico del repertorio linguistico del protagonista, cosa che, effettivamente avviene anche nell'originale russo [Cfr. esempio (32)]. Nel secondo caso, invece, ossia variando ogni volta la traduzione di *togo*, la particella perde la sua tipicità, viene a mancare il senso di ricorrenza e non emerge un intercalare caratteristico. In questi casi, il testo sembra più scorrevole e, come detto, riduce il non-senso degli enunciati in cui ricorre. La scelta di non usare un unico equivalente traduttivo per la particella *togo* potrebbe essere dettata dunque da motivazioni di stile, dalla volontà di rendere il testo italiano più naturale, senza ricorrere all'aggiunta di elementi superflui e ridondanti. Talvolta, infatti, *togo* viene omesso nel testo di arrivo, lasciando spazio solamente a tre puntini di sospensione. Vediamo ora in modo più dettagliato le versioni dei quindici traduttori da noi considerati.

Fra coloro che hanno reso *togo* sempre con uno stesso traduttivo rientrano: Reborà (1922), che lo traduce con *in quanto*; Landolfi (1941), *coso*; Pacini-Savoj (1957), *quella cosa*; Julovic (1964), *vero e coso*; De Nardis (1993), *bè*; Guercetti (1995), *ciò*²⁴⁶. Le soluzioni di De Nardis e Guercetti corrispondono a segnali discorsivi frequentemente usati nel linguaggio parlato dell'italiano contemporaneo e potrebbero pertanto soddisfare l'intento di mantenere un intercalare costantemente ripetuto dal protagonista, e garantire uno stile il più naturale possibile (anche se, lo ribadiamo, i dialoghi di Gogol' sono da considerarsi delle caricature, dunque non totalmente rispondenti allo stile naturale e scorrevole della lingua autentica). Le restanti soluzioni suonano più obsolete all'orecchio del parlante italiano e presentano realizzazioni al limite del "non-senso".

Per chiarire quanto detto, riportiamo (1), (2) e (6) con le rispettive versioni italiane in ordine cronologico. Il simbolo "Ø" indica che la particella è stata omessa nella traduzione:

(1) — *А я вот к тебе, Петрович, того...*

²⁴⁶ Come vedremo De Nardis e Guercetti traducono diversamente solo l'esempio (6).

REBORA - 1922: «Eccomi ora da te, Petròvic, in quanto...».

DUCHESSA D'ANDRIA - 1937: «Ecco, Petrovic, che ti porto questo...»

LANDOLFI - 1941: — Ecco, t'ho portato, Petrovic, coso...

DEL BUONO - 1949: - Sono venuto da te, Petròvic, poiché... in quanto... coso.

L. PACINI-SAVOJ - 1957: Io, ecco, Petrovič, sono venuto, quella cosa...

G. PACINI - 1963: — Ecco, io ti, Petrovich, questo...»

JULOVIC - 1964: — Ecco... io, caro Petròvic, sarei venuto, vero e coso...

ZVETEREMICH - 1967: «E io, ecco, per te, Petrovič, questo...»

BAZZARELLI - 1980: «A, ecco qua, a te, Petrovič, io...»

MARIANO - 1986: «E io ecco, Petrovič, a te questo...»

BEFFA - 1986: “E io, ecco qua, a te, Petròvich, questo... ”.

DE NARDIS - 1993: «Ti ho portato, Petrovič, be'...».

PRINA - 1994: «Ecco dunque che a te, Petrovič, quel...»

GUERCETTI - 1995: «Ecco, sono qui, Petrovič, cioè...»

LEGITTIMO - 2001: «Sì, ecco... io da te, Petrovič... cioè...»

(2) — А я вот мого, Петрович... шинель-то, сукно... вот видишь, (...) да вот только в одном месте немного мого... на спине, да еще вот на плече одном немного попотерлось, да вот на этом плече немножко — видишь, вот и все.

REBORA - 1922:

«Ora appunto eccomi in quanto, Petròvic... questo cappotto qua, il panno... ecco vedi, (...) ma ecco solo un po' in un posto in quanto... sul dorso, ma ancora ecco su una spalla è un po' liso, ma ecco su quest'altra è un pochino... vedi? Ecco appunto qui tutto.

DUCHESSA D'ANDRIA - 1937:

«Ecco, Petrovic, che ti porto questo... questo pastrano... il panno... vedi, (...) soltanto in un punto Q... sulla schiena, e anche qui, sulla spalla... è un po' lacerato... vedi? Soltanto qui.

LANDOLFI - 1941:

— Ed ecco io, Petrovic, coso... il mantello... il panno... vedi che in tutti i punti regge benissimo (...) ma qui soltanto in un posto è un po' ... coso... sul dorso, e qui su una spalla s'è consumato un tantino, e anche su quest'altra spalla un pochetto... vedi? Quest'è tutto.

DEL BUONO - 1949:

— Appunto, eccomi proprio qua, in quanto, Petròvic... coso... il cappotto... il panno qua... ecco vedi, (...) ma ecco soltanto in un punto... poiché... sul dorso, e qui su una spalla è un poco liso, ma ecco anche su quest'altra è un pochino ... vedi? Ecco qua tutto, appunto.

PACINI-SAVOJ - 1957:

Io, ecco, già, Petrovič, quella cosa... il cappotto, già, il panno., ecco, vedi, (...) ed ecco: soltanto in una parte è un po' quella cosa... sulla schiena, e poi, ecco, su una spalla è un po' consumato in una parte è un po' già, ecco, su questa spalla è un poco... vedi? Non c'è altro.

G. PACINI - 1963:

— Io, ecco... questo, Petrovich... il cappotto... il panno, ecco vedi, (...) soltanto in un punto è un po' ... così... sulla schiena, e anche su una spalla si è un po' logorato, ecco, qui su questa spalla, è tutto qui, vedi.

JULOVIC - 1964:

— Ecco... io, già, vero e coso... il cappotto, già, la stoffa... ecco, vedi, (...) Ecco... soltanto in un punto è vero e coso, sulla schiena e poi... su questa spalla è un po' consumato, già, ecco, proprio su questa spalla è un po' ... vedi? Per il resto...

ZVETEREMICH - 1967:

«E io, ecco, **che cosa**, Petrovič... il cappotto, già, il panno... ecco vedi (...) solo che in un posto è un poco **così**... sulla schiena, e poi anche su una spalla s'è un poco consumato; sì, ecco, su questa spalla un po'... ecco tutto.

BAZZARELLI - 1980:

«Ecco qua, Petrovič... **Ø** il cappotto, il panno... ecco vedi, (...), ma in un posto, però, è, **come dire**.... sulla schiena, e anche su una spalla, si è un po' sciupato, ecco qui, un po' su questa spalla, lo vedi? Ed è tutto qui.

MARIANO - 1986:

«E io **questo** ecco, Petrovič... questo cappotto... la stoffa... ecco vedi, (...) ma ecco solo in un punto c'è un po' **Ø**..., e poi ancora su una spalla s'è un po' consumato, ecco qui un pochino su questa spalla, vedi, ecco tutto.

BEFFA - 1986:

«Ecco qua, Petròvich...**Ø** il cappotto, il panno... ecco vedi, (...) solo che in un punto è, **come dire**, sulla schiena, e poi anche su una spalla, si è un po' consumato, ecco qui, un po' su questa spalla, lo vedi? ed è tutto qui.

DE NARDIS - 1993:

«Ecco... insomma, Petrovič... **Ø** il cappotto, il panno... vedi bene, (...) ecco in questo unico punto, però, si è un po'... **bè** sulla schiena, e anche qui su una spalla si è un po' consumato, ecco su questa spalla un pochino lo vedi? Tutto qui.

PRINA - 1994:

«Ecco dunque che io **questo**, Petrovič... la mantella, il panno... ecco, vedi, (...) ecco solo in un punto s'è un pochino **come**... sulla schiena, e anche sulla spalla s'è logorato un pochetto, ecco qui su questa spalla un pochettino, vedi, ecco tutto.

GUERCETTI - 1995:

«Ecco io, cioè, Petrovič... il cappotto, il panno..., ecco, vedi (...) solo in un punto è un pochino, cioè... sulla schiena e qui, ecco, su una spalla è un tantino liso e su quest'altra spalla un pochettino vedi, ecco tutto.

LEGITTIMO - 2001:

«Ecco io, Petrovič... cioè, il cappotto, la stoffa..., ecco vedi, (...) ecco solo in un punto è un po' Ø... sulla schiena, e anche sulla spalla è un tantino consumata, sì qui su questa spalla, vedi, tutto qua.

Si nota che i traduttori che cercano di mantenere una strategia di resa univoca, usano lo stesso termine indipendentemente dalla sua posizione all'interno dell'enunciato, mentre chi decide di modificare di volta in volta la traduzione di *togo* opta per varianti molto diverse anche all'interno dello stesso paragrafo. Bazzarelli, ad esempio, traduce con "io..." il *togo* presente in (1), omette il primo *togo* di (2) e inserisce il segnale discorsivo "come dire" per il secondo *togo* di (2).

Nell'esempio (6), invece, come si diceva, l'adattamento al contesto e al co-testo è più evidente anche nelle versioni di alcuni traduttori che in tutti gli altri casi si sono attenuti ad un unico termine per la resa di *togo* (cfr.: De Nardis e Guercetti). Molti qui trattano *togo* alla stregua di *tak*, traducendo con l'italiano "così" o con espressioni sinonimiche. La particella *togo* sembrerebbe quindi assumere il secondo sotto-significato introdotto dal MAS, ovvero di riempitivo "при сообщении о каком-л. неблагоприятии". Questo potrebbe giustificare il frequente impiego di "così" come variante traduttiva. Cfr. ad es. la versione di Pacini: "io davvero non mi sarei mai immaginato che andasse a finire così" (dove "così" potrebbe essere parafrasato come: "in modo così spiacevole"):

(6) «Этаково-то дело этокое, (...) я, право, и не думал, чтобы оно вышло **того**...

REBORA - 1922:

«Che razza di una faccenda è mai questa» (...) Io, in fede mia, non pensavo che la cosa avesse per risultato in quanto...»;

DUCHESSA D'ANDRIA - 1960:

«Che cosa! Ma che cosa!» (...) «davvero non pensavo che finirebbe così!... ».

LANDOLFI - 1941:

«Che razza di brutta faccenda, (...) certamente io non credevo che sarebbe andata...
coso...»

DEL BUONO - 1949:

«Che razza di storia è questa» (...) «in fede mia non pensavo proprio che sarebbe andata così...in quanto...».

PACINI-SAVOJ - 1957:

«Che specie di affare di che specie (...) «Io davvero non pensavo neanche che ne venisse fuori quella cosa...».

G. PACINI - 1963:

"Che faccenda è questa, (...) io davvero non mi sarei mai immaginato che andasse a finire così...

JULOVIC - 1964:

«Che pasticcio brutto brutto brutto...» (...) «E chi si andava a immaginare, vero e coso, che ne venisse fuori...».

ZVETEREMICH - 1967:

«Bella storia, bella,» (...) «davvero non l'avrei mai pensato che sarebbe andata a finire così...»

BAZZARELLI - 1980:

«Così proprio così! ecco in fondo che cosa, davvero, mai e poi mai da aspettarsi in quanto...mai e poi mai non si... un imbroglio di tal fatta!».

BEFFA - 1986:

"Bell'affare, questo" (...) "davvero non pensavo proprio che sarebbe andata a finire così..." "

MARIANO - 1986:

«Guarda che razza di storia, guarda» (...) «non pensavo proprio che sarebbe finita così...»

DE NARDIS - 1993:

«Questo è proprio un bell'affare» (...) «davvero non avevo previsto che sarebbe andata così...»

PRINA - 1994:

«Tale è dunque la faccenda» (...) «non avevo davvero nemmeno pensato che sarebbe andata a questo modo...»

GUERCETTI - 1995:

«Ma che razza di storia» (...) «cioè io davvero non pensavo neanche che sarebbe andata così...»

LEGITTIMO - 2001:

«Ma guarda che storia (...) io, davvero, non ci pensavo neanche che sarebbe andata a finire così...»

È interessante ora vedere come i traduttori che non optano per un unico equivalente di *togo* si sono cimentati nella traduzione della nota metalinguistica del narratore sull'uso cronico di *togo* da parte di Akakij, contenuta nell'esempio (32) che riproponiamo per intero:

(32) *Акакий Акакиевич уже заблаговременно почувствовал надлежащую робость, несколько смутился и, как мог, сколько могла позволить ему свобода языка, изъяснил с прибавлением даже чаще, чем в другое время, частиц «того», что была-де шинель совершенно новая, и теперь ограблен бесчеловечным образом (...)*

DUCHESSA D'ANDRIA - 1937:

(...) raccontò, con molti particolari, Ø (...)

DEL BUONO - 1949:

(...) spiegò, aggiungendo persino più del solito: poiché, in quanto e coso (...)

PACINI - 1963:

(...) e si mise a spiegare (...) aggiungendo più spesso del solito la particella "ecco" (...)

ZVETEREMICH – 1967:

(...) spiegò, intercalando ancor più del solito la parola «ecco» (...)

BAZZARELLI - 1980:

(...) spiegò, usando in modo ancor più frequente che altre volte la particella «quello» (...)

BEFFA - 1986:

(...) spiegò, intercalando ancor più di frequente la particella "quello" (...)

MARIANO - 1986:

(...) spiegò, aggiungendo persino più spesso del solito la parolina «ecco» (...)

PRINA -1994:

(...) chiari, con l'aggiunta persino più frequente del solito della paroletta «quello» (...)

LEGITTIMO – 2001:

(...) spiegò, (...) ricorrendo ancora più spesso del solito alla particella cioè (...)

In tutti i casi, ad eccezione della Duchessa D'Andria, che omette questa frase nel testo di arrivo, i traduttori scelgono una delle numerose varianti impiegate nella resa dei vari dialoghi, e talvolta non si tratta nemmeno della più usata.

Riassumiamo ora, nella tabella 18, tutte le possibili rese di *togo* riscontrate nelle quindici traduzioni:

Tab 18. La resa di *togo* nelle traduzioni italiane del *Cappotto*

RESA	TRADUTTORI
in quanto	Rebora, Del Buono
coso	Landolfi, Del Buono
poiché	Del Buono
quella cosa	Pacini-Savoj
questo	Pacini, Duchessa D'Andria, Zveteremich, Mariano, Beffa, Prina, Legittimo
vero e coso	Julovic
io...	Bazzarelli
cioè	Guercetti, Legittimo. Prina
ecco	Bazzarelli, Beffa, Mariano
come dire	Bazzarelli, Beffa
così	Pacini, Zveteremich, Bazzarelli, Beffa, De Nardis, Guercetti, Legittimo
quello/quel	Mariano, Bazzarelli, Prina
che cosa	Zveteremich
come	Prina
già	Zveteremich, Beffa
∅	Duchessa D'Andria, G. Pacini, Zveteremich, Bazzarelli, Beffa, Mariano, De Nardis, Legittimo

In base alla classificazione funzionale operata da Bazzanella in (1995: 232-257), dunque, *togo* è stato reso come riempitivo (*come dire*), indicatore di correzione (*cioè*), indicatore

di esemplificazione (*ecco*) o, ancora, come indicatore di riformulazione (*bé*). Non rientrano, invece, normalmente, nel novero dei segnali discorsivi i dimostrativi *questo*, *quello* o espressioni quali *in quanto*, *coso*.

Più in generale, il confronto tra le quindici versioni fa luce su un aspetto rilevante: le scelte traduttive possono essere dettate non solo da criteri di tipo linguistico, che considerano il singolo enunciato nel suo contesto, ma anche da scelte riferite all'economia del racconto o del romanzo nel suo complesso. Ciò consentirà di trarre alcune conclusioni in merito all'uso della linguistica dei corpora come strumento di analisi.

5.3.2. La resa dell'ellissi verbale in italiano

Gli esempi analizzati in questo paragrafo riguardano quei fenomeni di ellissi del verbo tipici del parlato russo e già descritti nel paragrafo 5.2.3:

(5) — Ну, а если бы пришлось новую, как бы она того...

(9) (...) а в это время я ему гривенничек и того, в руку, он и будет сговорчивее и шинель тогда и того...»

In (5) vediamo l'omissione di un infinito retto dalla costruzione impersonale “*a esli by prišlos Ø*”. Il verbo omissso, a sua volta, regge il sintagma nominale all'accusativo “*novuju*” anch'esso incompleto. Questo secondo tipo di ellissi, tuttavia, come è evidente dalla resa dei quindici traduttori, è possibile anche in italiano (spesso con l'ausilio del clitico *ne* e del pronome indefinito: *e se bisognasse farne uno nuovo?*) dal momento che il sostantivo omissso (*šinel'*) è menzionato nel contesto immediatamente precedente. L'ellissi verbale, al contrario, è un fenomeno che non si verifica di norma nelle lingue romanze, dove il nucleo principale della frase rimane il verbo, salvo in alcuni casi laddove un nome o un aggettivo svolgono una funzione predicativa (cfr. l'italiano.: *bella la tua giacca!* / *Ecco il treno*).

In (9) si evidenzia invece l'omissione di un verbo indipendente, appartenente al campo semantico del “dare”, qui espresso da una locuzione ellittica (*я эти гривенничек в руку*).

L'analisi di questo brano in modo contrastivo, con a fronte originale russo e traduzione, permette di fare luce sulle strategie che si possono adottare in italiano per rendere un verbo che di fatto non compare nel testo originale. A differenza dell'uso

eccessivo delle particelle – prima fra tutte *togo* – che anche nel testo fonte aveva lo scopo di rendere il linguaggio di Akakij Akakievič caricaturale, l’ellissi sintattica è impiegata dal protagonista in modo assolutamente ordinario. Si presuppone, pertanto, che i traduttori rendano queste porzioni di testo in modo che anche in italiano rispondano all’uso ordinario e risultino non marcate. Presentiamo dunque gli estratti tradotti così da verificare questa ipotesi e identificare le diverse strategie di resa.

(5) (...) *а шинель уж, видно, вам придется новую делать. При слове «новую» у Акакия Акакиевича затуманило в глазах, (...) — Как же новую?— сказал он (...) — Да, новую, — сказал с варварским спокойствием Петрович. — Ну, а если бы пришлось новую, как бы она того...*

REBORA - 1922:

(...) *E il cappotto allora, se non sbaglio, vi toccherà farlo nuovo». Alla parola «nuovo» Akàkii Akàkievic vide nero (...) «Come mai nuovo?» (...) «Già, nuovo» sentenziò con crudele tranquillità Petròvic. «Ma via, e se si fosse obbligati a uno nuovo, come sarebbe in quanto...».*

DUCHESSA D’ANDRIA - 1937:

(...) *Ma voi dovete farvi un pastrano nuovo». Alla parola «nuovo» gli occhidi Akaki Akakievic si velarono (...) «Come, nuovo?» «Si, nuovo», disse con crudele tranquillità Petrovic. «E, se mai... poi... uno nuovo? ...»*

LANDOLFI - 1941:

(...) *vi converrà farvene uno nuovo. Alla parola nuovo ad Akakij Akakievič si annebbiò la vista, (...) — Come, nuovo? (...) — Eh già, nuovo, — ridisse Petrovič con calma feroce— Ebbene, e se per caso uno nuovo, che cosa... coso...*

DEL BUONO - 1949:

Per quanto riguarda il cappotto, statemi a sentire, fatevene uno nuovo. Alla parola nuovo ad Akàkii Akàkievic si confuse la vista (...) — Come sarebbe a dire: nuovo? (...) — Nuovo, ci vuole, sentenziò, crudelmente tranquillo, Petròvic. — Via, facciamo il caso che uno... nuovo fosse obbligato... cosa... come... coso... in quanto...

PACINI-SAVOJ - 1957:

(...) in quanto al cappotto, è chiaro che vi toccherà farne uno nuovo. Alla parola «nuovo» ad Akakij Akakievič si annebbiarono gli occhi, (...) - Come, nuovo? (...) - Già: nuovo! disse con flemma barbara Petrovič. - E, a doverlo far nuovo, quella cosa, come...?

G. PACINI - 1963:

Quanto al cappotto è chiaro che ormai dovrete farvene uno nuovo.» A sentire la parola "nuovo" ad Akakij Akakievich si annebbiarono gli occhi (...) «Come sarebbe a dire nuovo? (...) — Già, nuovo, — ribadì Petrovich, con una calma crudele. — E se volessi farmelo nuovo, che cosa mi...

JULOVIC - 1964:

Quanto al cappotto, se ne dovrà fare uno nuovo, questo è certo. A sentir dire «nuovo», Akàkii Akàkievic si sentì mancare (...) — Come? nuovo? — biascicò. — Sicuro, nuovo — sillabò tranquillo Petròvic, con una certa crudeltà. — E a farlo nuovo, vero e coso... quanto...

ZVETEREMICH - 1967:

(...) e di cappotto dovrete farvene uno nuovo.» Alla parola «nuovo» Akàkij Akakièvič si sentì annebbiare la vista (...) «Come sarebbe, nuovo?» (...) «Sì, nuovo» disse con crudele flemma Petrovič. «Be', e se per caso uno nuovo, cosa, quanto...»

BAZZARELLI - 1980:

È chiaro che deve proprio farsi un cappotto nuovo.» Alla parola "nuovo" gli occhi di Akakij Akakievič si ricoprirono di nebbia (...) «Come, uno nuovo?» (...) «Sì, uno nuovo», disse Petrovič con feroce tranquillità. «Ma, se bisogna farne uno nuovo, allora, quello...»

BEFFA - 1986:

(...) e di cappotto dovrà farsene uno nuovo". Alla parola "nuovo" ad Akàkij Akakièvich si annebbiò la vista (...) "Come, uno nuovo?"(...) Sì, uno nuovo ", disse Petròvich con feroce tranquillità. "Ma, se bisogna farne uno nuovo, allora, quanto... "

MARIANO - 1986:

(...) «ed è evidente che dovrete farvi un cappotto nuovo.» Alla parola «nuovo» gli occhi di Akakij Akakievič s'annebbiarono, (...) «Come sarebbe nuovo?» «Già, nuovo» disse Petrovič con calma feroce. «Be', e se me ne dovessi **fare uno nuovo**, quanto..»

DE NARDIS - 1993:

(...) ma è evidente che ormai vi toccherà farvi un cappotto nuovo.» Alla parola «nuovo», ad Akakij Akakievič si velarono gli occhi di nebbia (...) «Come nuovo (...) «Sì, uno nuovo», disse con calma barbara Petrovič. «Ah, ma se toccasse **farne uno nuovo**, come sarebbe... be'...»

PRINA - 1994:

(...) ma è evidente che vi toccherà farvi una mantella nuova.» Alla parola «nuova» Ad Akakij Akakievič s'annebbiò la vista (...) «Come sarebbe nuova?» (...) «Sì, nuova», disse Petrovič con barbara calma. «E se per caso me ne toccasse una nuova, allora come quella...»

GUERCETTI - 1995:

(...) quanto al cappotto, si vede che le toccherà farsene uno nuovo.» Alla parola «nuovo» ad Akakij Akakievič si annebbiò la vista (...) «Come nuovo?» (...) «Sì, nuovo», disse con barbara tranquillità Petrovič. «Be', ma se ne occorresse uno nuovo come sarebbe...»

LEGITTIMO - 2001:

Le toccherà farsene uno nuovo». Quando sentì la parola «nuovo», ad Akakij Akakievič si ottenebrò la vista (...) «Come sarebbe a dire nuovo? (...) «Sì, nuovo» ripeté con barbara tranquillità. «Ma se ne servisse uno nuovo, come allora...»

La frase ellittica presenta qui una struttura sintattica analoga alla frase iniziale di Petrovič “вам придется новую делать”, di cui costituisce la variante incompleta con l'ellissi dell'infinito *delat'*. Vediamo che emergono due tendenze: da un lato quella di riprendere la struttura sintattica russa, violando le norme dell'italiano, che non prevedono in questi casi un'ellissi del verbo, e, dall'altro, rendere la naturalezza del testo fonte anche nella traduzione italiana inserendo un verbo, dal momento che l'omissione in russo non è marcata, bensì un espediente tipico della lingua parlata. La prima scelta si realizza in otto

casi, in cui l'infinito è omesso anche in italiano: Reborà, Landolfi, Del Buono, Duchessa D'Andria, Zveteremich, Prina, Guercetti e Legittimo. Prina, Guercetti e Legittimo, tuttavia, modificano le relazioni sintattiche originarie impiegando un unico verbo che mantiene la semantica del testo fonte (la necessità di avere un cappotto nuovo): “*se ne toccasse una nuova*”, “*se ne servisse uno nuovo*”, “*se ne occorresse uno nuovo*”. Questi tipi di resa non producono dunque alcun tipo di ellissi nel testo d'arrivo, se non quella del sostantivo che, come già detto, è normalmente utilizzata anche nella lingua italiana. Diverso è il caso degli altri cinque traduttori, dove l'omissione del verbo nel testo d'arrivo può essere percepita come un'anomalia. In questi casi si contribuisce ad accentuare ulteriormente il carattere disconnesso e insensato, tipico del modo di esprimersi del protagonista.

I restanti traduttori – Pacini-Savoj, Pacini, Julovic, Bazzarelli, Beffa, Mariano e De Nardis – inseriscono invece l'infinito mancante nelle loro versioni, eliminando così l'ellissi; nessuno si cimenta a ricercare un'alternativa semanticamente più precisa al polivalente verbo “fare”, che riprende il “fare” della frase completa pronunciata da Petrovič e assume in questo caso il suo significato ipsivalente, ossia “fabbricare”, “costruire”. Questa scelta consente di rispettare non solo la semantica del testo fonte, ma anche le norme sintattico-grammaticali del testo d'arrivo e quelle dello stile parlato italiano, che vedono il verbo “fare” come alternativa preferenziale rispetto a termini più precisi e ricercati.

(9) (...) *а в это время я ему гривенничек и того, в руку, он и будет сговорчивее и шинель тогда и того...»*

REBORA - 1922:

(...) **io** invece nel frattempo **li una monetina in mano**, in quanto... lui pure sarà più compiacente, e il cappotto allora pure, in quanto...».

DUCHESSA D'ANDRIA - 1937:

(...) **io**, in quel momento, **gli metterò in mano un grivennik**, sarà più abbordabile e allora, il pastrano...».

LANDOLFI - 1941:

(...) in questa **io gli allungherò dieci copechi**, e lui diventerà più conciliante, e allora il mantello...coso»

DEL BUONO - 1949:

(...) **io** invece, ecco, **una monetina in mano**... Poiché...coso...lui diventerà conciliante, ed il cappotto allora pure, poiché...».

PACINI-SAVOJ - 1957:

(...) e **io** allora, e quella cosa, **gli faccio scivolare in mano dieci copeche**: lui diventerà più trattabile, e allora il cappotto, e quella cosa...».

G. PACINI - 1963:

(...) ma proprio allora arrivo io, **gli metto in mano dieci copeche**, lui diventerà più conciliante e così la faccenda del cappotto..."

JULOVIC - 1964:

(...) Allora io, vero e coso, **con dieci copechi che gli faccio balenare sotto il naso**, me lo rendo più malleabile e il cappotto, vero e coso...»

ZVETEREMICH - 1967:

(...) e io allora proprio in quel momento **gli metto un grivennik nella mano**, e lui sarà più trattabile e allora pure il cappotto, già quello...»

BAZZARELLI - 1980:

(...) e io in quel momento **gli regalerò dieci copechi, in mano glieli darò**, e lui diventerà più ciarliero, e il cappotto, così lui...»

BEFFA - 1986:

(...) e io allora proprio in quel momento **gli metterò in mano un decino**, e lui diventerà più ciarliero e allora pure il cappotto, già quello... "

MARIANO - 1986:

(...) e allora io **gli faccio scivolare un grivennik nella mano**, e lui sarà più affidabile e allora il cappotto e quello...»

DE NARDIS - 1993:

(...) e in quel momento **gli metterò in mano un bel grivennik**, lui allora sarà più loquace e il cappotto quindi...bè...»

PRINA - 1994:

(...) e allora io **gli metterò in mano un soldo da dieci kopejki**, e quello, col soldo in mano, sarà anche più arrendevole, e allora anche la mantella quello...»

GUERCETTI - 1995:

(...) a quel punto arrivo io e **to' dieci copeche in mano**; allora lui diventerà più trattabile e così anche il cappotto, cioè...»

LEGITTIMO - 2001:

(...) e io al momento giusto **gli allungo una monetina**, lui diventa più trattabile e il cappotto allora...»

Nella resa di questo estratto, la maggior parte dei traduttori sceglie di inserire un verbo nel testo d'arrivo. Solo in tre casi ciò non si verifica e precisamente in Rebora, Del Buono e Guercetti. Quest'ultima sceglie tuttavia di inserire *to'*, un'interiezione colloquiale con valore verbale che funge da invito a prendere qualcosa che viene offerto e in contesti confidenziali può sostituire gli imperativi “prendi” o “tieni”, permettendo così di ovviare alla mancanza del verbo²⁴⁷. Le varianti di Rebora e Del Buono, invece, così come nelle traduzioni dell'esempio (5), mantengono l'ellissi, che essendo innaturale per l'italiano contribuisce a conferire alle parole di Akakij Akakievič un andamento impacciato e frammentario.

I restanti traduttori considerano, nella maggior parte dei casi, “*ja emu grivennik i togo, v ruku*” come un'unica proposizione semplice e riempiono quindi la posizione del

²⁴⁷ Riportiamo la definizione di Treccani: **to'** interiez. – Troncamento di *togli*, imperativo del v. *togliere*. **1.** Con valore verbale, equivale a *togli!*, *prendi!*, *tieni!*, e si dice, ma solo in tono molto confidenziale, porgendo ad altri qualche cosa: *to' un cioccolatino!*, *to', assaggia un goccetto!* (<http://www.treccani.it/vocabolario/to/>) (gennaio 2017).

predicato eliso con un verbo appartenente al campo semantico del “mettere”, che possa quindi creare una collocazione con il seguente “in mano”. Fanno eccezione Landolfi e Legittimo, con il loro “allungare” e la mancata resa di *v ruku*; e Bazzarelli, che tratta “*ja emu grivennik*” e “*v ruku*” come due proposizioni ellittiche separate, inserendo quindi due verbi: *gli regalerò dieci copechi, in mano glieli darò*. Negli altri casi è proprio il verbo “mettere” la variante prediletta. Ad esso si alternano alcune soluzioni più distanti dall'originale: “*gli faccio scivolare in mano*”, di Pacini-Savoj e Mariano e “*gli faccio balenare sotto il naso*” di Julovic, che sostituisce l’espressione “*in mano*” con “*sotto il naso*”. L’espressione “far scivolare in mano”, in particolare, rende l’idea di quello che il grivennik vuole essere: la “bustarella” usata da Akakij per corrompere il sarto Petrovič ed invogliarlo ad assecondare la sua richiesta.

Vediamo dunque che, sebbene le frasi siano complete e comprensibili dal punto di vista semantico, la resa dell’ellissi sintattica è fortemente legata al contesto e anche a fronte di esempi molto espliciti, quali “*ja domoj*”, può essere mantenuta solo tenendo in considerazione la *konsituacija*. Inoltre, anche in questo caso, vi sono eccezioni in cui la resa di una frase ellittica dipende da scelte legate all’universo del racconto, aspetto da non sottovalutare nell’indagine linguistica di un testo letterario.

Bibliografia

- Aijmer K., Foolen A., Simon-Vandenberghe A. (2006). *Pragmatic markers in translation: a methodological proposal*. In Fischer K. (ed.), *Approaches to discourse particles*. Oxford, Amsterdam: Elsevier, pp 101-114.
- Bazzanella C. (1995). *I segnali discorsivi*. In Renzi L., Salvi G., Cardinaletti A. (a cura di), *Grande grammatica italiana di consultazione III*. Il Mulino: Bologna, pp. 225-257.
- Bazzanella C. (2006). *Discourse markers in Italian: towards a compositional meaning*. In Fischer K. (a cura di), *Approaches to discourse particles*. Oxford, Amsterdam: Elsevier, pp. 449-464.
- Biber D (2011). *Corpus linguistics and the study of literature. Back to the future?* In *Scientific study of literature*. 1:1. Amsterdam: John Benjamins Publishing Company, pp. 15-23.
- Bonola A. (2008). *Strategie di organizzazione del testo russo. Particelle e diatesi*. Milano: Università Cattolica del Sacro Cuore – Diritto allo studio.
- Dibrova E. I. (2008) *Sovremennyj Russkij Jazyk. Teorija. Analiz jazykovnyjč edinic*. Moskva: Akademija.
- Dmitrievskaja L. N. (2009). *Novyj vzgljad na obraz "malen'kogo čeloveka" v povesti N.V Gogolja "Sinel"*. In Avgustovskij pedsovet. *Russkaja klassičeskaja literatura: sovremennoe pročtenie i metodika izučenija* –Moskva: MIOO, pp. 36-41.
- Ejchenbaum B. M. (1919). *Kak sdelana "Šinel" Gogolja*. (<http://www.opojaz.ru/manifests/kaksdelana.html>) (dicembre 2016).
- Evgeneva A. P. (1957-1984). *Malyj Akademičeskij slovar'*. Moskva: Institut russkogo jazyka Akademii nauk SSSR.
- Fernandez J. M. M. (1994). *Les particules énonciatives dans la construction du discours*. Presses universitaires de France.
- Fischer K. (2006). *Towards an understanding of the spectrum of approaches to discourse particles: introduction to the volume*. In Fischer K. (ed.), *Approaches to discourse particles*. Oxford, Amsterdam: Elsevier, pp. 1-19.
- Fischer-Starcke B. (2010). *Corpus linguistics in literary Analysis. Jane Austen and her Contemporaries*. London: Continuum.

- Gobber G. (1997). *Segnali discorsivi e teoria linguistica*. In *L'Analisi linguistica e letteraria*, 1, ISU, Università Cattolica del Sacro Cuore, pp. 262-273.
- Gobber G (2002). *La sartoria del testo: connettori fra semantica e pragmatica (con attenzione alla lingua francese)*. In «Mots Palabras Words», 1/2002.
- Gobber G. (2004). *Su alcuni usi dei termini Partikel e častica nelle grammatiche del tedesco e del russo*. In Milani C., Finazzi R. B. (a cura di), *Per una storia della grammatica in Europa*. Atti del convegno 11-12 settembre 2003, Milano: I.S.U., pp. 327-348.
- Gogol' N.V. (1937-1952). *Polnoe sobranie sočinenii v četyrnadcati tomach*. Leningrad: Izadel'stvo Akademii nauk SSSR, 1938, T. 3, Povesti, pp. 139-174.
- Hori M. (2004). *Investigating Dickens' Style. A collocational Analysis*. Basingstoke: Palgrave Macmillan.
- Inkova O. (2014). *Tradurre il titolo: le traduzioni italiane del "Cappotto" di Gogol'*. In *Kwartalnik neofilologiczny*. Université de Varsovie, LXI, 1, pp. 41-56.
- Inkova O. (2016). *Le prime traduzioni romanze del "Cappotto" di Gogol': precisazioni cronologiche e traduttologiche*. In *Diversité et identité culturelle en Europe*, tome 13/2, București: Editura Muzeul Literaturii Române, pp. 93-120.
- Kachaturyan E. (2011). *Una classificazione dei segnali discorsivi in italiano*. In Kachaturyan E. (ed), *Discourse markers in romance languages*. Oslo Studies in Language 3 (1), pp. 95–116.
- Kiseleva K., Pajar D. (2003). *Diskursivnye slova russkogo jazyka: kontekstnoe var'irovanie i semantičeskoe edinstvo*. In Kiseleva K., Pajar D. (a cura di) *Diskursivnye slova russkogo jazyka: kontekstnoe var'irovanie i semantičeskoe edinstvo*. Moskva: "Azbukovnik", pp. 8-26.
- Kovalev V. (2000). *Il Kovalev. Dizionario italiano-russo, italiano-russo*. Bologna: Zanichelli.
- Lekant P. A. (2004). *Sintaksis prostovo predloženiija v sovremennom russkom jazyke*. Moskva: Vyššaja Škola.
- Mahlberg M. (2007). *Corpus stylistics: bridging the gap between linguistic and literary studies*. In Hoey M., Mahlberg M., Stubbs M., Teubert W (eds), *Text, Discourse and Corpora: Theory and Analysis*. London: Continuum, pp. 219-246.
- Mathesius V. (1967). *O sistemnom grammatičeskom analize*. In Kondrašov N. A. (a cura di), *Pražskij lingvističeskij kružok*. Moskva: Progress, pp.226-239.

- Mazzucchelli S. (2005). *La letteratura russa in Italia nell'editoria dell'entre-deux-guerres (1919-1939)*. Tesi di dottorato, Università degli studi di Milano.
- Mel'čuk I. A. (1974). *O sintaksičeskom nule* in A. A. Cholodovič (pod red.), *Tipologija passivnych konstrukcij. Diatezy i zalogi*. Leningrad: Nauka, pp. 343-361.
- Nikolaeva T. M. (1985). *Funkcii častic v vyskazyvanii*. Moskva: Nauka.
- Onipenko N. K., Sidorova M. Ju., Zolotova G. A. (2004). *Kommunikativnaja grammatika russkogo jazyka*. Rossijskaja Akademija Nauk: Moskva.
- Ožegov S. I., Švedova N. Ju. (a cura di) (1992). *Tolkovyj Slovar' Ožegova*. Moskva: Russkij jazyk.
- Padučeva E. V. (1997). *Kto že vyšel iz "Šineli" Gogolja? (O podrazumevaemych subektach neopredelennyh mestoimenij)*. Izvestija AN, serija literatury i jazyka, t. 36, № 2, pp. 20-27.
- Pala V. (2009). *Tommaso Landolfi traduttore di Gogol'*. Roma: Bulzoni Editore.
- Peškovskij A. M. (2001). *Russkij sintaksis v naučnom osveščanii*. Moskva: Jazyki slavjanskoj kul'tury.
- Rebora C. (1990). *Annotazioni*. In Gogol' N. V. *Il Cappotto*. Milano: Feltrinelli, pp. 69-97.
- Rigotti E., Cigada S. (2004). *La comunicazione verbale*. Milano: Apogeo.
- Salvi G. (1988). *La frase semplice*. In Renzi L. (a cura di), *Grande grammatica italiana di consultazione*, vol. I, Bologna: Il Mulino, pp. 29-113.
- Scandura C. (1997). *Osservazioni a proposito delle traduzioni italiane del racconto di Gogol' Šinel'*. In «Russica Romana», IV, 1997, pp. 283-309.
- Serianni L. (1991). *Grammatica italiana: italiano comune e lingua letteraria*. Torino: UTET.
- Skovorodnikov A. P. (1973). *O kriterii elliptičnosti v russkom sintaksise*. In Voprosy jazykoznanija, Moskva, № 3.
- Širiaev E. N. (2001). *Semantiko-sintaksičeskaja struktura razgovornogo dialoga*. In Russkij jazyk v naučnom osveščanii, № 1, pp. 132-147.
- Šmelev A. D. (2002). *Russkij jazyk i vnejazykovaja dejstvitelnost'*. Moskva: Jazyki slavjanskoj kul'tury.
- Švedova N. Ju. (1980). *Russkaja Grammatika*. Moskva: Nauka. (<http://www.rusgram.narod.ru/>) (dicembre 2016).

- Tortorelli G. (1996). *«L'Italia che scrive» 1918-1938. L'editoria nell'esperienza di A.F. Formiggini*. Milano: Franco Angeli.
- Toury G. (2012). *Descriptive Translation Studies – and beyond. Revised edition*. Amsterdam, Philadelphia: John Benjamins Publishing.
- Ušakov D. N. (a cura di) (1935-1940). *Tolkovyj slovar' russkogo jazyka*. V 4 tomach. Moskva: Gosudarstvennyj institut "sovetskaja enciklopedija; OGIZ (t.1); Gosudarstvennoe izdatel'stvo inostrannyh i nacional'nych slovarej (t. 2-4).
- Valgina N. S. (2000). *Sintaksis Sovremennogo Russkogo Jazyka*. Moskva: 'Agar'.
- Vinogradov V. V. (1975). *O kategorii modal'nosti i modal'nych sloвах v russkom jazyke*. In Švedova N. Ju. (a cura di) *Issledovanija po russkoj grammatiki: izbrannye trudy*. Moskva: Nauka, pp. 53-87.
- Vinogradov V. V. (2001). *Russkij jazyk (grammatičeskoe učenie o slove)*. Zolotova G. A. (ed), 4-e izd. Moskva: Russkij Jazyk.
- Vjatkina S. V. (2009). *Sintaksis sovremennogo russkogo jazyka: učebnik*. Moskva: Akademija.
- Zemskaja E. A., Kitajgorodskaja M. V., Širjaev E. N. (1981). *Russkaja razgovornaja reč'. Obščie voprosy. Slovoobrazovanie. Sintaksis*. Moskva: Nauka.
- Zyrjanov O. V. (2004). *Kto vyšel iz gogolevskoj "šineli"?* Izvestija Ural'skogo Federal'nogo Universiteta. Serija: Gumanitarnye nauki, 5, № 24, pp. 123-143.

CONCLUSIONI

Le conclusioni che traiamo dal nostro lavoro sono di due ordini, conformemente al duplice scopo che ci eravamo prefissi, uno sul versante applicativo - l'ampliamento del corpus parallelo italiano-russo bi-direzionale presente nel NKRJa - e uno sul versante della linguistica contrastiva *corpus-based*, consistente in tre analisi che esemplificassero l'uso del CP ampliato, mostrandone l'utilità e i limiti eventuali in vista di uno sviluppo futuro. Questa duplice prospettiva ci permette dunque di fare alcune riflessioni in primo luogo metodologiche, relative al corpus parallelo come strumento d'indagine, e successivamente specifiche per ciascuna delle analisi contrastive da noi svolte, ossia sui verbi prefissali russi, sul causativo analitico italiano e sulle traduzioni italiane de *Il Cappotto* di N. V. Gogol'.

Per raggiungere il primo obiettivo, dopo un'attenta analisi dei CP esistenti - in particolare dell'ENPC (*English-Norwegian Parallel Corpus*), uno dei corpora paralleli al momento più validi in termini di struttura e bilanciamento - e dopo aver considerato i problemi teorici e applicativi che stanno alla base della loro compilazione, abbiamo stabilito alcuni criteri per la progettazione di un corpus parallelo di qualità (capitolo 1). Li elenchiamo di seguito:

- la presenza di testi russi e italiani, rispettivamente tradotti, in eguali quantità;
- l'appartenenza dei testi a diverse epoche, in modo da consentire studi di tipo diacronico;
- la varietà di autori e traduttori, per garantire rappresentatività e studiare la lingua come fenomeno generale e non particolare di un dato stile o autore.
- l'appartenenza, per quanto riguarda i testi di saggistica, a diversi ambiti del sapere, per garantire una sufficiente varietà terminologica.
- Infine, la presenza di più versioni tradotte di una stessa opera, per consentire la possibilità di condurre analisi comparate anche in questo senso.

Quindi abbiamo raccolto estratti di 10.000-15.000 parole di opere letterarie e di saggistica russe con la rispettiva traduzione italiana e viceversa, così da creare un corpus bi-direzionale bilanciato di almeno 3 milioni di parole (dimensione dell'attuale ENPC) e comprendente testi di vario tipo. Considerata la mole di testi da allineare e poi controllare manualmente, è per il momento disponibile on line alla consultazione (nel sito del NKRJa) solo la sezione dei testi letterari e una piccola parte dei testi di saggistica (cinque

estratti russi con la traduzione italiana); tuttavia, avendo già selezionato altri testi, pensiamo che in futuro si potrà procedere celermente al loro allineamento e inserimento nel NKRJa, soprattutto in vista di un ampliamento del sotto-corpus di saggistica.

Va tuttavia segnalato che non tutti i criteri di progettazione da noi stabiliti sono stati rispettati, dal momento che al nostro lavoro si è aggiunto quello di K. D. Balakina, studentessa di Mosca, la quale, dopo aver preso contatto diretto con i responsabili del NKRJa, ha lavorato in maniera indipendente, fornendo alcuni testi dai lei allineati che sono andati ad aggiungersi alla nostra selezione. Pertanto, al momento i testi russi tradotti in italiano e presenti nel corpus sono in numero maggiore rispetto alle opere italiane tradotte in russo, e alcuni sono disponibili in versione integrale. Ciò ha compromesso il bilanciamento inizialmente auspicato, ma dall'altro lato ha consentito di raggiungere le dimensioni prestabilite in minor tempo e di testare, durante la fase di analisi, l'utilità di avere a disposizione opere integrali anziché estratti. Questo fatto imporrà, in futuro, di compensare gli squilibri per tornare a un corpus bilanciato.

A dicembre 2015, il corpus parallelo italiano-russo ampliato è arrivato dunque a contare più di 4 milioni di parole (precisamente 4.066.172), così distribuite:

- 30 opere letterarie russe con la traduzione italiana (estratti e opere intere);
- 21 estratti di opere letterarie italiane con la traduzione russa;
- 5 estratti di saggistica russa con la traduzione italiana.

Due delle opere letterarie russe dispongono già di diverse versioni italiane: *I racconti di Kolyma* di V. Šalamov (2 versioni) e *Il Cappotto* di N. Gogol' (15 versioni).

Un altro dei nostri obiettivi era quello di rendere il corpus sistematico e comparabile, raccogliendo dunque tutti i dati relativi alla sua progettazione e rendendoli disponibili in un capitolo dedicato di questa tesi (capitolo 2). La disponibilità dei dati riguardanti sia i testi inseriti in un corpus, sia le scelte che ne hanno guidato la compilazione è infatti molto importante per valorizzare i risultati delle ricerche. Si prevede dunque di inserirli anche nel sito del NKRJa così da permettervi il libero accesso.

Infine, intendiamo proseguire l'ampliamento del corpus anche in futuro, pertanto elenchiamo di seguito i nostri obiettivi a breve e lungo termine:

1. Inserire almeno una versione delle opere letterarie degli autori da noi prescelti, ma non ancora presenti nel corpus parallelo: Lermontov (*Un eroe del nostro tempo*), Pasternak (*Il dottor Živago*) e Grossman (*Vita e destino*).

2. Acquisire e allineare ulteriori versioni delle opere già inserite per consentire indagini simili a quelle condotte nel capitolo 5 su *Il Cappotto* di Gogol’.
3. Creare un sotto-corpus di testi in versione integrale.
4. Ampliare la sezione di saggistica, come illustrato nel capitolo 2, allineando i testi già selezionati, appartenenti a diversi ambiti (arte, critica letteraria, filosofia-religione, scienze ecc.).
5. Più a lungo termine, introdurre testi appartenenti ad altre tipologie testuali, come ad esempio articoli di giornale o testi dello stile giuridico-amministrativo (documenti ufficiali delle istituzioni internazionali ecc.). Per questa parte del progetto ricordiamo che è difficile reperire originali e traduzioni in russo e italiano appartenenti a generi letterari diversi da quelli finora considerati. Alcune organizzazioni internazionali, ad esempio l’ONU, dispongono di documenti (perlopiù risoluzioni) tradotti sia in russo che in italiano, ma non è chiaro né se si tratti in entrambi i casi di traduzioni (verosimilmente dall’inglese), né come siano state portate a termine, aspetto fondamentale per conferire al corpus validità scientifica. Con la collaborazione di giornali e riviste (es.: «*Russia beyond the headlines*»), potrebbe essere invece più facile ottenere articoli redatti in una delle due lingue da noi considerate, e successivamente tradotti nell'altra. Questi testi potrebbero costituire una parte di corpus dedicata al linguaggio giornalistico. Infine, in un ulteriore sotto-corpus, sarebbe di grandissimo interesse inserire un corpus pilota degli errori (*Learner translation corpus*), vale a dire un corpus di testi e traduzioni parallele svolte da studenti e da traduttori principianti (Zanettin 2012: 28)²⁴⁸. Un simile corpus potrebbe trovare un buon numero di applicazioni anche negli studi di glottodidattica e di linguistica acquisizionale.

Per quanto concerne le analisi linguistiche *corpus based*, condotte sul nostro CP ed esposte nei capitoli 3, 4 e 5, esse ci permettono di trarre alcune conclusioni riguardo all'utilità di questo strumento che potranno essere tenute in considerazione per un suo futuro sviluppo e miglioramento, e più precisamente:

- **Dimensioni del corpus:** l’indagine condotta nel terzo capitolo sui prefissi verbali ha messo in luce la necessità di disporre di un corpus più voluminoso per poter analizzare l’uso di parole più rare. Infatti, alcuni verbi prefissali, appartenenti al gruppo dei verbi di memoria da noi esaminati, non sono potuti rientrare nella

²⁴⁸ Zanettin F. (2012). *Translation-driven corpora*. Routledge: New York.

nostra analisi (*popomnit'* e *upomnit'*) o non hanno ricevuto un'adeguata descrizione (*zapominat'*) perché non sono stati riscontrati nel nostro CP.

- **Inserimento di estratti vs testi integrali:** abbiamo constatato che per ampliare il corpus è preferibile inserire più brani tratti da varie opere anziché opere intere. Infatti, svolgendo le nostre analisi ci siamo limitati a considerare solo venti-trenta esempi per opera, anche là dove gli esempi per ogni unità linguistica ricercata erano molti di più. Ciò significa che l'elevatissimo numero di esempi, solitamente raggiunto mediante l'inserimento di opere integrali nel corpus, in ogni caso non è stato sfruttato. D'altro canto, avere a disposizione opere integrali può essere utile per rilevare aspetti legati non tanto alla lingua, quanto all'opera in sé e allo stile dell'autore o dei traduttori, tutti elementi che possono poi ricadere direttamente sull'uso di una data struttura linguistica. Nel quinto capitolo, ad esempio, solo avendo a disposizione la versione integrale de *Il Cappotto* abbiamo potuto constatare che la particella *togo* è talvolta impiegata in modo anomalo per caratterizzare il protagonista del racconto.
- **Validità di un corpus bi-direzionale vs uni-direzionale:** la bidirezionalità del corpus consente di disporre allo stesso tempo di un corpus parallelo e comparabile, ossia di un corpus che permette di confrontare non solo originali in lingua A e rispettive traduzioni in lingua B, ma anche originali in lingua A e traduzioni in lingua A (così come originali in lingua B e traduzioni in lingua B), per identificare le peculiarità e le eventuali differenze tra lingua originale e lingua della traduzione. Questa potenzialità è stata testata in particolare nel capitolo 4, dove la ricerca della costruzione causativa “*fare + infinito*” non solo in originali italiani, ma anche in traduzioni italiane di brani russi, ha permesso di confermare l'ipotesi di una funzione specifica di *fare* causativo – quella stilistica – confermando che l'uso di espressioni quali “*far vedere*” rispetto a “*mostrare*” si dà sia nei testi originali italiani, sia nelle traduzioni italiane di testi russi, dove nei dialoghi informali il russo “*pokazat'*” era tradotto sempre con “*far vedere*” e non con “*mostrare*”.
- **Corpus compilati in modo tradizionale vs web corpora estratti automaticamente:** possiamo confermare che un corpus tradizionale, composto da testi selezionati secondo precisi criteri e accompagnati da informazioni extra-linguistiche accurate, consente di valutare in modo più preciso i risultati ottenuti.

Ad esempio, grazie ai dati extralinguistici riguardanti le opere inserite, è stato possibile documentare che l'uso del verbo *pripomnit'* come sinonimo di *vspomnit'* è tipico del XIX secolo e in disuso in testi più recenti, mentre nel quinto capitolo è emerso che alcune strategie di resa della particella *togo* e delle frasi ellittiche sono legate all'economia del racconto in cui compaiono (*Il Cappotto*), che ha condizionato quindi, in molti casi, le scelte dei traduttori.

- **Annotazione:** la disponibilità di un corpus annotato grammaticalmente e semanticamente si è rivelata indispensabile per interrogare il motore di ricerca e ottenere gli esempi utili all'analisi. Tuttavia ci siamo resi conto, in particolare nel quinto capitolo, che la disambiguazione è un parametro altrettanto indispensabile per analizzare entità linguistiche altamente polisemiche o omonimiche, come nel caso della particella russa *togo*. Tra i nostri obiettivi futuri, dunque, rientrerà l'impegno di disambiguare almeno parte del corpus per poter ottenere dall'interrogazione risultati più precisi.

Nel nostro lavoro abbiamo anche approfondito alcuni temi linguistici in chiave contrastiva.

Nel terzo capitolo, abbiamo studiato mediante il CP la resa in italiano dei verbi prefissali di memoria: *vspomnit'-vspominat'*, *zapomnit'-zapominat'*, *pripomnit'-pripominat'*, con l'intento di chiarire le differenze semantiche di queste unità linguistiche, spesso trascurate dai dizionari sia monolingui sia bilingui. Abbiamo così visto che la complessità dei verbi prefissali russi risiede nell'includere in un'unica unità linguistica (il verbo) molti tratti semantici che, in lingue tipologicamente diverse quali l'italiano, non possono essere espressi da un unico verbo. La ricca varietà dei verbi di memoria russi in italiano viene così per lo più ridotta all'unico equivalente traduttivo *ricordare*. Abbiamo dunque approfondito la semantica di questo gruppo di verbi russi e verificato se nelle traduzioni in italiano dei verbi di memoria si verifici una perdita di informazione e, quando questo non accade, se esistano strategie di resa specifiche.

Per quanto riguarda il primo obiettivo – precisare la semantica dei verbi prefissali di memoria – l'indagine *corpus-based* è stata utile per identificare un significato specifico di *vspomnit'* e *pripomnit'*, quello conativo, che indica l'atto volontario di richiamare qualcosa alla memoria. Questo tratto, perlopiù ignorato dai dizionari, è in realtà molto frequente nell'uso e evidente anche in un corpus di soli 3 milioni di parole. Ciò testimonia

che i corpora, grazie ai numerosi esempi contestualizzati, possono colmare le lacune lessicografiche che sorgono talvolta nei dizionari compilati con metodi tradizionali e sulla base dell'intuizione linguistica del compilatore²⁴⁹. In aggiunta, abbiamo indentificato una corrispondenza tra le classi verbali vendleriane e la prefissazione verbale in russo: abbiamo infatti constatato che in russo la differenza tra “ricordare come azione volontaria”, “ricordare come stato involontario” e “ricordare come culminazione” è marcata in parte dall'aspetto, ma in parte anche dai prefissi. In italiano, al contrario, il verbo ricordare è più spesso usato come verbo stativo indicante un'azione involontaria. La volontarietà del ricordo può essere invece espressa grazie a espressioni aggiuntive con valore conativo (cfr.: “cercare di”), a sinonimi più specifici (cfr.: “pensarci”, “memorizzare” ecc.) e, in rari casi, all'uso dei tempi verbali (presente vs gerundio).

Infine, riguardo alla resa in italiano, abbiamo trovato conferma del fatto che l'uso di *ricordare* per tradurre i verbi di memoria russi comporta spesso una perdita di informazione se non è accompagnato da ulteriori elementi che permettano di rendere per intero la semantica dell'originale russo. È possibile ovviare a questa perdita grazie ad alcune strategie che abbiamo di volta in volta descritto e classificato.

L'analisi del quarto capitolo aveva lo scopo di studiare la costruzione causativa italiana *fare* + infinito in contrapposizione al russo. A differenza di quanto avvenuto per l'indagine precedente, in questo caso non siamo stati mossi da alcuna ipotesi preliminare. Pertanto questa seconda ricerca adotta un approccio più simile a quello *corpus-driven*.

Notiamo innanzitutto che, dal punto di vista teorico, abbiamo riscontrato una differenza sostanziale nel trattamento dei verbi causativi nelle due lingue: se negli studi di italianistica il causativo è studiato prevalentemente dal punto di vista sintattico e quasi esclusivamente in riferimento alla costruzione *fare* + infinito, nella russistica i numerosi verbi causativi vengono classificati in base alla loro semantica.

Analizzando poi gli esempi originali e le loro traduzioni è stato possibile approfondire sia la semantica e le funzioni di *fare* + infinito, sia le differenze tra russo e italiano nell'esprimere una situazione causativa. Oltre ad essere espressione di causativo, la costruzione analitica con *fare* funge spesso da marca stilistica (sostituisce verbi usati in contesti più formali: *far vedere* vs *mostrare*) o sintattica (consente all'infinito di un

²⁴⁹ L'attenzione verso questo tipo di problemi è già diffusa nella LC e la sfida dei linguisti specializzati nell'uso di corpora è di compilare, in futuro, dizionari che facciano uso esclusivamente dei corpora per le definizioni e gli esempi delle voci.

verbo intransitivo di reggere un complemento oggetto diretto: “*Luigi ha fatto scendere Giulia dall’auto*”). L’uso di *fare* + infinito con valore non causativo diventa immediatamente evidente nella traduzione russa, che in questi casi rende la costruzione italiana con verbi o espressioni non causative (cfr.: “*come ci ha fatto comprendere Berengario*” tradotto come “*По показаниям Берензара*”). Dal nostro corpus emerge anche che in italiano statisticamente prevale l’uso di *fare* + infinito, nonostante l’esistenza di molti verbi sinonimici con semantica più specifica (es.: *costringere, ordinare* ecc.), il che probabilmente è favorito anche dal fatto che *fare* è l’unico verbo con significato causativo che può essere seguito dall’infinito semplice senza preposizione.

Infine è emersa una differenza nell’espressione degli attanti sintattici: se la costruzione *fare* + infinito esprime due soggetti – quello del causatore e quello del causato -, il russo presenta più spesso un solo soggetto, creando così un’ambiguità, dal momento che “*korol’ ich zapereł*” può significare allo stesso tempo: “*il re li ha rinchiusi*” (causativo diretto) e “*il re li ha fatti rinchiusere*” (causativo indiretto). Questo fenomeno non si verifica solo con i verbi causativi, ma con molti altri verbi, anche intransitivi (es: *lečit’sja*). Ciò conferma la tesi di R. Govoruchò per cui il russo “a livello di gerarchia sintattica risulta essere meno coeso e più implicito” rispetto all’italiano (Govoruchò 2010: 187-88)²⁵⁰. L’autore qui si riferisce soprattutto all’aggiunta nel testo italiano di verbi proposizionali di percezione o di altri verbi semanticamente simili²⁵¹. Es: *Эраст Петрович подождет Ø, не последует ли выстрел... (Акунин)* in contrapposizione a *Erast Petrovič per un po’ attese per vedere se sarebbe seguito lo sparo...* (esempi di Govoruchò in 2010: 199); tuttavia, dopo l’analisi dei nostri esempi, si può affermare che questa tendenza interessi anche altri elementi.

Nel quinto capitolo, infine, abbiamo utilizzato il CP per un’indagine comparata di diverse traduzioni di un’unica opera. Anche in questo caso possiamo trarre alcune conclusioni preliminari non solo riguardo agli specifici temi di indagine (la particella *togo* e l’ellissi verbale), ma anche sull’uso dei corpora paralleli.

²⁵⁰ Govoruchò R. A. (2010b). *Qualche osservazione sull’uso dei verbi proposizionali in italiano e in russo*. In “L’Analisi Linguistica e Letteraria”, Anno XVIII, 2/2010, Facoltà di Scienze Linguistiche e Letterature Straniere, Università Cattolica del Sacro Cuore, Milano, pp. 187-200.

²⁵¹ “Nel testo italiano ricevono uno status grammaticale più ‘prestigioso’ quegli elementi del modus di enunciazione che si realizzano nella forma massimamente esplicita. Si tratta soprattutto di verbi proposizionali con semantica percettiva ed epistemica, nonché di verbi del dire ad essi vicini. Nel testo russo questi elementi del significato possono ridursi in parte o completamente e andare a far parte di altri elementi” (Govoruchò 2010: 199).

Abbiamo evidenziato, in primo luogo, il carattere russo-specifico sia della particella *togo*, priva di un esatto equivalente in italiano, sia dell'ellissi del verbo, inammissibile nella nostra lingua con la frequenza in cui si realizza in russo. In particolare riguardo a *togo*, l'assenza di un esatto equivalente è ancora più evidente quando le diverse traduzioni si riferiscono al medesimo testo fonte e l'elemento da tradurre è dunque inserito nello stesso contesto. Inoltre abbiamo notato che il modo di tradurre *togo* nelle versioni italiane da noi studiate non sempre è motivato dalla funzione discorsiva che esso assume nel contesto; più verosimilmente, i traduttori che hanno reso *togo* con forme quali *in quanto*, *coso*, *vero e coso* si sono ispirati alle indicazioni della voce narrante, secondo la quale Akakij Akakievič si esprimerebbe con particelle prive di senso, tra cui emerge in modo particolarmente insistente *togo*. Lo stesso può essere detto anche dell'ellissi: alcuni traduttori, seppur in numero minore, hanno ripetuto l'ellissi del verbo anche nella traduzione italiana, realizzando così un testo sconnesso che, tuttavia, non ha equivalente nell'originale russo.

La nostra analisi di diverse traduzioni di un testo letterario ci permette anche di completare le osservazioni sull'uso dei corpora paralleli per l'indagine linguistica. Infatti, se è vero che la lingua tradotta può essere considerata autentica e quindi adatta allo studio delle diverse strutture linguistiche, soprattutto in un corpus parallelo bi-direzionale dove l'indagine può essere condotta in entrambe le direzioni, è parimenti evidente che un testo letterario può utilizzare gli strumenti linguistici non secondo l'uso standard, come nel caso delle particelle, che nel *Cappotto* di Gogol' servono talvolta a creare non-senso, anziché svolgere la loro funzione ordinaria. Pertanto, tra le caratteristiche di un corpus di qualità sottolineiamo ancora una volta l'importanza di mettere a disposizione dei ricercatori i metadati necessari per risalire a informazioni extralinguistiche quali l'autore del testo, il titolo dell'opera, il periodo di composizione e tutti gli elementi che, come abbiamo visto, possono avere rilevanza anche nell'analisi linguistica.

Infine, lo studio delle particelle ha evidenziato la necessità di migliorare l'annotazione del nostro CP disambiguando l'omonimia, mentre quello dell'ellissi sintattica ha fatto emergere l'esigenza di mettere a punto sistemi di interrogazione più efficaci per estrarne degli esempi dal corpus.