

Università degli Studi di Trieste  
Dipartimento di Scienze Giuridiche, del Linguaggio,  
dell'Interpretazione e della Traduzione

# Rivista Internazionale di Tecnica della Traduzione

*International Journal  
of Translation*

EUT

n. 23/2021



*Rivista Internazionale di Tecnica della Traduzione (RITT)*  
*The Rivista Internazionale di Tecnica della Traduzione – International Journal of Translation of the IUSLIT Department and Scuola Superiore di Lingue Moderne per Interpreti e Traduttori, University of Trieste aims at providing a forum of discussion for the multifaceted activity of translation and related issues.*

*Scientific Committee*

Ovidi Carbonell i Cortés (Universidad de Salamanca)  
Marco A. Fiola (Ryerson University, Ontario)  
David Katan (Università del Salento)  
Heidi Salaets (Katholieke Universiteit Leuven)  
Michael Schreiber (Johannes Gutenberg-Universität Mainz)

*Editorial Committee*

Nadine Celotti  
Paola Gentile  
Pascale Janot  
Marella Magris  
José Francisco Medina Montero  
Stefano Ondelli  
Giuseppe Palumbo  
Manuela Raccanello  
Lorenza Rega  
Dolores Ross  
Federica Scarpa

*Editor*

José Francisco Medina Montero

*Co-Editor*

Floriana Sciumbata

© Copyright Edizioni Università di Trieste  
Trieste 2021

Proprietà letteraria riservata.

I diritti di traduzione, memorizzazione elettronica, di riproduzione e di adattamento totale e parziale di questa pubblicazione, con qualsiasi mezzo (compresi i microfilm, le fotocopie e altro) sono riservati per tutti i paesi

ISSN 1722-5906 (print)  
ISSN 2421-6763 (online)

EUT - Edizioni Università di Trieste  
via Weiss 21 – 34128 Trieste  
<http://eut.units.it>  
<https://www.facebook.com/EUTEdizioniUniversitaTrieste>

Rivista  
Internazionale di  
Tecnica della  
Traduzione

*International Journal of Translation*

Numero 23  
2021



# SOMMARIO

## TABLE OF CONTENTS

7	<i>José Francisco Medina Montero</i> Prefazione/Editorial Preface	91	<i>Rocío Luque</i> El prefijo s- en italiano y su exhaustiva correspondencia en la traducción al español
15	SEZIONE/SECTION 1 MISCELLANEA		
17	<i>Paolo Canavese</i> <i>Goranka Rocco</i> Versprachlichung der Rolle jüngerer Bevölkerungsschichten in Pandemiediskursen: syntaktische Funktionen, semantische Rollen und deontische Aspekte	105	<i>Rosaria Minervini</i> La traducción automática del género (español-italiano): análisis de ejemplos traducidos con DeepL y Google Traductor
37	<i>Antonella De Laurentiis</i> <i>Metegol, Fútbolín e Gool!:</i> tre versioni a confronto	129	<i>Giulia Pedrini</i> The “vaccine race”: Metaphorical conceptualizations of the search of an immunization against Covid-19
55	<i>Verónica Del Valle Cacela</i> Reflexiones sobre el error en la traducción jurídica	153	<i>Giuseppe Trovato</i> Hacia una sistematización de las disimetrías morfosintácticas en la traducción del régimen preposicional entre lenguas tipológicamente emparentadas: español/italiano
65	<i>Giovanna Ferrara</i> Turismo sostenible y culturemas en español e italiano		
79	<i>Giacomo Klein</i> Studenti di traduzione vs. studenti di lingua: chi traduce meglio? Uno studio greco-italiano	177	SEZIONE/SECTION 2 ESPERIMENTI DI DISTANT READING. ESTRAZIONE, ANALISI E VISUALIZZAZIONE DI DATI LINGUISTICI DA CORPORA LETTERARI

- 179 *Michele A. Cortelazzo*  
Corpora e storia della lingua
- 187 *Stefano Ondelli*  
Tempi verbali e perifrasi gerundivali in un corpus di italiano letterario (1800-2000)
- 213 *Floriana Carlotta Sciumbata*  
Premessa agli articoli di Sciandra, Trevisani e Tuzzi e di Sciumbata, Nadalutti e Tringali
- 219 *Andrea Sciandra*  
*Matilde Trevisani*  
*Arjuna Tuzzi*  
Sulle tracce dell'espressione dell'interiorità: analisi diacronica di un corpus di narrativa italiana del XIX-XX secolo
- 235 *Floriana Carlotta Sciumbata*  
*Paolo Nadalutti*  
*Luca Tringali*  
"Trovare lavoro" in un corpus di narrativa del XIX-XX secolo. Procedure, aspetti e problemi di creazione, estrazione e rappresentazione dei dati

# Prefazione

## Editorial Preface

JOSÉ FRANCISCO MEDINA MONTERO  
Università di Trieste

Nel ventitreesimo volume della *Rivista Internazionale di Tecnica della Traduzione (RITT)* presenteremo i contributi di diciotto ricercatori. Le università rappresentate sono italiane (Trieste, Udine, Ca' Foscari Venezia, Padova, Salerno, Basilicata, Modena/Reggio Emilia e Università del Salento) e straniere (l'Uniwersytet Marii Curie-Skłodowskiej, Polonia, e l'Università di Ginevra, Svizzera). I tredici articoli di questo numero sono stati redatti in quattro lingue (italiano, spagnolo, inglese e tedesco), il che aiuta ad aumentare il consueto taglio internazionale della rivista.

Gli argomenti sui quali si sono concentrati riguardano temi di linguistica contrastiva, testuale e di linguistica dei corpora, di formazione delle parole, di traduzione audiovisiva, automatica e giuridica, di lessico, di didattica della lingua e della traduzione e di analisi automatica dei dati testuali.

Abbiamo diviso il volume in due Sezioni, una miscellanea, che include nove saggi, e una monografica intitolata "Esperimenti di *Distant Reading*. Estrazione, analisi e visualizzazione di dati linguistici da corpora letterari", che ne comprende altri quattro. In ogni Sezione i lavori appariranno seguendo l'ordine alfabetico degli autori.

Nella prima Sezione troveremo i saggi di Paolo Canavese e Goranka Rocco, di Antonella De Laurentiis, Verónica Del Valle Cacula, Giovanna Ferrara, Giacomo Klein, Rocío Luque, Rosaria Minervini, Giulia Pedrini e Giuseppe Trovato. Di seguito riassumeremo con grande concisione di che cosa tratta ognuno di loro.



Lo studio di Paolo Canavese e Goranka Rocco, “Versprachlichung der Rolle jüngerer Bevölkerungsschichten in Pandemiediskursen: syntaktische Funktionen, semantische Rollen und deontische Aspekte”, svolto nell’ambito del più ampio progetto *COVIDisc – A focus on vector populations*, ha l’obiettivo di analizzare il discorso mediatico nella Svizzera tedesca e italiana sul ruolo dei giovani nella pandemia da Covid-19. Dopo l’identificazione dei principali mezzi lessicali utilizzati per riferirsi ai giovani, i due autori esaminano le funzioni sintattiche e i ruoli semantici in cui figurano, nonché gli aspetti deontici del discorso. Un’analisi cronologica permette inoltre di tracciare la costruzione di un *common ground* circa il ruolo dei giovani nel contesto pandemico e, in particolar modo, di investigare gli scenari deontici riguardanti i comportamenti da adottare o evitare, l’effettivo rispetto delle prescrizioni delle autorità e la responsabilità nella sensibilizzazione degli adolescenti. I dati riportati e illustrati con una serie di esempi costituiscono la base per futuri approfondimenti di taglio quantitativo, essenziali per affinare l’interpretazione dei trend emersi.

Il lavoro di Antonella De Laurentiis, “*Metegol, Fútbolín e Goooll!*: tre versioni a confronto”, prende come punto di riferimento la pellicola *Metegol* (2013), il primo film d’animazione girato in 3D in Argentina, diretto dal regista e premio oscar Juan José Campanella e liberamente ispirato al racconto dello scrittore argentino Roberto Fontanarrosa, *Memorias de un wing derecho*. De Laurentiis presenta la versione originale del film e la confronta con le versioni doppiate in italiano e nella varietà dello spagnolo peninsulare. L’autrice analizza alcune scene tratte dal film che mostrano tratti marcati della varietà argentina. In modo particolare si sofferma sulle locuzioni idiomatiche e sul lessico del *lunfardo*, con l’obiettivo di mostrare in che modo questi elementi siano stati tradotti, adattati o neutralizzati nelle due versioni doppiate. Inoltre, dimostra in che modo il destinatario influisce sulle scelte traduttive e quali sono le caratteristiche proprie del doppiaggio di prodotti audiovisivi destinati a un pubblico infantile.

Nel terzo articolo, “*Reflexiones sobre el error en la traducción jurídica*”, Verónica Del Valle Cacula riflette sugli errori commessi nel momento in cui lo studente deve affrontare per la prima volta il linguaggio giuridico, che determina una serie di difficoltà dovute anche al contesto in cui viene prodotto. L’autrice dedica la prima parte del contributo a un excursus dedicato alla definizione di linguaggio giuridico, dopodiché affronta le varie abilità che un traduttore deve acquisire cioè, comprendere il testo di partenza e produrre un testo scritto nella lingua di arrivo. Di seguito, Del Valle Cacula sottolinea il peso dell’errore e il suo valore didattico quando si impara a tradurre e segnala brevemente il ruolo che potrebbe avere nel momento in cui si lavora con un’altra lingua affine a quella italiana, ad esempio, con lo spagnolo. Infine, presenta le conclusioni con l’obiettivo di poter sviluppare più avanti delle ricerche pratiche sull’argomento.

Per Giovanna Ferrara, “*Turismo sostenible y culturemas en español e italiano*”, il turismo nasce dall’esigenza umana di esplorare, conoscere luoghi, persone e culture. Il linguaggio turistico è eterogeneo e si trova sulla linea di confine

tra altri linguaggi settoriali. Il suo lavoro si basa sull'analisi di alcuni aspetti nel campo della traduzione e della lessicografia del turismo sostenibile. Attraverso un approccio qualitativo, si sofferma su alcune caratteristiche linguistiche e traduttologiche da un punto di vista contrastivo (italiano/spagnolo), basandosi sul corpus Cometval, che contiene testi del portale ufficiale del turismo spagnolo (<https://www.spain.info/es/>), oltre che su diversi dizionari monolingui (il *Diccionario de la lengua española* della Real Academia Española e il *Clave*) e bilingui (Laura Tam). Ferrara evidenzia alcuni *realia* che possono presentare difficoltà nel processo di traduzione dallo spagnolo all'italiano (*alfajores, monas de Pascua* ecc.) e dall'italiano allo spagnolo (come per esempio *peperoni cruschi, tartellate* ecc.). L'obiettivo è quello di stabilire alcuni strumenti utili nel campo professionale della mediazione linguistica e culturale attraverso la traduzione. Per l'autrice, il mediatore linguistico e culturale trasmette conoscenze interculturali per facilitare lo scambio linguistico e per evitare fraintendimenti nella comunicazione. Non si tratta di una semplice traduzione, ma di trasmettere saperi unici e propri di un gruppo determinato e che spesso non hanno un equivalente concettuale e linguistico nella lingua d'arrivo. Per questo, il mediatore linguistico e culturale mette in pratica una serie di strategie e tecniche traduttologiche, col fine di raggiungere un grado di equivalenza quanto più fedele possibile.

L'articolo di Giacomo Klein, "Studenti di traduzione vs. studenti di lingua: chi traduce meglio? Uno studio greco-italiano", mira a confrontare i risultati di un esercizio traduttivo greco-italiano tra studenti di traduzione e studenti di greco L2. Il suo studio si concentra su un'analisi su diversi livelli degli errori nei testi di arrivo. Agli studenti della Sezione di Studi di Lingue Moderne per Interpreti e Traduttori dell'Università di Trieste e agli studenti che frequentano i corsi di greco della Fondazione Ellenica per la Cultura di Trieste è stato chiesto di tradurre un breve testo informativo utilizzando sia dizionari cartacei che online. I testi tradotti sono stati poi corretti e gli errori di senso, morfologia, lessico e sintassi sono stati individuati e confrontati utilizzando sia un approccio quantitativo che qualitativo. I risultati suggeriscono che gli studenti di traduzione hanno prodotto traduzioni più accurate, soprattutto per quanto riguarda la semantica e il lessico.

Nel suo elaborato, "El prefijo *s-* en italiano y su exhaustiva correspondencia en la traducción al español", Rocío Luque analizza l'uso del prefisso italiano *s-* con basi verbali, dal punto di vista morfosintattico e semantico, con l'obiettivo di conoscere le principali equivalenze traduttive in spagnolo nelle varie formazioni. Essendo due lingue latine con differente sviluppo cognitivo nel tempo, l'autrice riscontra la presenza di numerosi falsi amici con la stessa etimologia ma con degli usi pragmatici asimmetrici, oltre che diversi aspetti contrastanti, che vanno dalla loro origine all'uso attuale della lingua. Pur essendoci, difatti, dei principi morfologici che in qualche modo facilitano alcune equivalenze (per esempio, le corrispondenze con le forme spagnole *ex-* o *des-*), sono numerosi i casi in cui, a causa dell'elevata funzionalità del prefisso in italiano, in spagno-

lo bisogna ricorrere ad altre soluzioni, come l'assenza di questo tipo di affisso, l'uso di un altro o il ricorso alla parafrasi. Inoltre, Luque ricorda che il prefisso *s-* in italiano non solo è in grado di produrre dei cambiamenti semantici, apportando come valori principali quelli di intensità e di privazione, e dei cambiamenti morfologici (ricordiamo che, a differenza dallo spagnolo, dà origine a formazioni verbali parasintetiche a partire da nomi o sostantivi), ma possiede anche la caratteristica di modificare la sintassi delle accezioni di alcuni verbi, trasformandoli da transitivi in intransitivi e viceversa, rendendo ancora più interessanti i processi traduttologici.

Per Rosaria Minervini, "La traducción automática del género (español-italiano): análisis de ejemplos traducidos con DeepL y Google Traductor", in un mondo sempre più globalizzato, dove la necessità di comunicare e di avere scambi di tipo commerciale è sempre più impellente, la traduzione automatica ha acquisito un ruolo sempre più importante in quanto permette di tradurre una grande quantità di parole e di testi ad una velocità elevata e ad un costo esiguo. Per questi motivi, viene usata sempre più spesso sia dagli esperti della traduzione (in ambito professionale e/o accademico) che dalle persone non specializzate, che hanno bisogno di capire e creare testi in lingua straniera. La questione del genere nella lingua è ancora aperta e molto attuale e, sebbene sia stata discussa ampiamente nella teoria della traduzione, ha bisogno di essere approfondita ulteriormente nel campo della traduzione automatica. In questo articolo, Minervini descrive i principali sistemi di traduzione automatica ed esamina alcuni studi che si sono occupati dei problemi relazionati alla traduzione automatica del genere. Dopo questa parte teorica, attraverso alcuni esempi derivanti da discorsi politici e giornalistici, l'autrice osserva come due traduttori automatici disponibili oggi sul mercato, DeepL e Google Translator, traducono il genere nella combinazione linguistica spagnolo-italiano. Il motivo per cui ha scelto questi due sistemi è che sono tra i più utilizzati e gratuiti, con interfacce semplici, che rendono il loro utilizzo alla portata di tutti. Infine, Minervini parla delle differenze nel risultato dei due sistemi.

Nel suo articolo, "The 'vaccine race': Metaphorical conceptualizations of the search of an immunization against Covid-19", Giulia Pedrini presenta un'analisi multilingue (in italiano, inglese, tedesco e russo) delle metafore inerenti alla ricerca del vaccino per il coronavirus sulla base della Teoria della Metafora Concettuale di Lakoff e Johnson. Le metafore, infatti, illustrano concetti astratti tramite esemplificazioni concrete: perciò rappresentano uno strumento fondamentale nella comunicazione sanitaria in generale e nel discorso attorno al Coronavirus e i vaccini in particolare. Il corpus analizzato è costituito prevalentemente da articoli di giornale, ma anche da contenuti di canali televisivi e podcast rinvenibili online, mentre il periodo considerato va dall'inizio della pandemia fino a gennaio 2021. I domini sorgenti delle concettualizzazioni metaforiche individuate sono prevalentemente i seguenti: corsa, corsa a cavallo, gare automobilistiche, caccia, giochi di strategia, gioco d'azzardo, viaggio, guerra, mitologia e folclore,

nonché eventi storici. L'autrice esamina i contesti d'uso e le diverse sfumature che caratterizzano i ricorsi alle varie metafore, anche in prospettiva contrastiva tra le lingue oggetto d'analisi. Pedrini evidenzia alcune criticità relative all'uso di determinate concettualizzazioni, situabili al crocevia tra studi sulle metafore, terminologia e traduttologia: a puro titolo d'esempio, l'uso metaforico di termini tecnici che potrebbero dare adito ad ambiguità, o il ricorso a calchi di metafore utilizzate prevalentemente in un'altra lingua e non idiomatiche e immediate nella lingua d'arrivo.

Il contributo di Giuseppe Trovato, “Hacia una sistematización de las disimetrías morfosintácticas en la traducción del régimen preposicional entre lenguas tipológicamente emparentadas: español/italiano”, ha come obiettivo quello di analizzare la traduzione del regime preposizionale tra due lingue affini (spagnolo/italiano) da una prospettiva eminentemente contrastiva e traduttologica. È indiscutibile che la conoscenza profonda dell'uso delle preposizioni in una lingua straniera è un fatto che dimostra e corrobora un livello avanzato di *proficiency* linguistica. Tuttavia, si assiste spesso a fenomeni di incorrezione e incertezze nell'uso delle preposizioni. Mediante la traduzione di tale fenomeno linguistico è possibile avviare una riflessione sui contrasti esistenti tra l'italiano e lo spagnolo, dato che vi è una tendenza generalizzata a ritenere che il trasferimento linguistico tra queste due lingue avvenga in maniera naturale, immediata e trasparente. Per il presente studio, l'autore ha creato un corpus a partire dalla traduzione in italiano del romanzo *Caperucita en Manhattan*, della famosa scrittrice spagnola Carmen Martín Gaité. Data la vastità del fenomeno in questione, Trovato limita la sua analisi contrastiva e traduttologica a tre preposizioni spagnole molto comuni: *a*, *de*, *en*. Il corpus creato *ad hoc* consente di mettere in luce la presenza di numerose dissimmetrie morfosintattiche nell'ambito del processo traduttivo, il che dimostra in maniera abbastanza chiara che tradurre tra due lingue affini non rappresenta in assoluto un'attività priva di ostacoli.

Nella seconda Sezione introdurremo i quattro ultimi lavori cioè, gli elaborati di Michele A. Cortelazzo, Stefano Ondelli, di Andrea Sciandra, Matilde Trevisani e Arjuna Tuzzi e di Floriana Carlotta Sciumbata, Paolo Nadalutti e Luca Tringali.

La base di questa Sezione monografica è il progetto intitolato *Analisi automatica di un corpus multilingue di romanzi*, finanziato dal Fondo per la Ricerca di Ateneo (FRA) dell'Università di Trieste e coordinato da Stefano Ondelli. Il progetto prende le mosse dal COST-Action *Distant Reading for European Literary History* (COST Action CA16204: <https://www.distant-reading.net/>), sfruttando il corpus *European Literary Text Collection* (ELTeC: <https://distantreading.github.io/ELTeC/>) annotato in TEI (*Text Encoding Initiative*: <https://tei-c.org/>) in via di compilazione, che presto includerà 100 romanzi per ciascuna delle 13 lingue/letterature nazionali coinvolte.

Il progetto, teso a sviluppare e affinare software già disponibili sul mercato e in via di sviluppo, utili alla pulizia e preparazione dei testi e all'analisi auto-

matica e statistica dei dati linguistici (per esempio *POS-Tagger*, lemmatizzatori, analizzatori di dipendenze sintattiche multilingue ecc.), ha condotto analisi quali-quantitative condotte nell'ottica del *Distant Reading* (cfr. Franco Moretti, *La letteratura vista da lontano*, Torino, Einaudi, 2005) basate su strategie di *text mining*, *information retrieval* e *sentiment analysis* (cfr. Sergio Bolasco, *L'analisi automatica dei testi*, Roma, Carocci, 2013), ma comprendendo anche tentativi di indagine da una prospettiva più strettamente linguistica (per esempio modalità di discorso riportato, distribuzione dei tempi verbali, frequenza e distribuzione delle marche morfologiche di persona ecc.).

Il primo articolo che riassumeremo, “*Corpora e storia della lingua*”, appartiene a Michele A. Cortelazzo. Lui sostiene che nell'ultimo ventennio lo studio della storia della lingua italiana, soprattutto per quel che riguarda il lessico, ha tratto grande beneficio dall'esistenza di corpora letterari diacronici, come la “*Letteratura Italiana Zanichelli*” o la “*Biblioteca Italiana*” (e, in prospettiva, la sezione dell'ELTeC - European Literary Text Collection). Tuttavia, questi corpora sono fondamentalmente letterari, scelti in base ai dettami del canone letterario dominante e comprendono un numero esiguo di testi rispetto all'arco temporale che dovrebbero coprire. Questi limiti risultano particolarmente rilevanti per gli studi sull'evoluzione diacronica della lingua. Oggi è disponibile online un gran numero di testi (a cominciare da quelli raccolti in Google Libri), che coprono un ampio arco di tempo e un'ampia varietà di generi. L'analisi di tali grandi corpora richiede strumenti e approcci adeguati per garantirne un'efficace lettura a distanza. In questa prospettiva, la collaborazione interdisciplinare di linguisti, statistici e studiosi di letteratura può portare una prospettiva innovativa all'elaborazione dei dati contenuti nei corpora, allo scopo di valutare l'evoluzione dell'italiano letterario.

Stefano Ondelli sottolinea in “*Tempi verbali e perifrasi gerundivali in un corpus di italiano letterario (1800-2000)*” che tra le perifrasi verbali italiane, quelle gerundivali hanno avuto uno sviluppo reciprocamente inverso nell'ultimo secolo. Se la situazione fino al XIX secolo vede il predominio numerico della costruzione continua (*andare* o *venire* + gerundio), nel corso del XX secolo la costruzione progressiva (*stare* + gerundio) fa registrare un deciso sorpasso. Inoltre, alcune delle restrizioni a cui era sottoposta vengono meno e si amplia la gamma di tempi verbali e classi azionali disponibili. Sulla scorta di ricerche analoghe di tipo qualitativo e quantitativo svolte in precedenza, il suo contributo analizza un corpus di opere di romanzieri italiani e di traduzioni di romanzi stranieri relative ai periodi 1811-1896 e 1976-2005. Accanto ai riscontri quantitativi assoluti, che confermano la netta preminenza della perifrasi continua nella lingua letteraria alla fine del Novecento, i dati indicano che i traduttori hanno almeno parzialmente anticipato gli sviluppi delle perifrasi gerundivali nella lingua italiana nella selezione dei verbi modificatori e dei tempi verbali.

L'elaborato di Andrea Sciandra, Matilde Trevisani e Arjuna Tuzzi, “*Sulle tracce dell'espressione dell'interiorità: analisi diacronica di un corpus di narrativa*”

italiana del XIX-XX secolo”, illustra un esempio di analisi statistica di un corpus diacronico di grandi dimensioni, composto da 100 opere di narrativa italiana, con l’obiettivo di riconoscere tendenze prototipiche nella frequenza delle parole e identificare gruppi di parole con andamento temporale simile. A partire dalle indicazioni degli studiosi di letteratura, il lavoro prova a saggiare l’ipotesi, spesso avanzata nella storia della letteratura europea, secondo cui la vita interiore degli autori e/o dei loro personaggi abbia assunto maggiore importanza nel corso del XIX secolo e sia diventata uno dei capisaldi del modernismo. Sebbene non si possa rigettare l’ipotesi, i metodi statistici impiegati non evidenziano andamenti temporali netti: emergono alcune timide tendenze, ma molte parole mostrano a livello individuale un comportamento indifferenziato nel corso del tempo. L’idea di studiare con prospettiva diacronica la presenza nelle opere dell’espressione interiore degli autori è sicuramente molto ambiziosa. Questo primo esperimento mostra alcune potenzialità degli strumenti statistici impiegati ma anche alcuni limiti dei criteri adottati sia per la costituzione del corpus sia per la scelta dei fenomeni linguistici da osservare.

Per concludere, l’articolo di Floriana Carlotta Sciumbata, Paolo Nadalutti e Luca Tringali, “‘Trovare lavoro’ in un corpus di narrativa del XIX-XX secolo. Procedure, aspetti e problemi di creazione, estrazione e rappresentazione dei dati”, presenta metodi, strumenti e risultati preliminari di uno studio volto a creare una lista di nomi di mestieri e professioni finalizzata all’analisi di 100 opere di narrativa pubblicate tra il 1825 e il 1923. I nomi di mestieri e professioni costituiscono un elemento interessante perché riflettono i cambiamenti socio-economici occorsi nel corso del tempo, oltre a fornire importanti informazioni su come sono cambiati gli ambienti e i generi letterari. Dopo una breve introduzione sui nomi di mestieri e professioni dal punto di vista linguistico, gli autori presentano gli strumenti e i metodi utilizzati per la creazione di una lista di lemmi, l’estrazione dei dati e la loro rappresentazione, sia da un punto di vista linguistico sia da quello informatico. I risultati preliminari vengono mostrati e discussi con un approccio statistico: i dati non dimostrano particolari cambiamenti nel tempo e suggeriscono invece che molti mestieri sono rimasti costanti. Infine, i tre esaminano alcuni vantaggi e limiti del loro approccio. L’obiettivo del loro studio è quello di sviluppare un insieme di strumenti e metodi che possano essere facilmente riprodotti per costruire liste lessicali complesse, trovare i loro elementi e rappresentare i dati in corpora di qualsiasi genere e dimensione in modo semplice ed efficiente.

E senza perdere altro tempo Vi auguriamo, come sempre, buona lettura.



Sezione/Section 1

Miscellanea





# Versprachlichung der Rolle jüngerer Bevölkerungsschichten in Pandemiediskursen: syntaktische Funktionen, semantische Rollen und deontische Aspekte

PAOLO CANAVESE  
Université de Genève  
paolo.canavese@unige.ch

GORANKA ROCCO  
Università di Trieste  
grocco@units.it

## ABSTRACT

This study, which is part of the larger project COVIDisc - A focus on vector populations, sets out to analyse the media discourse in German and Italian-speaking Switzerland on the role of young people in the Covid-19 pandemic. After identifying the main lexical means used to refer to young people, the focus shifts to the syntactic functions and semantic roles in which they occur, as well as the deontic aspects of the discourse. Furthermore, a chronological analysis makes it possible to trace the construction of a common ground regarding the role of young people in the pandemic context. More specifically, it helps to identify deontic scenarios concerning the behaviours to be adopted or avoided, the actual compliance with the provisions issued by the authorities and the responsibility for raising awareness among young people. This data, reported and illustrated with a series of examples, will serve as a basis for future in-depth quantitative studies, which are essential for refining the interpretation of the trends that have emerged.

## KEYWORDS

Covid-19, contrastive discourse analysis, corpus study, Switzerland, vector population

Im Rückblick auf die ersten Monate der Covid-Pandemie kann angesichts der unsicheren Faktenlage und der zunächst nur vorläufigen Forschungsergebnisse von einer hochgradig komplexen Situation gesprochen werden, die durch „echte epistemische Unsicherheit“ (Stollorz 2021) gekennzeichnet ist. Zu den Fragen, die im Jahr 2020 Gegenstand öffentlicher (wissenschaftlicher, politischer, massenmedialer) Diskussionen waren, gehört auch die Frage nach der Rolle der jüngeren Bevölkerungsschichten in der Ausbreitung der Pandemie. Die vorliegende Arbeit, die die Ergebnisse einer ersten explorativen Untersuchung resümiert, nähert sich dieser Frage im Rahmen der Ziele des Projekts *COVIDisc – A focus on vector populations*<sup>2</sup> und fokussiert dabei die Covid-bezogenen Mediendiskurse in der deutsch- und italienischsprachigen Schweiz. Es gilt zu ermitteln, wie die Rolle und das Verhalten der jüngeren Bevölkerungsschichten (im Folgenden JB) versprachlicht werden. Im Mittelpunkt des Interesses stehen dabei die folgenden Aspekte:

(1) **Repräsentanten der JB.** Wie wird auf JB referiert? Welche lexikalischen Mittel dienen der Versprachlichung der JB (z.B. *Jugendliche, junge Menschen, giovani, ragazzi*) und wie variieren diese in morphologischer Hinsicht?

(2) **Syntaktische Funktionen und semantische Rollen.** In welchen syntaktischen Funktionen und semantischen Rollen treten JB-Repräsentanten auf? Sind beispielsweise JB in der Subjektrolle tendenziell eher Handlungsträger, Adressaten oder Betroffene und in welchen diskursiven Konstellationen treten einzelne Rollen auf?

(3) **Common Ground und Deontik des Pandemieverhaltens.** Wie bildet sich der Common Ground im Sinne des geteilten Wissens über die Pandemie heraus? Können Anhaltspunkte für eine Chronologie der Herausbildung gemeinsamer Wissensbestände in Bezug auf die Rolle der JB identifiziert werden? Wie wird das stets neu ausgehandelte Wissen zu den Verhaltensweisen versprachlicht, die als erforderlich bzw. notwendig gelten? Welches Verhalten gilt in unterschiedlichen Diskurszusammenhängen als riskant, gemeinschaftsgefährdend, verantwortlich oder unverantwortlich, angebracht oder unangebracht?

In Anlehnung an Grice (1989: 65, 274) beschreibt Stalnaker Common Ground (im Folgenden CG) als „presumed background information shared by participants in a conversation“ (2002: 701, Anm. 1) oder auch als „the mutually recognized shared information in a situation in which an act of trying to communicate takes place“ (ibid: 704). Die Dimension der Annahme einer generalisierenden Annahme bzw. Einstellung kommt dabei besonders in der folgenden Definition zum Ausdruck: „It is common ground that  $\phi$  in a group if all members accept (for the purpose of the conversation) that  $\phi$ , and all believe that all accept that  $\phi$ , and all believe that all believe that all accept that  $\phi$ , etc.“ (ibid: 716). In der vorliegen-

1 Der Aufsatz entsteht aus gemeinsamen Überlegungen und Analysen. Kap. 1 und 2.2 wurden von Rocco verfasst, 2.1 und 3 von Canavese und Rocco, 2.3 von Canavese.

2 Vgl. Anm. 4.

den Arbeit gilt das Interesse v.a. denjenigen Aspekten des CG, die die *Deontik*<sup>3</sup> *des Pandemieverhaltens* betreffen: Was kristallisiert sich in der neu entstandenen Situation diskursiv als das heraus, was *sein soll* bzw. *nicht sein soll*? Es geht also um den diskursiven Umgang mit Notwendigkeit, Pflicht und Verantwortung und um die Fragen, wer aus den rekurrenten Aussagen als verantwortlich für das Verhalten einzelner Gruppen oder als Adressat von Anweisungen und Aufklärungsversuchen hervorgeht und in welcher sprachlichen Form sich die jeweiligen Aussagen manifestieren.

Aus methodischer Sicht soll diese erste explorative Analyse dazu beitragen, die Schritte einer quantitativ orientierten Korpusanalyse auszudefinieren. Daher sollen im Folgenden die qualitativ untersuchten Aspekte auch im Hinblick auf ihre methodischen Implikationen erörtert werden.

Im Vorfeld zur Darstellung der Ergebnisse noch einige methodische Erläuterungen: Die korpuslinguistische Grundlage bilden die Schweizer Korpora S\_AL\_DE\_SNF\_COVID19 und S\_AL\_IT\_SNF\_COVID19,<sup>4</sup> wobei dem Untersuchungsgegenstand entsprechend nur mit den Subkorpora der Medientexte gearbeitet wurde. Insgesamt enthalten die zum Zweck der vorliegenden Studie untersuchten Subkorpora über 27 Mio. Wörter für das Deutsche und über 6 Mio. Wörter für das Italienische. Beide Korpora sind Teil von Swiss-AL, einem mehrsprachigen Schweizer Webkorpus (Krasselt et al. 2020) und wurden durch sogenanntes Webcrawling erstellt. Sie ermöglichen die Analyse aller im Netz frei verfügbaren Artikel der wichtigsten Massenmedien der Deutschschweiz und des Tessin, die auf den Zeitraum zwischen Dezember 2019 und Oktober 2020 zurückgehen und die Ausdrücke “Corona.\*”, “COVID.\*” und/oder “SARS-CoV.\*” beinhalten.

Ausgangsbasis für eine erste Zusammenstellung des zu analysierenden Sprachmaterials bildeten die Lemmata, die sich aufgrund einer ersten explorativen Durchsicht als Hauptredemittel zur Versprachlichung der JB erwiesen haben, für Deutsch z.B. *jung\**, *jugendlich\**, für Italienisch *giovan\**, *ragazz\**, *adolescent\** (vgl. 2.1).

Mit Blick auf die Teilziele (2) und (3) wurde aufgrund der ersten explorativen Durchsicht angenommen, dass bei der Korpusanalyse eine formalisierte Suche (z.B. Verschlüsselung als syntaktisches Subjekt, Modalverb) ohne eine vertiefende qualitative Untersuchung, die auch semantische Rollen einbezieht, potenziell zu Informationsverlusten und Fehlern führen kann. In besonderem Maße gilt

- 3 Die Deontik bezeichnet hier generell den grammatischen, lexikalischen oder pragmatischen Ausdruck deontischer Bedeutung, von den Modalverben (Harden 1989) und freien Infinitivkonstruktionen (Deppermann 2006) bis hin zu den lexikalisierten deontischen Bedeutungen im Sinne von Hermanns (2002: 347f., 1989) und Stocker (2005: 116f., 181f.).
- 4 Beide Teilkorpora gehören zum Korpus Swiss-AL (Krasselt et al. 2020) und wurden im Rahmen des SNF-Projekts COVIDisc - A focus on vector populations erstellt (SNF-Sonderauschreibung Coronaviren, Projekt-ID 196328). Die Autoren des Beitrags haben die Korpora als Partner dieses Projekts analysiert. Das Korpus S\_AL\_DE\_SNF\_COVID19 umfasst rund 29 Mio. Wörter. Das Korpus S\_AL\_IT\_SNF\_COVID19 umfasst rund 6.7 Mio Wörter.“

dies für komplexe Satzstrukturen (Relativsätze, Infinitivkonstruktionen, Spalt- und Sperrsätze und andere Formen der Herausstellung) und textgrammatische Aspekte (über Satzgrenzen hinausgehende Verweise und Funktionen) sowie für die vielfältigen sprachlichen Erscheinungsformen der Deontik. Die explorative Erkundung der chronologischen Aspekte (2.3) beschränkt sich auf die ersten Monate der Pandemie (Februar bis Mai 2020), wohingegen die restlichen Aspekte für den Zeitraum von Februar bis Oktober 2020 untersucht wurden.

## 2. ERGEBNISSE DER EXPLORATIVEN KORPUSUNTERSUCHUNG

### 2.1. REPRÄSENTANTEN DER JB IM DEUTSCHEN UND IM ITALIENISCHEN KORPUS

Die Versprachlichung der JB erfolgt im italienischen Subkorpus durch verschiedene mit dem bestimmten Artikel als Generalisierungssignal eingeleitete Syntagmen wie *i giovani, i ragazzi, gli adolescenti*, im deutschen Subkorpus tendenziell durch pluralische Substantive (z.T. substantivierte Adjektive) in meist artikellosen Syntagmen wie *Jugendliche, junge Menschen, junge Leute, junge Erwachsene, die Jungen*. Besonders in den ersten Wochen der Pandemieausbreitung sind aber auch (teilweise mit *si tratta di/es handelt sich um* eingeleitete) singularische Formen zu beobachten, die auf Individuen referieren und diese teilweise mit Blick auf die jeweiligen Aufenthalts- oder Reiseorte näher beschreiben: *un/una giovane proveniente da/di rientro da/che abita/ha soggiornato a...* bzw. *ein junger Mann, eine junge Frau, eine junge Baslerin, ein Jugendlicher aus...*

Bei den pluralischen Syntagmen fällt auf, dass sich die insgesamt schwach ausgeprägte Tendenz zu expliziter Benennung beider Geschlechter häufiger im Zusammenhang mit dem Thema Ausbildung und Berufseinstieg abbildet als in den Aussagen, die enger auf die Verhaltensweisen im Zusammenhang mit der Pandemieausbreitung bezogen und negativ konnotiert sind. So sind z.B. die meisten der relativ wenigen Belege für *jung\* Frauen und Männer* (16) bzw. *jung\* Männer und Frauen* (9) und besonders die ersteren positiv konnotiert und auf Schule, Studium, Berufsausbildung, Berufsperspektiven oder Militär bezogen: Junge Frauen und Männer sind hier Subjekte oder Nutznießer positiv dargestellter Initiativen, Veranstaltungen und Entwicklungen (s.u.) und nur vereinzelt Akteure von Handlungen, die im Pandemiezusammenhang als Risikoverhalten einstuftbar sind.

**Die jungen Frauen und Männer**, die kurz vor dem Ende ihrer Berufslehre stehen, sollen diese trotz Corona-Krise abschliessen können. (Südostschweiz, 6.4.20, Für die Lernenden zeichnet sich eine Notlösung ab)<sup>5</sup>

5 Bei längeren Zitaten wird neben der Quelle und dem Datum auch der Titel angegeben. Die Anführungszeichen und die typographischen Merkmale (z.B. Leerzeichen) wurden aus Übersichtlichkeitsgründen durchgehend vereinheitlicht; die Fettdruckhervorhebungen sind von den Autoren des Beitrags.

Trotz der aussergewöhnlichen Situation sollen **die jungen Frauen und Männer** ihren Lehrabschluss mit einem eidgenössischem Fähigkeitszeugnis oder einem eidgenössischen Berufsattest realisieren können. (Grenchner Tagblatt, 4.4.20, Lehrabschluss trotz Krise)

Dasselbe gilt auch für die lediglich viermal verzeichnete italienische Entsprechung *i giovani e le giovani* (*collocare tutti i giovani e le giovani che intendono iniziare un apprendistato*, tio 24.8.20; auch als Teil der Präpositionalergänzung: *con l'obiettivo di permettere ai giovani e alle giovani di entrare in un'azienda formatrice nel 2021*, tio 24.8.20) und für andere Formen, die eine explizite Bezugnahme auf beide Geschlechter durch den Einsatz morphologischer Mittel erlauben: *i ragazzi/le ragazze*, *gli/le adolescenti*. So sind z.B. bei insgesamt 43 Vorkommen von *(le) ragazze e (i) ragazzi* (20) und *i ragazzi e le ragazze* (23) lediglich drei eindeutig auf die Verantwortung bzw. das Verhalten im Zusammenhang mit der Covidpandemie bezogen: *stiamo chiedendo a ragazze e ragazzi di essere estremamente responsabili* (tio 10.5.20); *ragazzi e ragazze che bevono, chiacchierano senza mascherine* (tio 24.5.20); *le regole anti-Covid sono poco rispettate fra le ragazze e i ragazzi di quarta e terza* (rsinews 15.5.20). Dagegen wurden diese Formulierungen größtenteils im Kontext der (v.a. bildungs- und arbeitsmarktbezogenen) Bestrebungen, Maßnahmen und Initiativen angetroffen, die Jugendlichen und besonders vulnerablen Personengruppen zugute kommen sollen: *webinar in cui le ragazze e i ragazzi possono dare spazio alla loro creatività* (tio 23.3.20); *formazione a distanza e istruzione dei nostri ragazzi e ragazze* (tio 17.3.20); *corsi (...) destinati a ragazze e ragazzi provenienti da scuole pubbliche e private* (tio 25.5.20); *voler riportare ragazze e ragazzi fisicamente in aula* (tio 8.8.20).

## 2.2. SYNTAKTISCHE FUNKTIONEN, SEMANTISCHE ROLLEN UND DEONTIK DES PANDEMIEVERHALTENS

Die folgenden Ausführungen illustrieren die Vielfalt der möglichen massenmedialen Darstellungen verschiedener Ereignisse, Verhaltensmuster und Einstellungen seit Beginn des Pandemiegeschehens. Im Mittelpunkt des Interesses steht dabei die Frage nach der Repräsentation der JB durch verschiedene syntaktische Funktionen und semantische Rollen. Zur Sprache kommen dabei außerdem die möglichen methodischen Konsequenzen der einzelnen Befunde, u.a. die Besonderheiten, die sich beim explorativen Vergleich der deutschen und italienischen Korpusbelege ergeben und bei quantitativer Untersuchung größerer Textmengen berücksichtigt werden sollten.

Aufgrund der Analyse des italienischen Subkorpus erwies sich zum einen, dass die häufigste italienische Versprachlichung der JB *giovan\** an sehr unterschiedlichen syntaktischen Konstruktionen beteiligt ist (Infinitivkonstruktionen, Spalt- und Sperrsätze, *giovan\** als Subjekt von Zustandsbeschreibungen

und -veränderungen und passivischen bzw. passivähnlichen Konstruktionen, als Bestandteil von Lokalangaben, Bezugswort von Relativsätzen usw.).

Zum anderen weisen die Aussagen, in denen auf *giovan\** referiert wird, oft über die Satzgrenzen hinaus. Beide Beobachtungen sollten insbesondere bei der quantitativen Analyse berücksichtigt werden. So fungiert *giovan\** im nachfolgenden Beispiel nicht als syntaktisches Subjekt des im ersten Teilsatz ausgedrückten und durch das Adverb *incredibilmente* kommentierten Zustands *essere sprovvisti di mascherine*, und auch nicht als AGENS. Der erwähnte Zustand wird im nächsten Teilsatz und im darauf folgenden Satz durch *manca* sowie *rischia di compromettere* als unerwünschtes Risikoverhalten resümiert und besonders JB attribuiert.

**Incredibilmente** sui trasporti pubblici (bus, TILO) molti viaggiatori sono sprovvisti di mascherine, in particolare **tra i giovani in gruppo**. È una **manca** che **rischia di compromettere** la fase di riapertura delle attività iniziata in tutta la Svizzera. (tio, 17.5.20, Se fallirà la riapertura, sarà stata la colpa di... )

Eine über das Korpus hinausgehende Suche, die die kontextuelle Bedeutung von *manca* im zitierten Beispiel näher beleuchten sollte, ergab Folgendes : In einer ähnlichen Textvariante auf der Seite des VPOD (Verband des Personals öffentlicher Dienste) ist die Kritik am erwähnten Verhalten durch die Attributerweiterung zu *manca* (*manca di rispetto*) sowie durch den anaphorischen Bezug *questa negligenza* expliziter: “Incredibilmente sui trasporti pubblici (bus, TILO) molti viaggiatori in data 11 maggio erano sprovvisti di mascherine: questa negligenza ci sembra marcata tra i giovani in gruppo. È una manca di rispetto, che rischia di compromettere (...)”.<sup>6</sup>

Auch in den nachfolgenden Beispielen kommen JB innerhalb von Präpositionalausdrücken *tra/fra* (*i*), *in*, *nei*... vor, was methodisch gesehen auf die Bedeutung einer korpuslinguistischen Untersuchung der Präpositionalausdrücke mit dem Kern JB (*giovan\** u.a.) verweist. Das syntaktische Subjekt des jeweiligen (Teil-)Satzes ist infolge der Nominalisierung der jeweiligen Entwicklung oder der Verschiebung der thematisierten Gruppen in Relativsätze oft ein anderes, abstrakteres: *un calo nell'utilizzo della mascherina, la percentuale di chi non è disposto a, impennata di casi (...) legata all'ambiente dei giovani, la percezione del rischio*. So ist z.B. die Annahme, dass die Risikowahrnehmung bei jungen Erwachsenen reduziert ist, auf die folgende verbale Struktur mit JB als syntaktischem Subjekt zurückführbar: „Junge Erwachsene (SUBJ) haben eine reduzierte Wahrnehmung der Gefahren“.

“Certo. Non mancheranno quelli che si comporteranno come se nulla fosse successo. Chi è in là con gli anni si farà qualche problema in più. Ma soprattutto **tra i giovani**, che non percepiscono il Covid-19 come una minaccia mortale, ci sarà chi vuole dare libero sfogo agli ormoni”. (tio, 13.05.2020, Ecco come cambieranno sesso e amore per le persone single)

6 Vgl. *Per un uso obbligatorio della mascherina igienica sui trasporti pubblici* (15.5.20); <https://vpod-ticino.ch/news/2020/per-un-uso-obbligatorio-della-mascherina-igienica-sui-trasporti-pubblici/>

Abbiamo osservato che dall'inizio di agosto c'è stato un calo di quasi il 20 % **tra i giovani adulti** (20-35 anni) nell'utilizzo della mascherina nei luoghi pubblici e nei negozi, mentre tra gli adulti fino a 64 anni l'utilizzo viene mantenuto. (rsinews, 31.8.2020, Corona Immunitas, i primi risultati)

Questo ci dice che **tra i giovani adulti** la percezione del rischio è ridotta. (rsinews, 31.8.20, Corona Immunitas, i primi risultati)

La nuova impennata di casi sembra essere **legata all'ambiente dei giovani**, alle feste di fine anno scolastico e del weekend pre-partenza per la scuola reclute. (tio, 4.7.2020, 97 nuovi casi di Covid-19 da ieri)

In Bezug auf die semantischen Rollen erweist sich nach einer ersten qualitativen Durchsicht, dass die fokussierten Nominalsyntaxagmen wie bereits angekündigt nicht vordergründig in der AGENS-Rolle auftreten, sondern vielmehr als THEMA im Sinne des Gegenstands einer kognitiven, kommunikativen oder emotionalen Situation, als EXPERIENCER eines Zustands oder Geschehens, d.h. Träger eines psychischen Zustandes, also ein Wahrnehmender oder Empfindender oder als PATIENS im Sinne des physisch Affizierten einer Handlung ohne Kontrolle über diese Handlung:<sup>7</sup> *sono soprattutto i giovani/(non) sono (più) (solo) i giovani a essere infettati/a contrarre la malattia; ad ammalarsi sono anche i più giovani; il (corona)virus colpisce anche i giovani/può essere pericoloso anche per i giovani/può danneggiare anche i giovani; anche i giovani sono toccati/diffusori del virus/non sono immuni al virus.*

Im Hinblick auf die für die Ermittlung des CG relevante Frage, was bei bestimmten Formulierungen als gegeben bzw. selbstverständlich angenommen wird (Stalnaker 2002: 702), ist der Gebrauch von Fokuspartikeln von besonderem Interesse. Das Potenzial der Fokuspartikeln wurde im Rahmen dieser explorativen Studie in erster Linie aufgrund der im gegebenen Kontext besonders auffälligen Partikeln *anche* (it.) und *auch* (dt.) untersucht. Wie den oben zitierten Ausschnitten zu entnehmen ist, reflektieren diese ein ständiges soziales Aushandeln der Wissensbestände zum Thema Pandemie/Coronavirus: „(alle, auch) JB“ fungiert oft als affiziertes oder effiziertes Objekt, als PATIENS/EXPERIENCER/THEMA eines schwer beeinflussbaren Geschehens, bei dem eher dem Virus bzw. verschiedenen Personifizierungen des Pandemiegeschehens die aktive Rolle als Subjekt (von *causare*, *colpire*, *danneggiare*, *betroffen sein*, *verursachen* usw.) und AGENS zukommt.

Laut der Lagebeurteilung werden bei diesem wahrscheinlichsten Szenario zudem **auch junge Menschen** an Covid-19 sterben. (Blick, 21.3.20, Beatmungsplätze in der Schweiz werden knapp)

Verfechter einer Durchseuchung verweisen darauf, dass praktisch nur alte Leute mit Vorerkrankungen an Covid-19 sterben, während jüngere Menschen milde bis gar keine Krankheitssymptome zeigen. Das stimmt grundsätzlich, aber **auch junge und ver-**

7 Vgl. <https://grammis.ids-mannheim.de/terminologie/937>; letzter Zugriff am 23.9.21.



**meintlich gesunde Menschen** können schwer erkranken und müssen selbst bei einer Heilung Langzeitfolgen befürchten. (Watson, 27.7.20, Warum die Durchseuchung keine Alternative zum Lockdown ist)

Jüngere verkraften eine Infektion besser als ältere, wie Cerny sagt. Nichtsdestotrotz sind **auch junge Menschen** nicht vor Komplikationen gefeit und vor allem, sie können ältere Menschen anstecken. (Blick, 14.8.20, Sechsmal mehr Infizierte als vor zwei Monaten)

Der Bedeutungsbeitrag der additiven Partikel *auch* bzw. *anche* ist allgemein hin präsuppositioneller Art, während der des Restsatzes ohne *auch* assertiert wird (Sudhoff 2012: 206, Berger 2011: 136, Pasch et al. 2003: 154, Stalnaker 1973, Andorno 2005: 132-133, Lombardi Vallauri 2009: 18-19, Sbisà 2010: 1132). Im Unterschied zu *soprattutto*, *in particolare* bzw. *vor allem* und *besonders* scheint *anche* durch die präsuppositionsauslösende Rolle zumindest stellenweise im Dienste der euphemistisch versprachlichten Rollen und impliziten Zuschreibungen zu stehen. Das jeweilige fokussierende Adverb präsupponiert, dass es eine Gruppe gibt, auf die die jeweilige Aussage eher zutrifft bzw. nicht zutrifft: So präsupponiert die Aussage „Es ist notwendig, dass *auch* X begreifen, was um uns herum passiert“, dass X eben (noch) nicht zu denjenigen gezählt werden können, die den jeweiligen Sachverhalt „begreifen“, dass also der jeweilige Gehalt der Assertion (vor allem) für andere Gruppen gilt.

Eppure è **necessario** che **anche i giovani comprendano** cosa sta succedendo intorno a loro, sia perché possono essere veicoli di trasmissione del virus, sia perché le conseguenze sul rendimento scolastico potrebbero segnare a lungo e in modo drammatico il loro percorso di formazione. (rsinews, 26.3.20, Speciale “Whatsavirus”)

Anche l'estate sarà un periodo in cui **restare** particolarmente **vigili, anche i giovani**, perché il giovane magari non si ammala, ma nella sua cerchia potrebbe esserci qualcuno di più fragile. (ticinonews, 6.7.20, Un ticinese su dieci ha gli anticorpi)

Die über 65-jährigen sollen zu Hause bleiben. Aber **die Jungen** müssen **auch Verantwortung** übernehmen. (BaZ online, 16.3.20, „Ist das alles zu spät?“ Fragen über Fragen an den Bundesrat)

Das Video erreichte eine Viertelmillion Views auf Youtube, war der erste erfolgreiche Deutschschweizer Erklärfilm zum Virus. Es sei jetzt **extrem wichtig, auch die Jungen** zu erreichen, sagt Büsser am Telefon. (Tagesanzeiger, 24.3.20, Gefährdeter Spassmacher)

Die Deontik signalisieren dabei besonders die folgenden Redemittel: Infinitivkonstruktionen, Modalverben, (z.T. von intensivierenden Adverbien begleitete) Adjektive und auch Nomen mit deontischen Konnotationen wie *restare vigili*, *necessario*; *extrem wichtig* + Inf., *Verantwortung*, *sollen*, *müssen*, an anderen Stellen auch *mananza di rispetto*, *sprovisti di* (s.o.). Die Sequenz *auch/anche* + JB ist dabei stellenweise auch Teil von konzessiven oder adversativen Konstruktionen, die

vor allem ältere, aber auch jüngere Bevölkerungsschichten als anfällig ausweisen (*nichtsdestotrotz, aber*) und für die Bestimmung des CG und der Deontik des Diskurses ebenfalls aufschlussreich sind.

Von Relevanz für eine größer angelegte quantitative Untersuchung sind neben den hier illustrierten Präsuppositionsauflösern auch andere Fokuspartikeln (*solo, persino, nur, allein, sogar, usw.*), die im Zusammenhang mit dem Unterschied zwischen impliziter und expliziter Adressierung untersucht werden könnten (z.B. *anche* vs. *soprattutto*; *auch* vs. *vor allem, besonders*) und wie im folgenden Beispiel auch das Gegenteil von *giovan\** betreffen können (*non solo i vecchi* im ersten Satz und *giovani* erst im darauf folgenden Satz). In Frage kommen auch Präpositionen, Satzkonnectoren und andere Redemittel, die Konzessivität und Adversativität signalisieren.

Il coronavirus **non** è una malattia che colpisce **solo i vecchi**. A sostenerlo è Philip Tarr, medico dell'ospedale di Basilea Campagna, che ha parlato dei problemi avuti da due **giovani** (su una trentina di ricoverati) della struttura. (tio, 19.3.20, Il Covid-19 non colpisce solo i vecchi: casi di ventilazione artificiale per pazienti non anziani)

Im Zusammenhang mit der Versprachlichung der JB in Präpositionalausdrücken sind für das Italienische besonders auch die mit *a/ai* eingeleiteten Präpositionalsyntagmen zu erwähnen: JB können hier in der Rolle des ADRESSATS, d.h. Empfängers in einem kommunikativen Akt<sup>8</sup> auftreten, wobei das Empfangene hier eine Botschaft oder Empfehlung ist. So kann z.B. die jeweilige Gruppe ADRESSAT eines Appells, einer Ansprache, Warnung oder Aufforderung zu mehr Verantwortung oder zu einem bestimmten Verhalten sein, aber auch als BENEFATIV eines Erklärungsversuchs oder Ratschlags interpretiert werden.

Was die methodischen Konsequenzen anbelangt: Relevant für die Ermittlung der JB-Versprachlichungen dieser Art bei größeren Textmengen sind zum einen spezifische Präpositionalausdrücke, die mit der ADRESSAT- oder auch BENEFATIV-Rolle korrespondieren, z.B. *ai giovani/ragazzi, agli adolescenti* usw. Von Relevanz sind zum anderen spezifische Verben des Sagens und allgemeiner Verben, die eine Aufforderung, Mitteilung, Erklärung oder Empfehlung ausdrücken können (*rivolgarsi, spiegare, consigliare*), sowie ihre Ableitungen (*appello, invito, messaggio ai giovani*). Zu berücksichtigen ist allerdings, dass bei quantitativen Analysen größerer Textmengen durchaus auch andere, semantisch verwandte Verben einzubeziehen sind, die ebenfalls die Rolle des ADRESSATS einer mehr oder weniger nachdrücklichen Empfehlung oder Aufforderung zulassen.

Für das Deutsche spielt die Präposition *an* eine stellenweise mit *a* vergleichbare Rolle, doch insgesamt scheinen die Verben mit dem entsprechenden semantischen Gehalt bei der Untersuchung der ADRESSAT-Rollen zielführender zu sein als morphosyntaktische Mittel, da das ADRESSAT im Allgemeinen sowohl als reiner Dativ oder Akkusativ und auch als Präpositionalergänzung auftreten kann (*sagen, warnen, erreichen, appellieren*).

8 Vgl. <https://grammis.ids-mannheim.de/terminologie/937>; letzter Zugriff 23.9.21.

Die Weltgesundheitsorganisation (WHO) **appelliert an junge Leute**, sich während der Coronavirus-Krise an die Ausgangsbeschränkungen zu halten. (Südostschweiz, 20.3.20, WHO-Chef an die Jugend in Corona-Zeiten: Ihr seid nicht unbesiegbare)

Die erwähnten Verben erlauben ferner, die Redemittel zu identifizieren, die auch den CG und die Deontik des Diskurses abbilden. Im Einzelnen haben sich in dieser Hinsicht folgende Versprachlichungsstrategien und Redemittel als wichtig herausgestellt:

- Nomen, die Appelle, Botschaften und Initiativen versprachlichen. It. *appello, campagna, messaggio: Appello ai giovani: Esercitare la responsabilità personale, evitare assembramenti* (ticinonews, 11.3.20), *campagna “Stai a casa, leggi un libro” rivolta ai giovani* (ticinonews 12.3.20), *alcuni messaggi si rivolgono specificatamente ai giovani* (ticinonews 24.9.20). Dt. *Appell, Botschaft, Warnung, Kampagne: richtet (...) einen Appell an die Jugendlichen im Landratsaal. „Ich warne aber davor, das Coronavirus zu unterschätzen* (Grenchner Tagblatt, 9.9.20).
- Verben des Sagens i.w.S. und entsprechende elliptische Strukturen (Subjekt + Präposition + JB). It. *rivolgersi/rivolgere, spiegare, dire: bisogno di rivolgersi anche dire ai giovani, in ragione dell’abbassamento dell’età* (azione 24.8.20), *ai giovani bisogna spiegare che non è una applicazione* (ticinonews 10.7.20), *Voglio dire ai giovani: “non siete invincibili”* (ticinonews 20.3.20); *Koch ai giovani: “Cercate di proteggere gli altri”* (ticinonews 4.3.20). Dt. *appellieren, warnen (s.o.), WHO-Appell an die Jugend: Nicht unbesiegbare!* (Blick 20.3.20)
- Fokussierungsmittel. It. *anche, soprattutto, in particolare*. Dt. *auch, vor allem, besonders*
- auf die Deontik des Pandemiediskurses verweisende Redemittel. It. *appello perentorio, i ragazzi devono capire, bisogno di rivolgersi, bisogna spiegare, è il momento di + Infinitiv, non va sottovalutato, comprendere l’importanza, evitare assembramenti*. Dt. *extrem wichtig, Verantwortung, unterschätzen, Ausgangsbeschränkungen*
- verbale Aktionsart und Steigerungsmittel als Präsuppositionsauslöser. It. *È il momento di diventare responsabili. Lo dico ai giovani* (tio 20.3.20), *appelli rivolti ai giovani per una maggiore responsabilità* (tio 17.5.20). Dt. *Aber die Jungen müssen auch Verantwortung übernehmen* (BaZ online 16.3.20).

Die beiden letzten Aspekte sind eng miteinander verbunden: Die Formulierungen *È il momento di diventare responsabili* bzw. *müssen auch Verantwortung übernehmen* und *appelli rivolti ai giovani per una maggiore responsabilità* weisen die Notwendigkeit, verantwortlich zu handeln, als zentralen Aspekt der Deontik des Pandemiediskurses aus und signalisieren zugleich einen möglichen Mangel an Verantwortung, insofern die prozessuale Komponente der Ausdrücke *diventare responsabili* und *Verantwortung übernehmen* und der Ruf nach *mehr* Verantwortung zunächst einmal das Nichtvorhandensein oder die Unzulänglichkeit der jeweiligen Eigenschaft präsupponieren können. Zu erwähnen sind in diesem

Zusammenhang auch mehrere andere Formulierungen, die indirekt das Verhalten oder die Einstellung der adressierten Gruppe thematisieren: der oben zitierte WHO-Appell und mehrere Verweise auf die Notwendigkeit, den Ernst der Lage zu begreifen: *I ragazzi devono capire che andare in giro può essere pericoloso* (tio 19.3.20).

Abschließend seien noch die Fälle erwähnt, in denen ein JB repräsentierendes Syntagma als AGENS, also Urheber einer Handlung fungiert, der Kontrolle über diese Handlung hat. Solche Fälle wurden ebenfalls belegt, sind jedoch tendenziell seltener zu beobachten: *Secondo i primi dati i giovani adulti utilizzano sempre meno la mascherina* (rsinews 19.3.20).

Im nachfolgenden Abschnitt soll aufgezeigt werden, wie sich die skizzierten Tendenzen von Monat zu Monat im Sprachvergleich entwickeln.

### 2.3. CHRONOLOGIE DER DISKURSBILDUNG

Zum Zweck der chronologischen Übersicht wurden die einzelnen Subkorpora nach Monaten aufgeteilt, so dass verfolgt werden konnte, wie sich die Herausbildung des geteilten Wissens von Monat zu Monat gestaltet. Die Chronologie der Herausbildung diskursrelevanter Wissensbestände soll im Folgenden an einigen Beispielen illustriert werden. Durch eine Parallelbetrachtung des deutschen und italienischen Subkorpus soll dabei versucht werden, Ähnlichkeiten und Unterschiede in der Diskursbildung im deutsch- und italienischsprachigen Raum zu eruieren. Diese Chronologie ist auf die ersten Monate (Februar-Mai 2020) beschränkt, d.h. auf den Zeitraum, in dem der Diskurs zur Rolle der Jugendlichen in der Pandemie besonders dynamisch scheint, vermutlich auch weil die Herausbildung des CG noch *in fieri* ist.

Im Februar 2020 stehen JB noch nicht im Zentrum der Covid-Diskurse, da die Virusausbreitung erst vor kurzer Zeit begonnen hat und der erste in der Schweiz bestätigte Fall (ein 70-Jähriger in Tessin) auf den 25. Februar zurückgeht. In dieser Zeit besteht noch kein CG zur Pandemie und bei den im Zusammenhang mit Covid-19 thematisierten Repräsentanten der JB handelt es sich eher um Einzelfälle, im Ausland oder in der Schweiz, was sich indirekt auch in den niedrigen Trefferzahlen für Pluralformen reflektiert.

In Italien wurden drei neue Fälle von Coronavirus bestätigt. Darunter ist ein **Jugendlicher** aus dem Veltlin, der mit Fieber ins Spital in Sondrio eingewiesen wurde, wie „blick.ch“ berichtet. (Südostschweiz, 23.02.2020, Corona-Fall eine halbe Autostunde von Bündner Grenze entfernt)

Il laboratorio dell'ospedale universitario di Basilea ha rilevato due casi positivi di coronavirus. Si tratta di una **giovane donna** di Basilea Città e di un **giovane** che abita fuori cantone. (ticinonews, 27.2.20, Coronavirus: 9 casi totali in Svizzera)

Erst im März 2020, als sich die Lage in der Schweiz schnell verschlechtert und zur Verkündung der „ausserordentlichen Lage“ und somit des ersten Lockdowns am 16. März führt, beginnt auch die Auseinandersetzung um die Rolle der JB in der Corona-Krise. In dieser ersten Phase wird vor allem festgestellt, dass sich *auch* JB anstecken können. Der erste Beleg mit der Sequenz *auch* + JB erscheint im deutschen Korpus am 10. März, allerdings in Bezug auf die Erfahrungsberichte aus dem Ausland:

Beispiele aus dem Ausland würden zeigen, dass **auch junge gesunde Leute** an einer Beatmungsmaschine landen und schlimmstenfalls sterben können. (Blick, 10.3.20, Durchseuchung mit Coronavirus ist „unethisch“)

Nach wenigen Tagen wird aber *über die folgende* Aussage des Bundesamtes für Gesundheit berichtet:

Das Bundesamt für Gesundheit reagierte sofort in anderen Medien: **Auch junge gesunde Leute** könnten an einer Beatmungsmaschine landen und schlimmstenfalls sterben. (Blick, 17.3.20, Das Corona-Wörterbuch)

Ähnlich erscheint in der Tessiner Presse am 13. März ein Artikel, in dem die Erfahrung einer Italienerin in China thematisiert wird:

“Dobbiamo pensare ai più deboli, tenendo presente che **anche i giovani** non sono immuni al virus”. (tio, 13.3.20, Il coronavirus visto da un’italiana a Shanghai)

Den ersten auf die Schweiz bezogenen Beleg findet man erst zwei Wochen später, und zwar in einem Artikel mit dem folgenden Titel: *Il coronavirus colpisce **anche i giovani**: “Non capiamo perché alcuni muoiano”* (tio, 26.3.20). Zu beobachten ist auch ein Perspektivenwechsel: Er manifestiert sich wie in 2.1 angedeutet im Wechsel von den anfänglichen Berichten über Einzelschicksale zu den generalisierenden Pluralformen, der auf die allmähliche Herausbildung des CG hindeutet. Der Krankheitsverlauf bei JB wird als meist nicht allzu schwer beschrieben, obwohl vereinzelt auch über Komplikationen und Todesfälle berichtet wird. Insbesondere wenn es um Jugendliche im Schulalter geht, wird das Risiko tendenziell als geringer eingestuft:

Se la chiusura delle scuole dell’obbligo potrebbe addirittura peggiorare la situazione (molti ragazzi finirebbero con l’essere gestiti dai nonni, gruppo a rischio, mentre **per i ragazzi il Coronavirus è meno pericoloso dell’influenza comune**). (tio, 11.3.20, Coronavirus: niente panico, ma comportamenti adeguati e decisioni forti)

Die Tatsache, dass auch JB dem Risiko ausgesetzt sind, etabliert sich also ab März als Bestandteil des CG, es scheint jedoch weiterhin Unsicherheit darüber zu herrschen, inwieweit sie zur Diffusion des Virus beitragen. Erste Belege zu den JB als Vektoren sind in beiden Sprachen ebenfalls schon im März 2020 zu finden:

“Sentire dire che i **ragazzi** contraggono il coronavirus in forma lieve pur restandone vettori preoccupa i docenti”, ha aggiunto. (ticinonews, 11.3.20, I docenti sono preoccupati)

Außerdem werden JB teilweise als eine Gefahr für die ältere Bevölkerung (z.B. Enkelkinder für Großeltern) dargestellt, auch im Zusammenhang mit der schrumpfenden Zahl der Intensivbetten, die die Frage nach dem Vorrang bei schweren Erkrankungen aufwirft:

Sollen in der Corona-Krise **junge Kranke** bevorzugt behandelt und ältere sterben gelassen werden? (swissinfo, 16.3.20, Wie Schweizer Ärzte die Triage bei Corona-Patienten vornehmen werden)

Nichtsdestotrotz werden die JB nicht direkt angeschuldigt: Die Verantwortung für deren Sensibilisierung wird in mehreren Textbelegen, vor allem im italienischen Korpus, indirekt Erwachsenen zugeschrieben. Dies ist nicht immer der Fall in der deutschsprachigen Presse, in der die Verantwortung der JB teilweise expliziter ausgedrückt wird:

Molti lo hanno fatto notare, una delle categoria [sic] più difficili da convincere a tenere dei comportamenti responsabili sono i **giovani**. (ticinonews, 16.3.20, L'idea: un concorso su Instagram contro il coronavirus)

Ausserdem sollten speziell die **Jugendlichen** mithelfen, das Virus einzudämmen: „Sie haben viel mehr soziale Kontakte und sind mobiler. (Grenchner Tagblatt, 19.3.20, „Ich habe Vertrauen in die Schweizer Bevölkerung“: BAG-Spitzenbeamter Koch in der „Rundschau“)

Insgesamt sind Jugendliche, wie in 2.2 ausgeführt, eher Subjekte passivischer Konstruktionen; die Infizierung wird darüber hinaus auch durch Adjektive oder Nominalisierungen versprachlicht: *Il ragazzo trovato positivo* (ticinonews, 10.3.20). *Ein junger Mann ist [...] auf das Coronavirus getestet worden* (srf 3.3.20), *Il ragazzo infetto* (tio, 3.3), *della positività del ragazzo* (ticinonews, 10.3.20).

Wie bereits erwähnt, erscheinen die JB oft auch in der ADRESSAT-Rolle. Die Sensibilisierung der Jugendlichen durch Appelle zeichnet sich in beiden Sprachen schon ab Mitte März ab, wobei meist Zitate und direkte Rede verwendet werden. So veranschaulichen die nachstehenden Belege die Appelle des kantonalen Arztes Tessin Giorgio Merlani und der Zürcher Regierungsrätin und Bildungsdirektorin:

“Ragazzi state casa. Rispettate le regole”. **L'appello - rivolto ai bambini e ai giovani** che da oggi non andranno più a scuola - arriva direttamente dal medico cantonale ticinese Giorgio Merlani che durante l'odierna conferenza stampa ha più volte ripetuto l'importanza che “tutti facciano la propria parte”. (tio, 13.3.20, L'appello ai ragazzi: “State a casa. Rispettate le regole”)

Sie **appellierte deshalb an die Kinder und Jugendlichen**, dass sie, die ansonsten unbeschwert und manchmal auch unvorsichtig sein dürften, nun ihr Verhalten ändern und mithelfen sollten. (Grenchner Tagblatt, 14.3.20, Chaostage an den Schulen in Aussicht)

Ein parallel dazu verlaufender Diskursstrang betrifft die Einstellung der Jugendlichen gegenüber der pandemischen Lage: Einerseits wird ein zu lockeres Verhalten beklagt, andererseits werden sie als verantwortungsvolle Menschen beschrieben:

**I ragazzi** delle scuole medie di Gravesano **non sembrano essere particolarmente preoccupati** per la situazione e ai microfoni di Teleticino hanno spiegato cosa hanno appreso sul coronavirus. (ticionews, 2.3.20, "Il tasso di assenza sul 10%, siamo nella media")

Da un sondaggio esterno pubblicato il 24 marzo dall'istituto Sotomo, che ha interrogato 30'000 persone in Svizzera, risulta che **i giovani sono più favorevoli all'introduzione di misure più restrittive** da parte del governo. (swissinfo, 27.3.20, I numeri del coronavirus in Svizzera)

Im deutschen Subkorpus finden sich Belege, in denen der Vorwurf der Verantwortungslosigkeit vergleichsweise expliziter zum Ausdruck gebracht wird, wie z.B. *Polizei entrüstet über junge Corona-Ignoranten* (20minuten, 18.3.20), *Junge Menschen tendieren dazu, das Coronavirus zu unterschätzen, weil sie sich nicht betroffen fühlen*. (Watson, Die Superspreader, 28.3.20). Die Thematisierung der Verantwortung bzw. Verantwortungslosigkeit scheint aber stellenweise auch im italienischen Subkorpus einen anderen Ton anzunehmen, zumindest wenn es um JB in anderen Ländern geht. Die entsprechenden Textbelege enthalten negativ konnotierte Ausdrücke, wie die folgenden Aussagen über italienische und deutsche Jugendliche zeigen:

**Indifferenti** all'emergenza sanitaria del Coronavirus, circa 300 giovani - i dati ufficiali non sono ancora disponibili - si sono riuniti per un rave party in uno stabilimento industriale in disuso a Robecco d'Oglio, in provincia di Cremona, in piena zona gialla. (tio, 1.3.20, Fregandosene del coronavirus, un rave party nella "zona gialla")

La Germania è **indignata** per i Coronaparty dei giovanissimi  
In troppi aggirano i divieti di riunirsi in gruppo, mostrando di **non aver capito la gravità** della situazione (tio, 18.3.20, La Germania è indignata per i Coronaparty dei giovanissimi)

Im April finden sich verstärkt Belege für zwei neue Teildiskurse an. Einerseits werden erste Überlegungen zu den (überwiegend wirtschaftlichen) Folgen der Pandemie für JB angestellt und entsprechende Statistiken und Fakten dargeboten. Andererseits thematisieren v.a. italienischsprachige Medien geschlossene Schulen und das Recht der Jugendlichen, in die Schule zu gehen. Obwohl es zu diesem Zeitpunkt weitgehend gefestigt scheint, dass die JB ein kleineres Risiko tragen, bleibt jedoch weiterhin die Frage offen, inwieweit die möglichen Ansteckungen im Schulbereich zur Ausbreitung des Virus beitragen.

Covid-19: i disoccupati impennano

A "pagare" l'epidemia sono soprattutto i **giovani**. (tio, 7.4.20, Covid-19: i disoccupati impennano)



Besonders hart trifft die Coronakrise die Jugendlichen. Bei den 15- bis 24-jährigen stieg die Arbeitslosenquote von 2.3 auf 2.8 Prozent an. (srf, 7.4.20, Coronakrise hinterlässt erste Spuren auf dem Arbeitsmarkt)

Uno studio francese suggerisce che i giovanissimi non siano un vettore importante per l'infezione (tio, 21.4.20, Coronavirus: i bambini "sono poco contagiosi")

Gerade im Alter nach der offiziellen Schulpflicht werden die Jugendlichen immer mehr zu Vektoren des Virus. (Watson, 27.4.20, Wer für und wer gegen Maturaprüfungen ist)

In diesem Zusammenhang ist also hauptsächlich die Rede von JB im Schulalter als potentiellen Vektoren, wohingegen studierende und berufstätige JB noch nicht so oft in diesem Kontext thematisiert werden. Wie aber bereits in 2.2 erwähnt, treten dabei JB eher selten in der AGENS-Rolle auf, vor allem im italienischen Korpus. Die angesprochenen JB werden tendenziell nicht direkt wegen potentiell riskanter Verhaltensweisen angeschuldigt:

E poi c'è l'incognita scuola. **Impossibile pretendere che bambini e ragazzi** portino la mascherina ed evitino contatti fisici. (azione, 27.4.20, Tempo di superare la paralisi)

Das Ziel ist, dass **man die Jugendlichen** auf die Weisungen des Bundes **sensibilisiert** und die Situation erklärt. (sfr, 18.4.20, „Die virtuelle Welt ersetzt den physischen Kontakt nicht“)

Hier zeigen sich bei der Diskursbildung mehrere unterschiedliche Verläufe. Einerseits werden also die Jugendlichen nicht direkt als Schuldige bezeichnet, andererseits wird auch nicht klar dargestellt, wer (welche Erwachsenengruppen) die Verantwortung für ihr Verhalten und ihre Sensibilisierung tragen soll. Zur Abschwächung der stellenweise thematisierten Verantwortungslosigkeit kommen auch Aussagen vor, in denen die JB als bewusste und vernünftige Menschen beschrieben werden: *la maggioranza dei giovani svizzeri concorda con le misure imposte* (rsi, 2.3.20).

Im deutschen Korpus sind auch für den April z.T. auch Belege zu finden, in denen JB als Subjekt und zugleich Träger von Handlungen auftreten, die als verantwortungsloses Verhalten oder Verstöße gegen die Regeln gelten.

Im Kanton Aargau hat die Bevölkerung die Corona-Vorschriften über die Ostertage gemäss Polizeiangaben mehrheitlich gut eingehalten. **Besonders junge Menschen lebten jedoch den Verhaltensmassnahmen nicht konsequent nach.** Insgesamt gab es rund 300 Interventionen. [...] **Jugendliche und junge Erwachsene verstiesen oft gegen die Vorschriften** [...]. (Grenchner Tagblatt, 14.4.20, 300 Interventionen übers Osterwochenende)

Im Mai scheint sich der generelle Diskurs bezüglich des Verhaltens der JB auch weiter zu ändern. Dies betrifft die JB im weiteren Sinne und nicht hauptsächlich Schüler: Es zeigt sich zunehmend die Tendenz zur Fokussierung von Freizeitaktivitäten, Feiern und Spaß, die sich dann besonders auch in den Sommermonaten



fortsetzt. Dieser Perspektivenwechsel betrifft vor allem die italienischen Texte, in denen bisher eher vage oder abschwächende Formulierungen verwendet wurden, ohne die Verantwortung der JB direkt anzusprechen. Der Perspektivenwechsel ist auch an der syntaktischen und rollensemantischen Struktur der folgenden Beispiele erkennbar, in denen JB in der Subjekt- und zugleich AGENS-Rolle auftreten.

**Quasi mille giovani**, ignorando il divieto di assembramenti imposto per il coronavirus, si sono riuniti ieri sera a Losanna per una partitella di calcio. (tio, 22.5.20, Mille giovani attorno al campo in barba ai divieti)

Bis zu **300 Jugendliche feierten** am Wochenende eine Party am Neuenburgersee. [...] **Diese jungen Leute** hätten mit einem gewissen Egoismus gehandelt und nicht verstanden, dass die Weisungen im Zusammenhang mit der Coronavirus-Pandemie trotz Lockerungen weiterhin gelten. (Blick, 11.05.20, 300 Jugendliche feiern illegale Party und vertreiben)

Der Diskurs über die wirtschaftlichen Folgen der Krise für die JB, Arbeitslosigkeit und Schwierigkeiten bei der Suche nach Ausbildungsplätzen verdichtet sich im Vergleich zum Vormonat, was auch an einigen Schlagzeilen ersichtlich ist: *Generazione lockdown, 1 giovane su 6 è senza lavoro* (ticinonews, 27.5.20); *Bis zu 6000 Jugendliche könnten demnächst ohne Lehrstelle dastehen* (Watson, 1.5.20).

Die Rolle der Jugendlichen in der Ausbreitung der Pandemie bleibt allerdings weiterhin umstritten und reflektiert die teilweise entgegengesetzten Positionen auch innerhalb der Wissenschaft:

Ne è risultato che in Svizzera **i ragazzi dai 10 ai 20 anni** contribuiscono fortemente alla diffusione del virus e, in misura superiore alla media, anche la fascia dai 35 ai 45 anni. Gli anziani, invece, diffondono il virus in misura molto inferiore alla media. (ticinonews, 27.5.20, "Una nuova ondata potrebbe fare fino a 5'000 morti")

Die Rolle von **Kindern und Jugendlichen** bei der Übertragung des Coronavirus bleibt höchst unsicher. Das steht in einer am Freitag publizierten Untersuchung der wissenschaftlichen Covid-19-Taskforce des Bundes. (srf, 1.5.20, App des Bundes soll ab Mitte Mai verfügbar sein)

Zum Abschluss dieses Abschnitts kann festgehalten werden, dass die deutsch- und italienischsprachigen Schweizer Medien in den ersten Monaten der Pandemie die JB auf ähnliche Art und Weise abbilden. Die Rolle der JB in der Entwicklung der Pandemie wird ab März 2020 thematisiert. Die Ansteckungs- und Ausbreitungsgefahr bei den Jugendlichen erweist sich als sehr umstritten, auch über den hier näher analysierten Zeitraum hinaus. Diese Debatte korreliert mit der Unsicherheit in der Wissenschaft: Neue Studien werden durchgeführt und führen zu unterschiedlichen Erkenntnissen, die die Festigung einer eindeutigen Rolle der JB im pandemiebezogenen CG verhindern. Dies könnte z.T. auch mit den sprachlichen Abschwächungsstrategien im JB-Diskurs korrelieren.

Eine in chronologische Abschnitte aufgeteilte qualitative Analyse ermöglicht jedoch auch weitere Hypothesen zu diesem Phänomen, z.B. indem sie die Rolle der Heterogenität der JB deutlicher in Erscheinung treten lässt: In der ersten Phase (März und April) wird vor allem über JB im Schulalter berichtet und eine Zuschreibung der Verantwortung eher vermieden. Die Perspektive ändert sich im Mai, als vermehrt über Freizeitverhalten berichtet wird und dabei etwas expliziter auf den Verantwortungsmangel hingewiesen wird. Ein weiteres, damit verbundenes Ergebnis ist, dass sich die Haltung gegenüber den JB in den deutschsprachigen Texten als vergleichsweise kritischer erweist. Dieser Befund bedarf allerdings einer quantitativen Vertiefung und ist sicher auch vor dem Hintergrund der Unterschiede im Pandemieverlauf zu sehen.

### 3. FAZIT

Die wichtigsten Ergebnisse dieser explorativen Untersuchung zur Versprachlichung der Rolle junger Bevölkerungsgruppen und zur Herausbildung des Common Ground sollen im Folgenden mit besonderem Blick auf ihre methodischen Konsequenzen zusammengefasst werden.

An erster Stelle ist die grundsätzliche Polysemie und Kontextsensibilität der untersuchten lexikalischen Elemente zu erwähnen: Die anvisierte soziale Gruppe erweist sich als diskursrelevant und als äußerst heterogen (JB als Schüler, Studierende, Auszubildende, Ausbildungsplatzsuchende, Arbeitsuchende, Grenzpendler, Grenzgänger, Partyfeiernde, Konzertbesucher, Touristen, Urlaubsreisende, Reisende mit öffentlichen Verkehrsmitteln). Die jeweils unter JB subsumierbaren Ausdrücke wie z.B. *ragazzi*, *giovani*, *Jugendliche* und *junge Menschen* können im pandemiebezogenen Diskurs je nach Zeitraum (z.B. März 2020 vs. Mai 2020) auf andere Akteure und andere Situationen referieren (z.B. die Schweiz vs. Ausland, Tourismus vs. Arbeit).

Ein weiterer, besonders mit Blick auf die Korpusannotierung und -abfrage wichtiger Punkt ist die Tendenz zu sprachlicher Abschwächung, die bei den JB-bezogenen Aussagen auffällt. Sie manifestiert sich in der Vermeidung direkter Prädikationen und v.a. darin, dass die Adressierung der Akteure, die potenziell gegen die Deontik des Pandemieverhaltens „verstoßen“, durch höchst verschiedene syntaktische Funktionen und semantische Rollen erfolgt.

Insgesamt zeigt sich, dass die diskursive Konstruktion des CG zum Pandemiegeschehen im Allgemeinen und zur Rolle der JB im Besonderen auf viele unterschiedliche Gruppen und Situationen referiert. Sie ist gekennzeichnet durch eine beträchtliche Dynamik und Polyphonie im Sinne einer großen, für bestimmte Adressatengruppen möglicherweise kaum überschaubaren Vielfalt von jeweils unterschiedlich versprachlichten Sichtweisen, Empfehlungen, Expertenaussagen und Regeln. Eine systematische und umfassende Analyse dieser polyphonen Vielfalt scheint auch mit Blick auf aktuelle und künftige gesellschaftsrelevante Fragen (z.B. Impfung) unabdingbar.

- Andorno C. (2005) *Linguistica testuale: un'introduzione*, Roma, Carocci.
- Berger F. (2011) "Die Fokuspartikel »auch« im Erstspracherwerb. Früh vorhanden – spät verstanden? Methodologische Maßnahmen zum Nachweis eines frühen Verständnisses", *Spektrum Patholinguistik*, 4, pp. 135-139.
- Deppermann A. (2006) "Deontische Infinitivkonstruktionen. Syntax, Semantik, Pragmatik und interaktionale Verwendung". In S. Günthner & W. Imo (Hrsg.), *Konstruktionen in der Interaktion. Bd. 20, Linguistik – Impulse & Tendenzen*, Berlin/New York, De Gruyter, pp. 239-262.
- Grice P. (1989) *Studies in the Way of Words*, Cambridge MA, Harvard University Press.
- Harden T. (1989) "Ausdrucksweisen der deontischen Modalität im Deutschen und Portugiesischen". In A. Franco & J. Schmidt-Radefeld (Hrsg.), *Das linguas em contraste: Português e Alemão*, Porto, Editora da Universidade do Porto, pp. 123-136. <http://ler.letras.up.pt/uploads/ficheiros/artigo5921.pdf> (letzter Zugriff 27.4.21)
- Hermanns F. (1989) "Deontische Tautologien. Ein linguistischer Beitrag zur Interpretation des Godesberger Programms (1959) der Sozialdemokratischen Partei Deutschlands". In J. Klein (Hrsg.), *Politische Semantik*, Opladen, Westdeutscher Verlag, pp. 69-149.
- Hermanns F. (2002) "Dimensionen der Bedeutung I: Ein Überblick". In D. A. Cruse, F. Hundsnurscher, M. Job & P.R. Lutzeier (Hrsg.), *Lexikologie. Ein internationales Handbuch zur Natur und Struktur von Wörtern und Wortschätzen*, 1. Halbband (*Handbücher zur Sprach- und Kommunikationswissenschaft*, Bd. 21.1), Berlin/New York, De Gruyter, pp. 343-350.
- Krasselt J., Dreesen P., Fluor M., Mahlow C., Rothenhäusler K. & Runte M. (2020) "Swiss-AL: A Multilingual Swiss Web Corpus for Applied Linguistics". In *Proceedings of the 12th International Conference on Language Resources and Evaluation*, Marseille, European Language Resources Association, pp. 4138-4144. <https://www.aclweb.org/anthology/2020.lrec-1.510> (letzter Zugriff 27.4.21)
- Lombardi Valluri E. (2009) *La struttura informativa. Forma e funzione degli enunciati linguistici*, Roma, Carocci.
- Pasch R., Brauße U., Breindl E. & Waßner U. H. (2003) *Handbuch der deutschen Konnektoren. Linguistische Grundlagen der Beschreibung und syntaktische Merkmale der deutschen Satzverknüpfers (Konjunktionen, Satzadverbien und Partikeln)*, Berlin/New York, De Gruyter.
- Sbisà M. (2010) "Pragmatica". In R. Simone (a cura di), *Enciclopedia dell'italiano*, Roma, Istituto della Enciclopedia Italiana, pp. 1127-1134.
- Stalnaker R.C. (2002) "Common Ground", *Linguistics and Philosophy*, 25:5-6, pp. 701-721.
- Stalnaker R.C. (1973) "Presuppositions", *Journal of Philosophical Logic*, 2, pp. 447-457.
- Stollorz V. (2021) "Herausforderungen für den Journalismus über Wissenschaft in der Coronapandemie – erste Beobachtungen zu einem Weltereignis", *Bundesgesundheitsblatt – Gesundheitsforschung – Gesundheitsschutz*, 64, pp. 70-76. <https://link.springer.com/article/10.1007/s00103-020-03257-x> (letzter Zugriff 27.4.21)
- Stocker C. (2005) *Sprachegeprägte Frauenbilder: Soziale Stereotype im Mädchenbuch des 19. Jahrhunderts und ihre diskursive Konstituierung*, Band 262, Reihe Germanistische Linguistik, Tübingen, Niemeyer.

Sudhoff S. (2012)  
“Fokuspartikelinventare des  
Niederländischen und Deutschen”.  
In B. Rothstein (Hrsg.), *Nicht-  
flektierende Wortarten*, Berlin/  
Boston, De Gruyter, pp. 203-223.



# Metegol, Futbolín e Gool!: tre versioni a confronto

ANTONELLA DE LAURENTIIS  
Università del Salento  
[antonella.delautentiis@unisalento.it](mailto:antonella.delautentiis@unisalento.it)

## ABSTRACT

*Metegol* (2013) is the first real 3D animated feature film produced in Argentina, directed by the Oscar winner and film director Juan José Campanella, and loosely based on the book by the Argentinian writer Roberto Fontanarrosa, “Memorias de un wing derecho”. This paper compares the original version of the film with the Italian and peninsular Spanish dubbed versions. The analysis will focus on some scenes including evident features of the Argentinian variety of the Spanish language, with particular attention to the idiomatic locutions and to the lexis of lunfardo. The aim of this analysis is to show how these elements have been translated, adapted or neutralized in the two dubbed versions. Furthermore, the study will describe the way the target audience may influence the translation strategies adopted and will outline the features of audiovisual dubbed products addressed to children.

## KEYWORDS

Dubbing, cartoons, dual audience, idiomatic locutions, translation strategies

## 1. INTRODUZIONE

In questo contributo presentiamo un'analisi comparativa tra la versione originale di *Metegol*, film di animazione del premio oscar Juan José Campanella, e le due versioni doppiate rispettivamente in Spagna e in Italia.<sup>1</sup> L'obiettivo è quello di verificare in che modo siano stati tradotti, adattati e/o neutralizzati quegli elementi appartenenti alla categoria della cultura linguistica (Molina, 2001) considerando anche il duplice destinatario di questo genere di film costituito da un pubblico infantile e un pubblico adulto. Attraverso un'analisi descrittiva di tipo qualitativo cercheremo di individuare le principali strategie impiegate nella traduzione di elementi linguistici e culturali specifici della varietà dello spagnolo rioplatense, seguendo la catalogazione proposta da Gottlieb e riadattata ai fini della nostra ricerca.

*Metegol*, primo film d'animazione girato in 3D in Argentina, è uscito nelle sale cinematografiche nel mese di luglio del 2013 e ha subito ottenuto un grande successo vincendo vari premi, tra cui il *Premio Sur* per la migliore sceneggiatura (2013) e il *Premio Goya* come miglior film d'animazione (2014). La versione doppiata in Spagna è stata presentata per la prima volta in occasione del Festival del cinema di San Sebastián nel mese di settembre del 2013, con il titolo di *Fútbolín* mentre nelle sale italiane uscirà l'anno successivo, nel maggio del 2014 con il titolo di *Gool!*

Il film è liberamente ispirato al racconto dello scrittore argentino Roberto Fontanarrosa, "Memorias de un wing derecho", nel quale un giocatore di biliardino racconta le sue vittorie, paragonandosi ai veri campioni del calcio, come Maradona o Pelé, ed esprime tutta la sua dedizione per questo sport, considerato la sua "única verdad".

Se nel breve racconto di Fontanarrosa la narrazione è affidata al monologo dell'anonimo giocatore di biliardino, nel lungometraggio di Campanella la voce narrante è quella di Amadeo ormai padre, che in un lungo flashback racconterà a suo figlio e a tutti noi la storia di un Amadeo prima bambino e poi ragazzo, fanatico del biliardino, che trascorre le sue giornate nel bar di un paese senza nome. Ad un certo punto, posto davanti alla possibilità che il progresso distrugga la sua comunità, decide di trasformarsi nell'eroe del suo paesino, accompagnato dagli stessi giocatori del biliardino che, animandosi, lo seguiranno nelle sue avventure.

1 Esiste anche la versione doppiata in spagnolo neutro che è stata oggetto di uno studio approfondito e che sarà pubblicato sul prossimo numero della rivista *Lingue e Linguaggi* (De Laurentiis, 2021).

L'intero sviluppo del lungometraggio avvolge lo spettatore nella grande passione per il calcio, soprattutto grazie all'ironia che caratterizza i piccoli giocatori del biliardino, i quali dimostrano come sia possibile collaborare e instaurare un'amicizia anche con gli avversari, in quanto "lo que nos une es mucho más que lo que nos separa". Quattro di loro sono particolarmente importanti per lo svolgimento della storia: Capi, il capitano della squadra de *los rayados*, è il giocatore che più trasmette il senso della condivisione e dell'amore per il calcio, è colui che incoraggia Amadeo a non arrendersi aiutandolo a riconquistare il biliardino, il paese e la sua amata Laura; Beto è l'egocentrico del gruppo, parla di sé in terza persona ed è convinto di essere l'unico talentuoso in una squadra mediocre, mentre Loco rappresenta il mondo della filosofia, sempre pronto a proporre una riflessione cosmica che i suoi compagni difficilmente sono in grado di comprendere, e infine Lisandro, il capitano della squadra avversaria de *los lisos*, il quale ha una forte personalità ed è costretto ad ammettere di aver bisogno dei suoi nemici storici per poter recuperare il suo gruppo.

Il linguaggio utilizzato dai protagonisti del film rispecchia perfettamente la cultura argentina e la passione di questo Paese per il mondo del calcio: "Una película como ésta está empapada de la pasión y la melancolía argentinas, lleva la denominación de origen en la etiqueta".<sup>2</sup> Ma cosa succede quando questa passione è trasportata e riadattata in altri contesti linguistico-culturali?

## 2. TRADUZIONE AUDIOVISIVA E FILM D'ANIMAZIONE

Nel doppiaggio di prodotti audiovisivi, ciò che attira subito l'attenzione è la traduzione del titolo, che nel caso del nostro film inquadra immediatamente il tema trattato. Il titolo originale, infatti, fa riferimento alla presenza di un particolare campo da calcio, il *metegol*, intorno al quale ruotano la maggior parte delle vicende del film. Nella traduzione in spagnolo peninsulare si è optato per un titolo equivalente: ciò che in Argentina si conosce come *metegol*, è maggiormente identificabile in Spagna come *futbolín*. In Italia, invece, il film è stato doppiato con il titolo di *GOOOL!*, espressione che sebbene non si riferisca concretamente al gioco del biliardino e quindi ai suoi piccoli giocatori fa però riferimento all'esultanza presente in una delle prime scene, quando Amadeo sconfigge Ezequiel in una partita di biliardino e Laura non riesce a trattenere l'euforia e alla medesima esclamazione della scena finale, durante la partita di calcio che questa volta si disputerà in uno stadio a misura d'uomo e che vedrà scontrarsi la squadra di giocatori professionisti guidata da Ezequiel e la squadra di "dilettanti", rappresentata dagli abitanti del paesino e guidata da Amadeo.

Prima di procedere con l'analisi delle versioni doppiate riteniamo necessario presentare quelle che sono le principali caratteristiche della traduzione di film di

2 <http://www.eldesconcierto.cl/2013/12/26/campanella-y-metegol-la-importancia-de-tener-algo-que-decir-en-3d/>.



animazione prestando particolare attenzione al possibile destinatario di questo genere, quello infantile, che condiziona molto spesso le scelte traduttive con il fine di rendere accessibile la storia anche al pubblico meta. A tal proposito il regista argentino ha dichiarato, in un'intervista, di aver condiviso l'idea di doppiare le voci dei personaggi sia in Argentina (varietà dello spagnolo rioplatense) che in Spagna (varietà dello spagnolo peninsulare) con l'obiettivo di rendere la storia più comprensibile al pubblico, in particolare a quello infantile, poiché è proprio quest'ultimo il principale destinatario del suo film.<sup>3</sup>

Il film d'animazione si configura come un genere molto importante per lo sviluppo cognitivo del bambino, dal momento che, oggi, un prodotto audiovisivo attrae molto di più rispetto alla lettura di un libro. Il tempo passato davanti alla televisione, infatti, occupa gran parte della giornata di un bambino. In Italia come in Spagna, la maggior parte delle produzioni audiovisive destinate ad un pubblico di minori viene importata dall'estero, per cui la traduzione assume un ruolo fondamentale e, inoltre, trattandosi di un genere che vede come suo principale destinatario un pubblico infantile, la modalità che prevale in ambito audiovisivo è quella del doppiaggio anche nei paesi in cui si privilegia la sottotitolazione.

Chaume Varela definisce il testo audiovisivo come un "texto multidimensional cuya coherencia depende extraordinariamente de los mecanismos de cohesión establecidos entre la narración verbal y la narración visual" (2001: 67): il legame originale tra parole ed immagini conferisce coerenza al testo e la sua trasposizione nella lingua d'arrivo deve riproporre la stessa funzione che il testo aveva per il pubblico di partenza.

La traduzione audiovisiva, quindi, è condizionata soprattutto dall'industria cinematografica e televisiva e dal bisogno di produrre un testo che abbia lo stesso grado di successo presso l'audience di arrivo. A differenza del settore editoriale, quello cinetelevisivo, dispone di un mercato più ampio e ha la responsabilità di soddisfare un pubblico molto più ampio e variegato:

Publicidad y sociedad de consumo obligan, por una parte, a que los niños dejen de ser niños lo más rápidamente posible para que se conviertan en «jóvenes» consumidores cuanto antes, y, por otra parte, a que los viejos consumidores quieran seguir siendo eternamente jóvenes (Yuste Frías, 2006: 200).

I prodotti audiovisivi per bambini presentano delle peculiarità che li differenziano dai lungometraggi destinati agli adulti. Trattandosi di personaggi animati, e non di persone vere e proprie, i movimenti labiali sono meno visibili e molto spesso non coincidono con il testo realmente emesso. Inoltre, il sincronismo espressivo è ridotto e le inquadrature ravvicinate sono rare. Per quanto riguarda il linguaggio, per facilitare la comprensione da parte del pubblico, si tende ad utilizzare un parlato più semplice, caratterizzato da un lessico elementare, pro-

3 <https://doblaje.fandom.com/es/wiki/Metegol>

posizioni omogenee, turni brevi, scanditi e con una maggiore attenzione alla dizione. Anche la presenza degli elementi tipici dell'oralità, quali esitazioni, false partenze, riformulazioni e frasi interrotte è ridotta (De Rosa, 2010: 8-9).

Il traduttore di un film d'animazione, dunque, se da una parte può godere di un grado di libertà maggiore per quanto riguarda i vincoli tecnici del doppiaggio, dall'altra è fortemente influenzato dal destinatario soprattutto in relazione alla traduzione di giochi di parole, di elementi culturali, di espressioni idiomatiche e di varianti sociolinguistiche. Per quanto riguarda il processo traduttivo, il traduttore è posto dinanzi a una macro-scelta, deve decidere se adottare una strategia tesa al mantenimento delle strutture morfosintattiche, del lessico, dello stile della lingua di partenza (straniante), oppure una strategia volta a commutare determinati aspetti morfosintattici, lessicali e stilistici per rendere così la traduzione più vicina alla lingua e alla cultura di arrivo (addomesticante). Quest'ultima risponde a ragioni più pratiche, talvolta di ordine sociale o politico ma molto spesso di ordine puramente commerciale, che non possono essere ignorate e che tendono sempre più a prevalere (García de Toro, 2014).

Per quanto riguarda il pubblico, nel caso di prodotti audiovisivi come i film d'animazione o i cartoni animati possiamo parlare di un "duplice destinatario", il bambino e l'adulto. Tuttavia, nel tradurre elementi legati alla cultura linguistica di un paese, come giochi di parole, locuzioni idiomatiche, metafore, ecc., si è soliti optare per una strategia addomesticante considerando soltanto ciò che il bambino della cultura meta riuscirà a comprendere e provocando, in tal modo, una perdita di elementi da parte dello spettatore adulto (De los Reyes Lozano, 2015).

Il compito principale del traduttore-adattatore di un prodotto audiovisivo è quello di creare un testo sufficientemente "trasparente" per il pubblico di arrivo (Pavesi, 2005: 22). In questo senso, si tende spesso a neutralizzare determinate espressioni culturali o a modificare frasi idiomatiche e giochi di parole con lo scopo di avvicinare il testo allo spettatore della cultura meta.

Nei prossimi paragrafi presenteremo esempi concreti tratti dal film per mostrare le principali strategie traduttive impiegate per tradurre alcuni elementi linguistico-culturali specifici della varietà argentina, confrontando la versione originale con le versioni doppiate in spagnolo peninsulare e in italiano.

### 3. ASPETTI METODOLOGICI

In questa fase della nostra ricerca ci soffermeremo sull'analisi delle scelte traduttive riferite alle categorie che, secondo la definizione di Lucia Molina, appartengono all'ambito della *Cultura Linguistica*: "falsi amici culturali", "espressioni idiomatiche", "figure retoriche" e "giochi di parole".

Con cultura linguistica intendiamo tutte quelle espressioni che nel testo originale presentano delle specificità e che entrando in contatto con un'altra cultura nel processo traduttivo (nel nostro caso ci riferiamo alla modalità del doppiag-

gio) possono presentare, nella maggior parte dei casi, un problema di tipo culturale tra il testo originale e il testo meta (Cfr. Molina 2001: 78 e ss.). Nel nostro caso specifico le scene che abbiamo selezionato presentano interessanti esempi di locuzioni idiomatiche ed espressioni riconducibili al lessico del lunfardo che caratterizzano la parlata argentina e veicolano magistralmente quella sottile ironia tipica della produzione filmica di Juan José Campanella.

Per individuare e analizzare le scelte traduttive degli elementi appartenenti alla cultura linguistica presenti nella versione originale, abbiamo deciso di basarci sulle strategie proposte da Henrik Gottlieb (citato da Diaz Cintas 2002, p. 16) e riadattate ai fini del nostro lavoro. Riportiamo, qui di seguito, una loro breve descrizione:

- a. Congruenza: uso di un'espressione idiomatica che si basa sulla stessa metafora presente nel testo originale (l'E.I. della LM mantiene stesso significato e stessa forma).
- b. Equivalenza: uso di un'espressione idiomatica che si basa su una metafora simile a quella del testo originale (si traduce l'E.I. con una che presenta un significato simile ma struttura diversa).
- c. Corrispondenza: uso di un'espressione idiomatica che si basa su una metafora diversa (sia la struttura che il significato differiscono dall'originale).
- d. Riduzione: uso di una sola parola per tradurre l'espressione idiomatica dell'originale.
- e. Parafrasi: uso di una frase per spiegare il significato dell'espressione idiomatica dell'originale.
- f. Omissione: si elimina l'espressione idiomatica dell'originale (In questo caso non si traduce l'E.I. o perché non esiste un equivalente nella lingua meta o a causa di restrizioni spazio-temporali/sincronismo, o perché non incide sulla comprensione generale del testo originale).
- g. Compensazione: si inserisce un'espressione idiomatica non presente nel testo originale (questa strategia si usa soltanto nei casi in cui si voglia o debba equilibrare e compensare un'E.I. presente in un'altra parte del testo della quale non esisteva un modismo equivalente).

Queste stesse strategie possono applicarsi anche alla traduzione dei termini provenienti dal lunfardo o specifici della varietà dello spagnolo rioplatense, che riflettono un uso molto colloquiale della lingua e che sono quindi connotati sia dal punto di vista diatopico sia dal punto di vista diafasico. Cercheremo di capire se e come queste espressioni siano state tradotte nelle versioni spagnola e italiana utilizzando, da un lato, termini appartenenti allo stesso registro linguistico ma riadattati alla cultura linguistica del testo meta oppure innalzando il registro linguistico attraverso l'uso di un lessico standard e, di conseguenza, neutralizzando quei tratti caratteristici presenti nella versione originale.

Per rendere più comprensibile il lavoro di confronto e analisi delle due versioni doppiate, abbiamo creato delle tabelle in cui si riportano i seguenti dati: nella parte superiore, la trascrizione del testo originale preceduta dall'intervallo

di tempo della scena; nella parte inferiore, due colonne in cui presentiamo le versioni doppiate con le relative strategie e, in corsivo, gli elementi rappresentativi delle aree tematiche.

### 3.1. METEGOL, FUTBOLÍN E GOOOL!: ANALISI DELLE STRATEGIE TRADUTTIVE

Per quanto riguarda le locuzioni idiomatiche, si tratta di espressioni ancorate nella lingua di partenza, per le quali è necessario trovare un equivalente nella versione tradotta o procedere con la neutralizzazione e spiegazione degli elementi costitutivi. Le scene che abbiamo selezionato per questo lavoro presentano interessanti rese traduttive e, come vedremo, esemplificano alcuni di quei tratti caratteristici del linguaggio utilizzato dai personaggi del film. In *Metegol* sono piuttosto ricorrenti, inoltre, espressioni colloquiali o appartenenti al lessico del lunfardo, spesso associate al mondo del calcio e che, nella maggior parte dei casi, non possono essere tradotte letteralmente ma hanno bisogno di un equivalente altrettanto diffuso nel contesto di destinazione.

In questa prima scena che presentiamo, Capi (il capitano della squadra di biliardino) dimostra fin da subito il suo legame con il mondo calcistico: dopo l'arrivo di Ezequiel e la conseguente distruzione della città, Amadeo appare amareggiato perché teme di aver perso tutto ciò che aveva. Il piccolo giocatore, che prende vita grazie alle lacrime di Amadeo, cerca di incoraggiarlo paragonando la situazione ad una partita di calcio.

Versione originale (24:58-25:03)	
Capi: ¡No te me caigas ahora, papá! No te me caigas ahora que <i>a este partido todavía se puede dar vuelta</i> , viejo.	
Amadeo: ¿Qué? ¿Me estás cargando?	
Doppiaggio spagnolo peninsulare	Doppiaggio italiano
Capi: ¡No te derrumbes ahora, amigo! No te derrumbes, hombre, que todavía <i>podemos darle la vuelta al marcador</i> .	Capi: Non devi crollare proprio adesso, amico! Non devi crollare, possiamo ancora <i>ribaltare la partita</i> , sai?
Amadeo: ¿Me estás vacilando?	Amadeo: Mi stai prendendo in giro?
Strategia	Strategia
Congruenza	Parafrasi

Tabella 1

In questa scena ritroviamo l'espressione idiomatica "dar la vuelta al partido", che verrà ripresa più volte nel corso del film per sottolineare il forte legame esistente

tra vita reale e mondo del calcio. L'espressione, piuttosto ricorrente negli articoli sportivi argentini, è sostituita nella versione in spagnolo peninsulare dalla locuzione "darle la vuelta al marcador", che si basa sulla stessa metafora presente nella versione originale ma che è più diffusa nel contesto calcistico spagnolo. Entrambe provengono dalla locuzione verbale colloquiale "dar alguien la vuelta a la tortilla" il cui significato è quello di "Invertir las circunstancias o producir un cambio total en una situación" (DRAE). Se nel doppiaggio spagnolo la strategia utilizzata è quella della congruenza, nella versione italiana si perde la locuzione idiomatica che viene esplicitata attraverso la frase "ribaltare la partita", espressione, questa, diffusa nell'ambito del giornalismo sportivo.

Nella scena successiva il piccolo capitano crede che ritrovare gli altri componenti della squadra sia l'unica cosa indispensabile per superare questo momento difficile. Amadeo, però, pensa soprattutto alla sua amata Laura che è stata rapita da Ezequiel el Grosso.

Versione originale (26:34-26:45)	
<p>Capi: No necesitamos a nadie. A nadie. [...] Ni al pueblo... ni al bar... ni a Laura.            Amadeo: Ni al pueblo... ni al bar... ni a Laura... no. A Laura sí.            Capi: Y la vas a tener. Con el equipo completo, otra que el Grosso. Ya vas a ver cómo viene pidiendo la escupidera. Ya vas a ver.</p>	
Doppiaggio spagnolo peninsulare	Doppiaggio italiano
<p>Capi: No necesitamos a nadie. A nadie. [...] Ni al pueblo... ni al bar... ni a Laura.            Amadeo: Ni al pueblo... ni al bar... ni a Laura... no.            A Laura sí.            Capi: Y pronto la tendrás. Con el equipo completo, seguro que pasa del crack. Ya verás cómo vuelve toda loquita por tus huesos. Ya verás.</p>	<p>Capi: Non ci serve nessun altro. Nessuno. Né la città... né il bar... né Laura.            Amadeo: Né la città... né il bar... né Laura... no. Laura sì.            Capi: E allora sarà tua. Con la squadra al completo altro che il Grosso. Vedrai che tornerà indietro pentita. Vedrai!</p>
Strategia	Strategia
Corrispondenza	Parafrasi

Tabella 2

Nella versione originale Capi, per convincere Amadeo del fatto che Laura scapperà dalle grinfie de el Grosso per tornare da lui mettendo da parte il suo orgoglio, utilizza l'espressione colloquiale "pedir la escupidera", che secondo la definizione del DRAE significa "1. loc. verb. Arg. y Cuba. Acobardarse, tener miedo". La versione in spagnolo peninsulare adotta la strategia della corrispondenza, utilizzando una locuzione idiomatica che si basa su una metafora diversa da quella dell'origi-

nale e presenta una struttura differente: la locuzione “estar loco por los huesos de alguien” sta a indicare il fatto che Laura sarebbe tornata follemente innamorata di Amadeo. Nel caso del doppiaggio italiano, potremmo parlare di una parafrasi anche se il senso dell’espressione non corrisponde a quello dell’originale poiché sembra trasmettere il messaggio che Laura abbia seguito volontariamente il suo rapitore e che quindi la squadra al completo la farà di certo pentire della sua scelta.

In questa scena, quando i giocatori del biliardino finalmente si ritrovano nella discarica, Beto dà subito dimostrazione del suo carattere, vantandosi di essere il giocatore più forte della squadra. Capi, però, non è d’accordo e i due iniziano a discutere.

Versione originale (31:06-31:13)	
Beto: El goleador siempre fue el Beto, todos lo sabemos. Y es muy difícil ser goleador en este equipo <i>de morondanga</i> .  Capi: Equipo de... vos sos <i>de morondanga</i> .	
Doppiaggio spagnolo peninsulare	Doppiaggio italiano
Beto: El goleador siempre fue el Beto, todos lo sabemos. Y es muy difícil meter goles en este equipo <i>de chichinabos</i> .  Capi: Equipo de... tú, tú sí que eres <i>de chichinabo</i> .	Beto: È sempre stato Beto il bomber, lo sappiamo tutti. E non è semplice essere il bomber di una squadra <i>scadente</i> .  Capi: Una squadra? Tu sei <i>scadente</i> !
Strategia	Strategia
Equivalenza	Riduzione

Tabella 3

Beto, per affermare il suo talento e la sua superiorità, cerca di sminuire il gruppo di cui fa parte con un aggettivo che, spesso, viene attribuito alle squadre che non dimostrano di valere particolarmente.

La locuzione aggettivale “de morondanga” è utilizzata con il significato di squadra “de poca calidad, sin valor” (*Diccionario de americanismos*) e ricorre negli articoli giornalistici che mettono in risalto la debolezza o l’inferiorità di una particolare squadra.

Il doppiaggio in spagnolo peninsulare mantiene la stessa metafora presente nell’espressione argentina, utilizzando una locuzione idiomatica più facilmente identificabile dal pubblico di arrivo: “de chicha y nabo”, infatti, fa riferimento ad una squadra “de poca importancia, despreciable” (*DRAE*). Nel caso della versione italiana, ancora una volta la strategia utilizzata è quella della riduzione, infatti l’espressione idiomatica non viene tradotta ma sostituita dal semplice uso dell’aggettivo “scadente”.

Gli insulti agli avversari sono molto frequenti, prima che le due squadre riescano ad instaurare un rapporto di collaborazione. Quando Capi ritrova il tavolo

del biliardino, scopre che i rivali non hanno avuto la fortuna, come lui e i suoi compagni, di liberarsi e camminare con le proprie gambe. Quando si accorge che il portiere della sua squadra è stato ridipinto del colore della squadra avversaria, il bordeaux, inizia una discussione accesa tra i due capitani.

Versione originale (47:24-47:32)	
Lisandro: ¿No te da vergüenza gritarles a compañeros en desgracia?	
Capi: Estos no son compañeros. Desgraciados sí.	
Lisandro: ¿A quién le decís desgraciado?	
Capi: Si toda la vida te tuvimos de hijo...	
Doppiaggio spagnolo peninsulare	Doppiaggio italiano
Lisandro: Qué vergüenza, gritar a compañeros en desgracia.	Lisandro: Non ti vergogni a insultare degli amici in disgrazia?
Capi: Estos no son compañeros. Desgraciados sí.	Capi: Questi non sono miei amici, disgraziati sì.
Lisandro: ¿A quién llamas desgraciados?	Lisandro: Chi hai chiamato disgraziati?
Capi: Pero si toda la vida no has sido más que un paquete...	Capi: Ma se è tutta la vita che vi stracciamo.
Strategia	Strategia
Equivalenza	Riduzione

Tabella 4

Nella versione argentina Capi, per riferirsi al suo avversario, usa l'espressione "Tener de hijo (o de hija) a alguien" il cui significato è il seguente: "tratar a alguien como si se gozara de su tutela, tiranizarlo; en el ámbito deportivo, superar reiteradamente un competidor a su adversario (*Diccionario etimológico del lunfardo*). Nel nostro caso specifico l'accezione che più si adatta al contesto è quella relativa al mondo del calcio, cioè la capacità di un giocatore più forte di vincere continuamente contro uno più debole, riuscendo a sottometterlo in qualsiasi situazione. Una espressione equivalente è quella utilizzata nella versione spagnola, "ser un paquete", molto diffusa in ambito calcistico e che fa riferimento alla scarsa qualità di un giocatore. Nel doppiaggio italiano l'espressione idiomatica è sostituita, in questo caso, dal verbo "stracciare" che riproduce il senso dell'originale, sottolineando la superiorità di una squadra sull'altra.

La stessa scena, riportata nella tabella 5, continua presentando altri elementi interessanti alcuni dei quali appartengono al lessico calcistico sia argentino che spagnolo.

Versione originale (47:40-47:44)	
Capi: Salí de acá, <i>pecho frío</i> .	
Lisandro: ¿A quién le dijiste pecho frío?	
Capi: Además, sos un <i>morfón</i> .	
Lisandro: ¿A quién le dijiste pecho frío?	
Capi: A vos, <i>pecho frío</i> .	
Doppiaggio spagnolo peninsulare	Doppiaggio italiano
Capi: Venga ya, <i>sangre de horchata</i> .	Capi: Ma vattene, <i>smidollato</i> .
Lisandro: Suéltame. ¿A quién llamas horchata?	Lisandro: Espera, a chi hai detto smidollato?
Capi: Además, eres un <i>tufo</i> .	Capi: Sei anche un <i>egoista</i> .
Lisandro: ¿A quién llamas de horchata?	Lisandro: A chi hai detto smidollato?
Capi: A ti, <i>chupón</i> .	Capi: A te (Ø)!
Strategia	Strategia
Equivalenza/Omissione/Compensazione	Riduzione/Riduzione/Omissione

Tabella 5

La prima espressione usata da Capi per offendere l'avversario è “pecho frío”, il cui significato riportato nel *Diccionario etimológico del lunfardo (DEL)* è il seguente: “imperturbable, carente de amor propio o de emoción; cobarde, pusilánime”.<sup>4</sup> Nel doppiaggio spagnolo ritroviamo una espressione equivalente, “sangre de horchata”, che si riferisce a una persona con “carácter calmoso que no se altera por nada” (*DRAE*). La versione italiana presenta ancora una volta una riduzione, sostituendo l'espressione originale con l'aggettivo “smidollato”, che si riferisce a una persona di carattere debole e priva di coraggio.

Il secondo termine appartenente al lessico del lunfardo, *morfón*, proviene dal mondo del calcio e si riferisce a un “jugador que retiene excesivamente la pelota” (*DEL*). In questo caso il termine lunfardo è reso nella versione spagnola con *tufo* e nella versione italiana con *egoista*, in entrambi i casi si perde il riferimento al les-

4 Riportiamo in nota la definizione completa dell'espressione presente nel dizionario di Conde (*DEL*): pecho. m. En la expr. PECHO FRÍO: imperturbable, carente de amor propio o de emoción; cobarde, pusilánime. (Del esp. pecho: parte del cuerpo humano que se extiende desde el cuello hasta el vientre y en cuya cavidad se contienen el corazón y los pulmones, por alusión a la falta de calor en el pecho o a la falta de sangre bombeada desde el corazón, que darían cuenta del entusiasmo y del compromiso.)



sico del calcio. Interessante vedere come in realtà un'espressione perfettamente equivalente a *morfón* è usata nella battuta successiva per tradurre l'espressione argentina *pecho frío*. In questo caso il doppiaggio spagnolo usa il termine *chupón*, che in ambito calcistico si riferisce proprio al giocatore individualista, mentre la versione italiana presenta un caso di omissione.

Il prossimo esempio (tabella 6) vede sempre protagonisti i due capitani delle squadre avversarie che, grazie alla saggezza orientale del giocatore Loco, hanno riscoperto il valore dell'amicizia e del rispetto reciproco.

Versione originale (48:26-48:31)	
Capi: Qué linda época. Cómo nos divertíamos.	
Lisandro: Él era imposible de marcar. Le teníamos un <i>julepe</i> bárbaro.	
Doppiaggio spagnolo peninsulare	Doppiaggio italiano
Capi: Era otra cosa. Qué época.	Capi: Erano altri tempi, noi sì che ci divertivamo.
Lisandro: Este era imposible de marcar. Este nos daba súper <i>yuyu</i> .	Lisandro: Eri impossibile da marcare. Ci facevi una <i>paura</i> tremenda.
Strategia	Strategia
Equivalenza	Riduzione

Tabella 6

In questo esempio Lisandro confessa a Capi che lui e i suoi compagni di squadra lo temevano tantissimo e usa l'espressione colloquiale “le teníamos un *julepe* bárbaro” dove il termine colloquiale *julepe*<sup>5</sup> si riferisce proprio a questo senso di paura. Nella versione spagnola è stato tradotto con un termine equivalente, *yuyu*, usato in modo colloquiale per riferirsi alla sensazione di paura e inquietudine che si prova soprattutto dinanzi a ciò che non si conosce. Il doppiaggio italiano tende, invece, a elevare il registro linguistico, usando il sostantivo, in italiano standard, *paura* neutralizzando e “riducendo” in questo modo l'elemento marcato dal punto di vista diatopico e diafasico presente nella versione originale.

La prossima scena (tabella 7) che presentiamo rappresenta uno dei pochi riferimenti diretti al racconto di Fontanarrosa presenti nel film: si parla del personaggio di Tachola il quale, sebbene avesse la passione per l'alcool, era bravissimo a giocare al biliardino.

5 *julepe*. I.1.m. Bo:O, Ch, Ar, Ur, pop + cult → espon; Pe, Ch, Py, pop. Miedo, susto. (*culepe*). (*Diccionario de americanismos*).

Versione originale (49:00-49:05)	
Capi: Al final estaba tan borracho que me hizo sacar un <i>chutazo</i> de molinete que rompí un vidrio.	
Lisandro: ¡Me acuerdo!	
Capi: ¡Y el <i>atorrante</i> quería que lo pague yo!	
Doppiaggio spagnolo peninsulare	Doppiaggio italiano
Capi: Y al final tenía tal pedal que me hizo pegar un <i>chute</i> de molinete, y rompí un cristal.	Capi: Alla fine era così ubriaco che mi ha fatto fare un (Ø) mulinello e mi ha fatto rompere una finestra.
Lisandro: ¿Te acuerdas?	Lisandro: Mi ricordo!
Capi: ¡Y el muy <i>cachondo</i> quería que lo pagara yo!	Capi: Quel <i>mascalzone</i> voleva che la pagassi io!
Strategia	Strategia
Equivalenza/Corrispondenza	Omissione/Riduzione

Tabella 7

In questa scena ritroviamo due termini colloquiali, *chutazo*<sup>6</sup> e *atorrante*:<sup>7</sup> il primo appartenente al lessico calcistico e indica un calcio violento dato alla palla, il secondo è usato soprattutto in Argentina e Uruguay per indicare una persona sfacciata. Nella versione spagnola il primo termine trova un corrispettivo equivalente in *chute*<sup>8</sup> mentre il termine colloquiale *cachondo*<sup>9</sup> non mantiene lo stesso valore semantico poiché si riferisce a una persona scherzosa. Per quanto riguarda il doppiaggio italiano, nel primo caso siamo davanti a una omissione mentre la traduzione di *atorrante* con il termine *mascalzone* determina, ancora una volta, un innalzamento diafasico.

Gli esempi che presentiamo nelle tabelle 8 e 9 si riferiscono alla stessa scena in cui le due squadre si sono finalmente riunite e Capi chiede al suo amico Amadeo di lanciargli la palla in modo da poter giocare una partita tutti insieme.

6 *chutazo*. m. (fútb.) Puntapié violento dado a la pelota/ efecto de la pelota. (DEL)

7 *atorrante* 2. adj. despect. Arg. y Ur. Desfachatado, desvergonzado. U. t. c. s. (DRAE)

8 *chutar* 1. intr. En el fútbol y otros juegos, lanzar el balón o la pelota con fuerza (DRAE)

9 *cachondo* 4. adj. coloq. Burlón, divertido, bromista. Apl. a pers., u. t. c. s. (DRAE)

Versione originale (52:13-52:18)	
Capi: Amadeo. Tirá la pelota y empezamos un <i>picadito</i> .	
Doppiaggio spagnolo peninsulare	Doppiaggio italiano
Capi: Amadeo. Tira la pelota, y echamos una <i>pachanguita</i> .	Capi: Amadeo, lancia la palla che facciamo una <i>partita</i> .
Strategia	Strategia
Equivalenza	Riduzione

Tabella 8

In questo caso, per riferirsi alla partita di calcio Capi usa *picadito*, diminutivo di *picado*,<sup>10</sup> termine colloquiale che si riferisce a una partita che si gioca tra amici. Nella versione in spagnolo europeo si opta per un termine equivalente, *pachanguita*, diminutivo di *pachanga*,<sup>11</sup> che mantiene lo stesso registro colloquiale e lo stesso valore di partita informale. Nella traduzione in italiano il termine è tradotto con il sostantivo *partita*, eliminando il riferimento al contesto informale in cui si giocherebbe la stessa.

Presentiamo, infine, l'ultimo esempio dove i due capitani, per convincere Amadeo a dimenticare la sua amata Laura, affermano completando l'uno la frase dell'altro, che le donne sono come "las minas", termine lunfardo che riecheggia in molti testi di tango e nella ben nota letteratura argentina.

Versione originale (53:19-53:25)	
Capi: Pero, por favor, Amadeo. Las mujeres son...	
Lisandro: Como las <i>minas</i> .	
Doppiaggio spagnolo peninsulare	Doppiaggio italiano
Capi: Pero, por favor, Amadeo. Las mujeres son...	Capi: Ma per favore, Amadeo. Le donne sono...
Lisandro: Como las <i>pibas</i> .	Lisandro: Come le <i>mine</i> .
Strategia	Strategia
Equivalenza	Omissione

Tabella 9

10 picado. m. Partido de fútbol informal y amistoso, fulbito. (Del esp. picado: dividido en trozos muy menudos y, por ext., mezclado; por alusión a que en un partido de este tipo los jugadores que habitualmente juegan juntos en un equipo se mezclan con otros). (DEL)

11 pachanga. 3. f. coloq. Partido informal de fútbol, baloncesto u otros deportes. (DRAE)

Questa scena presenta interessanti scelte traduttive in entrambe le versioni doppiate. Per quanto riguarda la versione spagnola, sebbene il termine *mina* sia sicuramente comprensibile da parte di qualsiasi pubblico ispanofono e non solo, risulta alquanto “originale” la scelta di tradurlo con un altro termine, *piba*,<sup>12</sup> che di certo ha ben poco a che vedere con la varietà dello spagnolo peninsulare. Il doppiaggio italiano ha probabilmente commesso un errore di comprensione dovuto ai cosiddetti “falsi amici” tra lingue affini, poiché il termine *mina* in italiano non corrisponde assolutamente al significato dell’originale (*donna*). Si tratta, in questo caso, di omissione del riferimento linguistico-culturale presente nella versione originale.

#### 4. RIFLESSIONI CONCLUSIVE

A partire dagli esempi riportati nelle tabelle e prendendo in considerazione altre scene del film che non sono state riportate nel presente studio, possiamo presentare alcuni risultati della nostra analisi basata sul confronto tra la versione del film nella varietà argentina, e le due versioni doppiate nella varietà dello spagnolo peninsulare e in italiano.

Per quanto riguarda il doppiaggio realizzato in Spagna prevale una traduzione di tipo addomesticante: le scelte traduttive mostrano una tendenza ad avvicinare il testo al pubblico di arrivo, adottando espressioni prevalentemente equivalenti che conservano il significato di partenza adattandolo alla lingua e cultura meta. Le strategie che predominano negli esempi riportati sono, infatti, l’equivalenza, la congruenza e la corrispondenza; queste ultime determinano un processo di adattamento degli aspetti specifici della cultura linguistica argentina al contesto linguistico e culturale spagnolo attraverso una ricerca costante di termini ed espressioni appartenenti alla varietà dello spagnolo peninsulare.

Possiamo dunque affermare che il doppiaggio spagnolo, pur presentando una diversa varietà diatopica, riesce a preservare lo stesso uso di un registro colloquiale e informale e a caratterizzare le specificità dei personaggi del film attraverso il loro modo di esprimersi e i loro continui riferimenti al mondo calcistico. In particolar modo, le espressioni idiomatiche, nella maggior parte dei casi, trovano un equivalente nella lingua meta mantenendo lo stesso livello diafasico della versione argentina.

Nella versione doppiata in italiano risulta evidente da un lato, un processo di neutralizzazione delle espressioni idiomatiche tipiche della varietà argentina, con la predominanza di strategie quali la riduzione, la parafrasi e l’omissione e, dall’altro, un innalzamento diafasico poiché la maggior parte dei termini appartenenti al registro colloquiale o marcati dal punto di vista regionale è tradotta

12 *piba*. I.1. f. Ar, Ur. *Mujer joven, especialmente la hermosa y atractiva*. pop + cult → espon. (*Diccionario de americanismos*).

utilizzando un lessico appartenente all'italiano standard. In questo modo anche i giochi di parole o gli effetti umoristici legati alla caratterizzazione linguistica dei personaggi risultano, nella versione italiana, meno incisivi e, in alcuni casi, possono determinare una minore comicità delle scene rispetto alla versione originale.

Ulteriori aspetti che sarebbe interessante analizzare riguardano la versione doppiata in spagnolo peninsulare, nella quale abbiamo riscontrato l'aggiunta di battute e dialoghi assenti nella versione originale; ciò accade in alcune scene in cui si inquadra soltanto l'ambiente esterno, come accade ad esempio nella scena ambientata nella discarica. Si tratta, probabilmente, di una scelta finalizzata ad enfatizzare la comicità del testo di partenza e, probabilmente, a rendere più comprensibili alcune scene in funzione del principale destinatario del film, il pubblico infantile.

- ASALE, *Diccionario de americanismos*. <https://www.asale.org/recursos/diccionarios/damer>
- Campanella J. J. (2013) *Metegol* [DVD], Argentina, Universal.
- Chaume Varela F. (2001) “Más allá de la lingüística textual: cohesión y coherencia en los textos audiovisuales y sus implicaciones en traducción”. In Miguel Duro (ed.), *La traducción para el doblaje y la subtitulación*, Madrid, Cátedra, pp. 65-81.
- Conde O. (2004) *Diccionario etimológico del lunfardo*, Penguin Random House Grupo Editorial Argentina, Edizione del Kindle.
- De Laurentiis A. (2021) “De *Metegol* a *Futbolín*: análisis de las tres versiones en español de la película de Juan José Campanella”, *Lingue e Linguaggi*, 46, pp. 53-70.
- De los Reyes J. (2015) *La traducción del cine para niños. Un estudio sobre recepción*, Tesis doctoral de la Universitat Jaume I de Castellón y la Université de Reims Champagne-Ardenne.
- Díaz Cintas J. (2002) “El subtítulado de expresiones idiomáticas al castellano”. En *Actas de las III Jornadas de Doblaje y Subtitulación de la Universidad de Alicante*, Alicante, Universidad de Alicante, pp. 13-28.
- De Rosa G. L. (2010) (ed.) *Dubbing Cartoonia. Mediazione interculturale e funzione didattica nel processo di traduzione dei cartoni animati*, Casoria (Napoli), Loffredo Editore.
- García de Toro C. (2014) “Traducir literatura para niños: de la teoría a la práctica”, *TRANS. Revista de Traductología*, 18, pp. 123-137.
- López P. (26/12/2013) “Campanella y Metegol: la importancia de tener algo que decir”. <http://www.eldesconcierto.cl/2013/12/26/campanella-y-metegol-la-importancia-de-tener-algo-que-decir-en-3d/> (consultato il 16/09/2021).
- Molina L. (2001) *Análisis descriptivo de la traducción de los culturemas árabe-español*, Tesis doctoral, Universitat Autònoma de Barcelona.
- Pavesi M. (2005) *La traduzione filmica. Aspetti del parlato doppiato dall'inglese all'italiano*, Roma, Carocci Editore.
- Real Academia Española, *Diccionario de la Real Academia Española de la lengua*. <http://www.rae.es>
- Yuste Frías J. (2006) “Traducción y paratraducción de la literatura infantil y juvenil”. En Luna, A. y Montero, S. (eds.), *Traducción e Política editorial de Literatura infantil e xuvenil*, Vigo, Universidade de Vigo, col. T&P, 2, pp. 189-201.



# Reflexiones sobre el error en la traducción jurídica

VERÓNICA DEL VALLE CACELA  
Uniwersytet Marii Curie-Skłodowskiej  
[veronica.delvallecacela@mail.umcs.pl](mailto:veronica.delvallecacela@mail.umcs.pl)

## ABSTRACT

In this article, we intend to reflect on the essential skills that all translators working in the legal field must possess, as well as to highlight the importance of discussing, analysing and reflecting on the mistakes that the student may have made. Understanding translation errors in the legal field seems to us to be a fundamental requirement for carrying out this translation task, as the legal content of these documents means that the seriousness of the error is greater than in other contexts. Likewise, in some sections, we have sought to highlight the value that the didactics of error can have when the pair of languages with which we are working share a certain linguistic affinity, as is the case of Spanish and Italian. In doing so, we have only intended to offer a brief reflection on the need to carry out more empirical studies on how to take advantage of error to solve problems of syntactic calques or false friends.

## KEYWORDS

Legal Translation, Didactics of Specialized Translation, Romance Languages, Law, Translator Training



## 1. LOS LENGUAJES DE ESPECIALIDAD

La lengua engloba multitud de características y matices en cualquier sistema lingüístico. Esta riqueza provoca, a su vez, que se identifiquen particulares concretas en determinados contextos en los que participa:

Mais on donne aussi le nom de langage, au sein d'une langue, à la façon particulière dont celle-ci est parlée dans un groupe ou dans un secteur d'activité, si du moins cette façon présente assez de propriétés linguistiques pour être isolée comme un parler particulier. On dit volontiers qu'il s'agit d'un langage spécial ou spécialisé. C'est en ce sens, que le droit a son langage, de même que l'économie, la médecine ou la sociologie. (Cornu, 2005: 16)

Estos llamados lenguajes de especialidad, como bien señala Cornu, tienen como principal nexo común la disciplina en la que se desarrollan, fomentando una serie de características desde varios aspectos de la lengua: léxico, gramatical, sintáctico, entre otras. Dichas peculiaridades que conforman el lenguaje empleado en un determinado ámbito tienen como principal hilo conductor el distanciamiento que se crea entre el lenguaje en modalidad oral o escrita, ya sea en su uso diario o cuando se emplee en un determinado contexto:

In ogni ambito disciplinare, anche idiografico, tendono a costituirsi amalgami lessicali specifici, linguaggi settoriali che segnano, se non barriere, almeno specifiche modalità di distanziamento rispetto al comune esprimersi parlato e scritto. (De Mauro, 2017: 714)

Este distanciamiento indicado por De Mauro implica que, solo los especialistas en la materia en cuestión, puedan comprender completamente el discurso, por ejemplo, sobre inteligencia artificial, biología o economía. No obstante, el constante intercambio léxico entre los campos de especialidad y la lengua estándar provoca que se transfieran conceptos entre uno y otro, aunque con connotaciones distintas ya que difícilmente se mantiene el significado conocido por algunos de estos grupos, a saber, legos *versus* no legos en la materia. Esto será un aspecto destacable a lo largo del presente artículo que parte desde el punto de vista del traductor pues como indica Trovato (2018: 99):

El hilo conductor que se puede localizar entre lengua y traducción se sitúa en las llamadas lenguas de especialidad, esto es, variedades funcionales de la lengua que permiten la comunicación del saber especializado entre especialistas y legos en una determinada materia. Las lenguas de especialidad cuentan con características comunes a la variedad estándar de la lengua pero se diferencian de esta sobre todo desde una perspectiva léxica y terminológica.

Sin embargo, cuando el ámbito de especialidad en el que nos movemos abarca situaciones estrechamente ligadas con el Derecho, la terminología no será el único escollo que tendrá que afrontar el traductor dado que las características gramaticales y sintácticas cobran relevancia en este tipo de lenguaje.

El Derecho es una disciplina cuyas ramificaciones impregnan cada acción que una persona realiza pues, todo es proclive a tener valor legal explícita o implícitamente. Dentro de esta disciplina, el lenguaje empleado se torna esencial para poder aplicar la norma jurídica o interpretarla: «la linguistique juridique examine les signes linguistiques que le droit emploie et les énoncés que le droit produit» (Cornu, 2005: 1). Una relación entre la lengua y el Derecho que lleva a Cornu (2005: 1) a plantearse si «le droit même n'est pas, en soi, un langage».

El sistema jurídico en el que se engloba la temática de este artículo es el Derecho continental que se caracteriza, entre otras cuestiones, por tener sus normas escritas y regirse por ellas. Necesita de la lengua para dar sentido a los términos que emplea en ella dando lugar a que se hable de una terminología jurídica (Cornu, 2005: 13-14) o de un lenguaje jurídico:

Quando si dice «linguaggio giuridico» o, con un'espressione assimilabile alla corrispondente inglese, «linguaggio legale», si ricorre a un'etichetta di estensione variabile, adattata, dai non-giuristi almeno, a un universo testuale composito, in cui si riconoscono varietà di lingua concorrenti. (Mortara Garavelli, 2001: 7)

Se han realizado numerosos estudios con el fin de enumerar y establecer aquellos puntos coincidentes en textos o discursos desarrollados en ámbito legal para señalar unas particularidades que donen un carácter especial a la lengua empleada en contextos jurídicos porque:

La linguistique est ici juridique à cause de l'imprégnation du langage par le droit. Elle a dans son objet les interactions du langage et du droit, c'est-à-dire aussi bien l'action du droit sur le langage que l'action du langage dans le droit. Il est essentiel de comprendre, à ce point, que l'étude (linguistique) du langage du droit conduit nécessairement à l'étude (juridique) du droit du langage (Cornu, 2005: 2).

No obstante, como indica Cornu, la estrecha relación entre la función del lenguaje dentro del Derecho y del Derecho dentro del lenguaje deriva en un estudio poliédrico que, además, no es estático porque las leyes se modifican, se derogan y se emanan constantemente. Por tanto, se crean nuevas realidades legales que necesitan una nueva nomenclatura o bien, una revisión de la que ya podía existir con nuevos matices. Por consiguiente, también desde una perspectiva lingüística es necesario hacer una revisión frecuente, sobre todo desde el punto de vista léxico.

Esto nos lleva a nuestra siguiente cuestión: la formación de los futuros traductores jurídicos, ya que estamos hablando de un lenguaje comprendido por un grupo de personas dentro de un sector conformado por tecnicismos y por un carácter arcaico que puede ser considerada una lengua plural desde un punto de vista funcional y dimensional (Cornu, 2005: 17-24).

### 3. LA FORMACIÓN DE TRADUCTORES JURÍDICOS

Hoy en día podemos encontrar distintos cursos para preparar al estudiante que quiere especializarse en traducción dentro del ámbito legal, más allá del propio grado en Traducción e Interpretación como, por ejemplo, el Experto en Traducción e Interpretación Judicial de la Universidad Jaume I<sup>1</sup> o el *Corso di alta formazione in Assistenza Linguistica per l'Ambito Giudiziario* de la Universidad de Bolonia.<sup>2</sup> Estos cursos tienen una estructura común. Por un lado, se introduce al estudiante en el sistema jurídico que corresponda dependiendo de su combinación lingüística y a continuación, se le prepara para la actividad traductora. De este modo, se favorece la adquisición de competencias en ámbito jurídico y traductológico. Queremos recoger, en este punto, las recomendaciones propuestas por Ortega Arjonilla, Martínez López y Alarcón Navío (2017) quienes ponen el foco en que el traductor jurídico debe, principalmente, dominar dos destrezas: la lectora y la escrita. Dentro de estas, enumeran una serie de competencias básicas que necesariamente se deben adquirir.

Con respecto a la comprensión lectora, los autores nos señalan ocho puntos que se deben considerar para alcanzar esta destreza:

- a. entender totalmente los hechos y datos, aunque se presenten en un lenguaje jurídico complejo;
- b. identificar las referencias sociales y culturales en el texto y sus implicaciones para la comprensión total del mismo;
- c. identificar el nivel de formalidad del texto y sus implicaciones para la comprensión total del mismo;
- d. identificar los distintos registros y su razón de ser en el documento, entender totalmente las ideas abstractas y los detalles teóricos, aunque se presenten en un lenguaje jurídico complejo;
- e. identificar los puntos clave de los textos jurídicos;
- f. identificar las diferencias y similitudes existentes entre las ideas y las teorías contrarias;
- g. identificar el tono, el significado y los significados implícitos aparentes del texto;
- h. identificar y aclarar cualquier ambigüedad aparente en el texto, ya sea ésta intencional o no, y, si fuera necesario, pedir asesoramiento mediante consultas a expertos o investigaciones, siguiendo las condiciones de la prueba, e indicar que esto es lo que sería necesario realizar en los encargos de traducciones reales. (Ortega Arjonilla, Martínez López y Alarcón Navío, 2017: 336-337)

Y es que la comprensión lectora es el primer paso para poder realizar adecuadamente nuestra labor traductora. Sin comprender el texto, no podremos trasladar el mensaje a la lengua meta. Por ello,

L'occhiatale del linguista cerca di isolare nei testi giuridici i tratti che li caratterizzano in quanto appartenenti a varietà di lingua distinta nel tempo, nella distribuzione ge-

1 <https://www.fue.uji.es/formacion/EX190094>

2 <https://site.unibo.it/assistenza-linguistica-ambito-giudiziario/it>

ogáfica, nel mezzo di attuazione (scritto o parlato), nei registri; cerca di riconoscere i profili argomentativi, le strutture sintattiche, semantiche e testuali, le implicazioni pragmatiche, sociolinguistiche e via elencando. (Mortara Garavelli, 2001: 4)

El lenguaje jurídico encierra características lingüísticas que dificultan su comprensión cuando no se es lego en la materia y, en ocasiones, incluso a pesar de tener cierta familiaridad con este campo de especialidad. Por todos es conocido el arcaísmo que impregna el discurso jurídico con una sintaxis bastante compleja y caracterizada por párrafos unioracionales repletos de incisos con el fin de ser un texto objetivo pero que, para el lingüística, dificulta su tarea de comprensión.

Before translating, the translator must understand the legal context in which the source text has been written to fully understand the issues involved in the text and its translation. In other words, s/he must (1) understand why, for whom and for what purpose law is being translated in a context, then (2) be able to measure its level of legal complexity, depending on the context. (Monjean-Decaudin y Popineau-Lauvray, 2019: 118)

La buena comprensión del texto de partida nos permitirá traducirlo a la lengua meta, es decir, procederemos a realizar la producción escrita de nuestro encargo. En relación a esta, Ortega Arjonilla, Martínez López y Alarcón Navío (2017: 339) nos detallan seis destrezas que el traductor debería adquirir:

- a. utilizar las estrategias que les permitan lograr el propósito de comunicación con más probabilidades de éxito;
- b. utilizar la lengua, incluyendo el lenguaje especializado y los registros, y los matices para comunicar la información de una manera que sea apropiada para el tema, el contexto y la intencionalidad, teniendo en cuenta las referencias culturales;
- c. utilizar la lengua para transmitir ideas, teorías y opiniones;
- d. expresar el significado claramente y sin ambigüedades;
- e. utilizar términos especializados y específicos de forma apropiada según el contexto jurídico;
- f. identificar y utilizar, cuando sea posible, las fuentes de referencia necesarias e indicarlas de forma adecuada.

Es esencial que el traductor sepa trasladar el mensaje a la lengua meta y por tanto, no se trata de que el estudiante simplemente sepa el funcionamiento de un sistema lingüístico (o de dos), sino que debe comprender el contexto en el que se desarrolla, o ha desarrollado, dicho discurso y conocer todas sus particularidades. De lo contrario, como indican Guerrero Ramos y Carparrós Sánchez (2017: 196):

Si el que trabaja con el texto es un traductor, éste intentará ante todo que su traducción sea comprensible para el lector, que tenga buen estilo y que no contenga vicios lingüísticos tales como repeticiones, cacofonías, calcos, etc. Sin embargo, esta actitud del traductor en el ámbito jurídico es un arma de doble filo, ya que en muchas ocasiones puede caer en errores de contenido por querer quizás dar un toque de elegancia a su

traducción, algo imperdonable en los textos jurídicos cuyo contenido puede ser vital para una organización o un particular.

Estamos de acuerdo con los autores mencionados en este apartado en que no podemos preparar a los estudiantes como futuros traductores jurídicos sin desgranar todos los aspectos que competen a la lengua del Derecho. Recordemos que se tratan de documentos originados en un contexto específico con un efecto legal concreto que puede solicitar un organismo internacional, nacional, empresas u otros clientes (Monjean-Decaudin y Popineau-Lauvray, 2019: 115) por lo tanto, «un nivel elevado de competencia lingüística constituye un requisito básico para los intérpretes y traductores jurídicos, quienes han de ser fiables y precisos en cada detalle y en todo momento» (Ortega Arjonilla, Martínez López y Alarcón Navío, 2017: 312).

#### 4. EL ERROR EN EL AULA DE TRADUCCIÓN

A pesar de que estas destrezas (lectora y producción escrita) son cualidades inherentes a la actividad del traductor independientemente del par de lenguas de trabajo, consideramos que cuando la afinidad lingüística entre ambos idiomas es próxima, dominar estas destrezas se vuelve indispensable porque la similitud entre ambas lenguas puede derivar en cometer más errores tanto desde el punto de vista lingüístico como pragmático o cultural (Hatim, 2013: 180).

No obstante, definir que se entiende como un «error de traducción» es complejo, pues cabría considerar que estamos ante un error cuando difiere de la idea de traducción correcta por parte del docente (Hatim, 2013: 180). De hecho, clasificar los tipos de errores que se pueden dar en un texto traducido no es fácil de establecer. Sin ir más lejos, Battaner Arias y López Ferrero (2019: 428) nos hablan de cuatro tipos de errores: «el error de precisión de una palabra, el error de colocación, el error de registro léxico, el error relacionado con el esquema sintáctico en el que se utiliza la palabra».

Los errores clasificados por Battaner Arias y López Ferrero ponen el foco, principalmente, en la cuestión léxica. No en vano, posiblemente sea uno de los aspectos que más dificultad la tarea de traductor, especialmente en ámbito jurídico:

La terminología jurídica también refleja una fuerte inestabilidad, en contra de lo que muchos opinan, pues dicha terminología depende en gran medida de leyes; y las leyes son inestables, varían con el tiempo, dando lugar a fenómenos como la sinonimia o la polisemia, lo cual provoca que el lenguaje jurídico sea aún más confuso para el profano. Por tanto, la terminología jurídica está en constante evolución pero, a diferencia del vocabulario del lenguaje ordinario, esta evolución es muy lenta y casi no se percibe, (Guerrero Ramos y Caparrós Sánchez, 2017: 194-195)

Por tanto, estamos ante la necesidad de adquirir conocimientos lingüísticos y extralingüísticos pues la falta de estos o el desconocimiento a la hora de emplear las

técnicas de traducción originan muchos errores de traducción (Gile, 2009: 118). Sin embargo, estos no se limitarán al léxico:

È evidente che non si tratta solo di intervenire ragionevolmente sulla terminologia; si tratto soprattutto di controllare la sintassi e l'organizzazione testuale. Se la difficoltà di capire il lessico specialistico è cosa da dare per scontata, non trovano giustificazioni di sorta gli inciampi provocati da vizi e intemperanze sui piani sintattico e testuale: basti pensare all'uso sfrenato di cumuli di subordinate, ghiotti e indigesti esempi di *congeries* infarciti, in aggiunta, di infelici «tecnicismi collaterali». (Mortara Garavelli, 2001: 18)

El lenguaje jurídico italiano y español, por ejemplo, comparten innumerables características morfosintácticas como la presencia de las formas impersonales del verbo, las estructuras unioracionales, la nominalización, las perífrasis de obligación, las construcciones pasivas, entre otros. Sin embargo, la frecuencia de uso de estas características divergirán de una lengua a otra en determinadas ocasiones. Por consiguiente, analizar los posibles errores que hayan podido surgir por la proximidad entre ambas lenguas se convierte en una herramienta esencial para intentar corregir aquellas interferencias interlingüísticas o, incluso, intralingüísticas (Santos Gargallo, 1993). En palabras de Miletich (2008: 453):

No se trata de corregir la traducción del alumno proponiendo la de estudiantes más avanzados, o la del propio profesor, sino de estudiar, colectiva e individualmente, todos aquellos errores, descuidos o desaciertos que hacen que una traducción no posea el nivel o la finalidad deseada.

Se trata de hacer visible una técnica para identificar, aclarar e interpretar de forma sistemática los posibles errores del estudiante al traducir. El análisis de los tipos de errores que se manifiestan en el texto meta, tanto colectiva como individualmente, nos sirve para determinar el paradigma de errores que existen, el origen y las causas de los mismos y su peculiar naturaleza.

A través de esta reflexión, el estudiante puede adquirir nuevas estrategias que le ayuden a solucionar posibles problemas de traducción en su vida profesional o en otras traducciones propuestas en clase, pues no hay una única forma de solventar un problema de traducción y por ello, «no existe una única técnica de traducción a la hora de traducir un determinado segmento textual» (Trovato, 2018: 95). De lo contrario, nos encontraremos ante el riesgo de que el estudiante no se atreva a buscar otras posibilidades para traducir algunas cuestiones y, de este modo:

Ils pourraient alors être tentés de réduire les risques d'erreurs en restant aussi près que possible du texte de départ et en évitant de prendre des décisions audacieuses, pourtant nécessaires dans bien des cas. Cette frilosité les empêcherait de déployer toute leur créativité linguistique, et les priverait d'une partie du plaisir que l'on peut éprouver à chercher, trouver et mettre en œuvre des solutions créatives. Tous ces facteurs ne sont pas favorables à un développement optimal de la compétence traductive. (Gile, 2015: aptdo. 6.1)

Volviendo a la relación entre el italiano y el español, la proximidad de ambas puede provocar que los calcos sintácticos o los falsos amigos deriven en un texto sin sentido. Evidentemente, en cualquier ámbito, si un texto no es coherente puede perder su función. Sin embargo, estos contrasentidos adquieren otra cariz en el ámbito jurídico pues, la gravedad que encierran desencadenan que se pierda valor legal o pueden provocar que se tomen decisiones equivocadas al no contar con información veraz.

## 5. CONCLUSIONES

Indiscutiblemente, la identificación de los problemas de traducción, los posibles errores que se podrían cometer al respecto y las soluciones que se podrían proponer han estado siempre presentes en el aula de traducción. Sin embargo, establecer un criterio que nos dicte qué debemos considerar como un error y en qué medida este afecta a nuestra traducción y la gravedad del mismo no se ha conseguido todavía ya que son muchos factores a tener en cuenta: ¿es más grave el contrasentido que la imprecisión? ¿O quizás lo es la sustitución sinonímica por no contar el término elegido con el matiz que necesita nuestro texto? ¿O puede que el problema sea haber realizado calcos semánticos en nuestro texto meta del documento original?

Cuando estamos ante un documento con contenido jurídico, determinar todo aquello que pueda considerarse un error es más complejo porque estos textos adquieren un valor legal o conllevan unas obligaciones jurídicas. Por tanto, el traductor jurídico debe ser capaz de comprender perfectamente el contenido del texto de origen para, posteriormente, realizar la traducción a la lengua meta. La buena comprensión lectora del documento servirá de auxilio para solucionar problemas de traducción. Asimismo, será necesario que el traductor domine a la perfección la producción escrita pues esta destreza es la parte visible de su labor traductora. Esto implica que los conocimientos que debe adquirir el traductor jurídico no se centran exclusivamente en aspectos lingüísticos ni tan siquiera desde una perspectiva interlingüística puesto que también habrá que considerar los problemas intralingüísticos. Tengamos en cuenta que en el lenguaje jurídico también existe el fenómeno de la polisemia, el cual no hemos tratado en el presente artículo por haber querido presentar un trabajo reflexivo que deberá ser explorado en futuros trabajos de carácter empírico.

Cuando tratamos el error en el aula de traducción como un elemento más de aprendizaje podemos conseguir que los estudiantes creen nuevas estrategias transversales que les ayuden a solventar posibles problemas de traducción en otros encargos académicos o cuando ya entren de lleno en el mercado laboral. De lo contrario, el estudiante intentará no alejarse demasiado del texto de partida a la hora de llevar a cabo su traducción por temor a que, si se aleja, pueda cometer errores. Esto, paradójicamente, puede provocar que el texto traducido contenga

más errores de lo esperado. Consideramos que esto es especialmente así cuando las lenguas con las que se trabaja comparten características lingüísticas similares como en el caso del italiano y el español. Esto se debe a que, a pesar de compartir muchos rasgos, la frecuencia de uso de estas no será la misma así como pueden existir divergencias de uso en determinados contextos. Si favorecemos que el estudiante permanezca lo más “próximo” posible al texto, acabará fosilizando algunas estructuras que, una vez que empiece su recorrido profesional, pueden provocar graves errores de traducción. Sin embargo, el conocimiento de la materia, puede facilitar que el estudiante detecte problemas gramaticales, léxicos, sintácticos por lo que a través del análisis del error podemos proporcionarle herramientas metodológicas para afrontar estos problemas (Giles, 2009: 122).

Por consiguiente, consideramos que sería necesario profundizar más en el análisis del error como herramienta didáctica en el aula de traducción especializada, en este caso jurídica, para que el estudiante se sienta más preparado al enfrentarse a este tipo de traducción tan necesaria y tan compleja a la vez.



- Battaner Arias P. y López Ferrero C. (2019) *Introducción al léxico, componente transversal de la lengua*, Madrid, Cátedra. (ebook)
- Cornu G. (2005) *Linguistique juridique*, 3ª edición, París, Montchrestien.
- De Mauro T. (2017) *Storia linguistica d'Italia dall'Unità a oggi*, Roma-Bari, Laterza.
- Gile D. (2009) *Basic Concepts and Models for Interpreter and Translator Training*, edición revisada. Amsterdam, John Benjamins Publishing Company.
- Gile D. (2015) *La traduction. La comprendre, l'apprendre*, París, Presses Universitaires de France.
- Guerrero Ramos G. y Caparrós Sánchez R. (2017) "Situación actual de la terminología jurídica en España", en *La traducción e interpretación jurídicas en la Unión Europea. Retos para la Europa de los Ciudadanos*. Dir. por Ortega Arjonilla, E., Martínez López, A.B. y Delgado Pugés, I, 2ª edición, Granada, Comares, pp. 191-210.
- Hatim B. (2013) *Teaching and Researching Translation (Applied Linguistics in Action)*, 2ª edición, Londres, Routledge. (ebook)
- Miletich M. (2008) "Pedagogía del error: aspectos didácticos de la traducción", en *Traducción: contacto y contagio: actas del III Congreso "El español, lengua de traducción"*. Dir. por González L. y Hernández P., Puebla, Universidad Autónoma de Puebla México, pp. 453-465.
- Monjean-Decaudin S. y Popineau-Lauvray J. (2019) "How to apply comparative law to legal translation: a new juritraductological approach to the translation of legal texts", en *Research Methods in Legal Translation and Interpreting: Crossing Methodological Boundaries (Law, Language and Communication)*. Dir. por Biel, L., Engberg, J., Martín Ruano, M. y Sosoni, V., Londres, Routledge, pp. 115-129. (ebook)
- Mortara Garavelli B. (2001) *Le parole e la giustizia. Divagazioni grammaticali e retoriche su testi giuridici italiani*, Torino, Piccola Biblioteca Einaudi.
- Ortega Arjonilla E., Martínez López A.B. y Alarcón Navío E. (2017) "Estándares lingüísticos para intérpretes y traductores jurídicos a nivel de grado y de posgrado: traducción-resumen del capítulo 2", en *La traducción e interpretación jurídicas en la Unión Europea. Retos para la Europa de los Ciudadanos*. Dir. por Ortega Arjonilla, E., Martínez López, A.B. y Delgado Pugés, I, 2ª edición, Granada, Comares, pp. 311-352.
- Santos Gargallo I. (1993) *Análisis contrastivo, análisis de errores e interlengua en el marco de la lingüística contrastiva*, Madrid, Síntesis.
- Trovato G. (2018) *Lingüística española y traducción desde la contrastividad*, Roma, Aracne editrice.

# Turismo sostenible y culturemas en español e italiano

GIOVANNA FERRARA  
Università della Basilicata  
[giovanna.ferrara@unibas.it](mailto:giovanna.ferrara@unibas.it)

*Según conocemos una cosa, así la nombramos.*  
Tomás de Aquino, *Summa Theologica*

## ABSTRACT

Tourism was born from the human need to explore, knowing places, people and cultures. The tourist language is heterogeneous, and we can find it on a boundary line among other sectorial languages. The present work is based on an analysis of some aspects in the field of translation and lexicography of the sustainable tourism. Through a qualitative focalisation, we will focus on some linguistic and translational characteristic from a contrastive point of view. The reference points will be some corpus like the “Official Portal of tourism of Spain” (“Portal Oficial de turismo de España”) (<http://www.spain.info/es/>), Cometal, the beginning of monolingual (DRAE and Keys) and bilingual (Laura Tam) dictionaries. We will highlight some culturemas that can give problems at the moment in which they will be translated from Spanish to Italian (alfajores, monas de Pascua, etc.) and from Italian to Spanish (like, for example, “peperoni cruschi”, “Tartelate”, etc.). The goal is to establish some useful instruments for the translation in the professional field of the linguistic and cultural mediation.

The linguistic and cultural mediator transmits intercultural knowledges to ease the linguistic exchange and to avoid misunderstandings during the communication. It neither deals with a simple and mere translation nor transfer and transmit unique and typical knowledges of a specific group that often does not own a conceptual and linguistic equivalent in the given language. For that reason, they come into play a series of translational strategies and techniques to accomplish a grade of equivalence as much accurate as possible.

#### KEYWORDS

Sustainable Tourism, language of tourism, culturemas, strategies of translation, Spanish/Italian translation

#### 1. TURISMO SOSTENIBLE

El turismo es sustentable cuando su desarrollo conserva sus actividades en un tiempo ilimitado sin alterar el ambiente natural, social, artístico presente en un territorio. En la primera Conferencia mundial del turismo sostenible se focalizó la atención en la necesidad de reconocer el impacto de la industria del turismo sobre el ambiente.<sup>1</sup> Los resultados (*Carta del turismo sostenible*) son una garantía y tutela hacia un turismo basado en la sustentabilidad, integración del ambiente natural y cultural, favoreciendo la aplicación concreta del desarrollo sustentable, capaz de responder a las necesidades presentes, respondiendo al sentido original del viaje turístico, es decir descubriendo lugares, entrando en contacto con nuevas realidades en un intercambio mutuo.

Un aspecto importante es la interacción entre la lengua del turismo y su medio ambiente de uso. El discurso turístico encaja perfectamente en los estudios de ecolingüística pues sus recursos léxicos están plasmados por el entorno físico, cultural y social en el que se inscribe. El término ha sido acuñado por Haugen (1972) adelantando la perspectiva de Halliday (1978) y su visión del lenguaje como semiótica social que hace hincapié en las relaciones entre el medio ambiente, las lenguas y el hablante, es decir la manera en que usamos la lengua para comunicarnos que incluye el dónde, cuándo, porqué se usa o no.

El lenguaje turístico es holístico (relación total con la realidad que designa); dinámico (en continua transformación y, por lo tanto, tecnologizado); interactivo y localizado (según el marco espacio-temporal del lugar concreto).

Hablar una lengua es una cuestión cultural y pragmática además de lingüística, ya que somos usuarios de la lengua. Hurtado Albir (2001: 611) habla de “elementos característicos de una cultura presentes en un texto que, por su especificidad pueden provocar problemas de traducción”. Se trata de encaminar hacia

1 Realizada en Lanzarote, Islas Canarias, España, en abril de 1995.

una competencia lingüística y constatar los problemas interculturales que comporta la comunicación en otra lengua (Pérez Vicente, 2012: 630).

Para la Organización Mundial del Turismo (OMT) el turismo abarca más de 600 millones de personas. El estudio del léxico ayuda a captar los argumentos utilizados para potenciar el sector. El turismo es un motor de riqueza en numerosos enclaves y ofrece diferentes actividades.

Una premisa necesaria es la dicotomía con la que se presenta la palabra en español de sostenible y sustentable frente al término sostenible en italiano. Para el DRAE sostenible en ecología y economía es algo que se puede mantener durante largo tiempo sin agotar los recursos o causar grave daño al medio ambiente; mientras que sustentable es algo que se puede mantener sin agotar los recursos. La OMT utiliza el término sostenible para el turismo, como aquel tipo de turismo que satisface las necesidades de los destinatarios y de determinadas localidades receptoras, al mismo tiempo que protege e incrementa oportunidades laborales para el futuro. Holísticamente abarca todos los recursos: económicos, sociales, estéticos, etc., manteniendo su integridad ecológica y cultural (Chierichetti, Garofalo y Mapelli 2019).

El valor de sostenible impregna muchos ámbitos de la vida actual y entre ellos el del turismo, que implica la promoción y protección del medio ambiente (Sanmartín, 2016). Incluye el interés hacia la cultura, las tradiciones, lo autóctono, la sustentabilidad, gracias a su capacidad de transversalidad en diversos sectores económicos y sociales, favoreciendo la preservación de tradiciones locales y tutela de la biodiversidad. Los productos turísticos son únicos y auténticos a la vez y representan una oportunidad de crecimiento y desarrollo territorial.<sup>2</sup> Del mismo modo, la Agenda 2030 para el desarrollo sostenible de la ONU afirma que el turismo es uno de los sectores en donde se pueden lograr los objetivos de sustentabilidad, pues a través de la promoción se crean áreas de trabajo y se incrementa el conocimiento de determinados productos territoriales. Dentro de este marco, el turismo gastronómico promueve la cultura a través de la comida, y favorece el turismo a través de la preservación de un patrimonio cultural e identitario.

## 2. ELEMENTOS CULTURALES

El lenguaje es la condición que hace posible la cultura (Coseriu, Sapir, Whorf). Estamos de acuerdo con Katan (1999) en que la lengua y la cultura son filtros de la percepción de la realidad. Se trata de la relación entre las palabras y la realidad que describen. De modo que la traducción desenvuelve un rol primordial en los estudios culturales: “En tanto que la cultura y el lenguaje generan una distorsión sobre la realidad del mundo, el papel de la traducción es reparar esa distorsión (Molina 2006: 26). Para Calvi (2006: 56) el léxico del turismo:

2 La OMT (2018) apunta hacia el uso de la tecnología como instrumento de promoción y preservación de los productos de un determinado territorio.

Refleja la variedad de los sectores profesionales relacionados con la industria de la hospitalidad y la heterogeneidad de los componentes temáticos implicados en la descripción del producto turístico.

Su riqueza léxica se debe a la complejidad de los subcampos semánticos que conlleva, pues abarca infinitos campos semánticos concernientes diferentes aspectos y realidades (geográficas, históricas, gastronómicas, etc.). Como apunta Calvi (2006: 17) intervienen también otros ámbitos semánticos (historia del arte, artesanía, etc.) y el trasvase de elementos culturales representa el mayor escollo en la traducción (Newmark). Los elementos culturales influyen en la expresión de nuestros pensamientos y sentimientos. Estos

mantienen un comportamiento dinámico en el proceso de traducción en consonancia con los principios de funcionalidad y dinamismo que rigen la equivalencia traductora y que por tanto la manera en la que se resuelve su transferencia depende del conjunto de circunstancias y propósito que rodea el texto traducido” (Molina 2006: 14).

Nida (1945), Newmark (1995), Vlachov (1970), Koller (1992), Bödeker (1987), Vermeer, Molina, Hurtado Albir, House y Nord entre otros, representan los estudios y aportaciones más significativos por lo que concierne el tratamiento de los elementos culturales en la traducción. A su vez, el proceso de traducción de textos turísticos es un campo ampliamente estudiado por Calvi (2006), Pérez Vicente (2012), López González (2017), Castellano Martínez (2018) entre otros.

Los conceptos culturales que conciernen la geografía, historia, folklore, gastronomía, etc., reciben diferentes denominaciones: palabras culturales, términos culturales, elementos culturales, *realia* o culturema (Calvi 2006: 36).

El estudio de las dificultades culturales es uno de los problemas cruciales en la traducción. Newmark habla de “foco cultural” refiriéndose a la profusión de términos de un determinado campo léxico en una lengua. Nida es el primero en hablar de elementos culturales como punto focal de la traducción. Para Vlachov y Florin se trata de *realia*, es decir esos elementos textuales que expresan color local e histórico, distinguiendo 4 categorías: geográficos y etnográficos; folklóricos y mitológicos; objetos cotidianos y sociales e históricos. House (1977)<sup>3</sup> por su parte habla de traducción patentada y encubierta en función de la distancia cultural de ambas lenguas y propone un “filtro cultural” para que la traducción mantenga la misma función que el término de partida. De esta forma se superan las dificultades.

### 2.1. CULTUREMAS

Si bien Vermeer da el concepto de culturema, basándose en los estudios de *Kulturtheorie*, es Molina la que presenta la definición más completa (2006: 79):

3 Citado en Molina 2006.

Elemento verbal o paraverbal que posee una carga cultural específica en una cultura y que al entrar en contacto con otra cultura a través de la traducción puede provocar un problema de índole cultural entre los textos origen y meta.

Ahondando en la dimensión dinámica de la palabra, lo más relevante es que “los culturemas no existen fuera de contexto, surgen en el seno de una transferencia cultural entre 2 culturas” (Molina: 78). Este es, por supuesto, uno de los problemas que tiene que enfrentar el traductor, y para ello se hace necesario elaborar una serie de estrategias que aseguren la transferencia de su auténtico significado a otra lengua y, por ende, a otra cultura.

El culturema representa una realidad sociocultural que no siempre posee un equivalente en otra lengua (López Santiago, 2016: 82)<sup>4</sup> y su inadecuada traducción puede generar malentendidos o transmisión errónea de información.

Por motivos de espacio y tiempo, se ha decidido limitar el corpus del trabajo a determinados términos del área temática gastronómica española e italiana (y en lo específico a la de la región de la Basilicata), pues estos últimos aspectos se engloban en un proyecto mucho más amplio del turismo sostenible que se está llevando a cabo dentro del ámbito de la *Notte della ricerca Europea 2021: Superscienceme - Research is your Re-Generation* (European Researchers' Night).

La gastronomía tiene un peso relevante dentro del lenguaje turístico, en lo que Gamero Pérez (2001) denomina foco exhortativo, pues a través de la exposición de contenidos (gastronómicos) incita al destinatario a cumplir una acción. Hay que considerar, además, el aspecto relevante multimodal,<sup>5</sup> pues casi siempre está acompañado por imágenes que ayudan y persuaden al lector, formando muchas veces un “todo indivisible” denotando además un “sabor local” (Pérez Vicente: 630). En la traducción de los culturemas la integración de imágenes es un refuerzo para ayudar al destinatario de texto a descodificar el verdadero significado de una entrada, eliminando casos de ambigüedades y malentendidos.

Los culturemas aportan mucho más que el simple y mero lenguaje. Nos limitaremos a observar cuales mecanismos interculturales se activan en la traducción de dichos términos y cuáles pueden ser las dificultades de comprensión. Veremos pues como las posibles técnicas de traducción son siguiendo a Hurtado Albir (2001) y Molina (2006):

- traducción literal, en donde se privilegia la cultura extranjera, opacando de alguna manera el término;
- uso de equivalentes, que favorecen la comprensión, pero pueden modificar el significado, perdiendo a su vez el “sabor local”;

4 López Santiago habla, por ejemplo, de la procesión de los “Facundillos”, que son los niños que portan la imagen del Niño Jesús; o también de la Romería, la peregrinación dedicada a un Santo.

5 Las páginas web presentan un lenguaje multimodal (audiovisual) pues presentando diferentes tipos de textos informa y persuade a la vez.

- paráfrasis, en donde la explicación puede extender el significante alterando la estructura del texto.

Como en cualquier tipo de traducción, todo dependerá del *skopos* de la misma. La función es informativa, apelativa y exhortativa, son textos híbridos o multifuncionales (Hatim y Mason 1995).

El fin es posibilitar de forma inmediata la comunicación en la cultura final “lo que induce a optar por una traducción comunicativa y a sustituir las convenciones lingüísticas y de estructuración textual de la cultura de partida por las convenciones propias de la cultura final” (Reiss y Vermeer 1996: 171).

### 3. CULTUREMAS ESPAÑOLES Y SU TRADUCCIÓN

La relación entre la gastronomía y la cultura hace que se tenga que “recurrir a verdaderas explicaciones que arrojan un poco de luz para un lector extranjero, y más teniendo en cuenta que en muchas ocasiones ni los propios nativos de un país conocen determinados platos de una zona muy concreta de la que desconocen sus exquisiteces culinarias” (Ponce Márquez 2006: 12).

El campo gastronómico está lleno de culturemas. Tomaremos como corpus de estudio, desde un punto de vista cualitativo la página del Portal oficial de turismo de España ([spain.info/es/](http://spain.info/es/)). Los términos a los que vamos a pasar revista son los siguientes: Alfajores, flores de Semana Santa, leche frita, monas de Pascua, polvorones, roscas de vino. Los diccionarios utilizados son el DRAE y el Clave con sus versiones en línea y el diccionario bilingüe de Laura Tam.

Por otra parte, el corpus multilingüe del turismo de la universitat de València COMETVAL (2016) da indicaciones para evitar los errores de traducción de los culturemas, sugiriendo algunas estrategias a seguir, que son los ya evidenciados por Hurtado Albir (2001: 614): transcripción, calco, formación de una nueva palabra, asimilación cultural, traducción aproximada, préstamo, neutralización, traducción literal, etc.

Alfajor: Clave nos da la siguiente definición: Dulce hecho con una pasta de almendras, nueces, miel, pan rallado y tostado u otros ingredientes: En Navidad solemos tomar polvorones y alfajores. El diccionario bilingüe lo traduce como “panforte, pan pepato, ciambelline di panforte”. En la versión italiana de la página oficial de turismo español, encontramos la receta de los “alfajores andalusí” y de los “alfajores di Murcia”, pero sus ingredientes (latte, zucchero, farina, cannella, buccia di un limone) no corresponden a los de la receta original. Sin embargo, en la descripción detallada de los platos típicos de la gastronomía andaluza refiere que son dulces de derivación árabe hechos con almendras:

La pasticceria andalusa rivela le radici arabe nell'uso delle mandorle e del miele. La tradizione si è mantenuta nei conventi, che offrono yemas di San Leandro, [...], alfajo-

res, mantecados, polvorones, ciambelline al vino, ecc., sono invece dolci di produzione quasi sempre industriale che meritano di essere conosciuti.<sup>6</sup>

Flores de Semana Santa: en el Drae se encuentra como “dulce que imita en su forma a la flor, hecho con huevos, leche y harina, que se fríe en aceite y se rocía con azúcar o miel”. La página web presenta el término original y una posible traducción entre paréntesis como “sfoglie croccanti” lo cual no basta para entender el producto, de modo que una traducción posible sería: sfoglie croccanti dolci a forma di fiori, tipiche della Settimana Santa e del Carnevale. La traducción ofrecida por el diccionario bilingüe es “frittelle di semolino al latte”. El portal del turismo en cambio no traduce el término, sino que lo deja tal cual indicando que se trata de un “dolce típico del Nord della Spagna”. Sin embargo, habla de las “flores de cuaresma” y las describe como:

Una croccante e delicata massa frita a forma di fiore, decorata con zucchero e cioccolato (nonostante si trovino in tutto il Paese, sono tipiche della Castiglia - La Mancia, Castiglia e León, Estremadura e Galizia).<sup>7</sup>

Leche frita: El Drae nos dice que se trata de un dulce de harina cocida con leche y azúcar, y frita; mientras que para Clave es un dulce elaborado con una masa de leche y harina que se fríe rebozada. Para el diccionario bilingüe son: “frittelle di semolino al latte”. En el portal del turismo, encontramos la leche frita entre los dulces típicos de Semana Santa: “Nel nord della Spagna, in queste occasioni è molto diffusa la leche frita”.<sup>8</sup>

Mona de Pascua: Para el Drae es un hornazo (rosca con huevos) que es costumbre comer, en algunos pueblos, en la Pascua de Resurrección. Para Clave es un pastel con figuras de chocolate que se hace en las fechas próximas a la fiesta cristiana de Pascua de Resurrección. El diccionario bilingüe traduce como “ciambella pasquale”. Una traducción posible sería “torta pasqualina, (focaccia o ciambella guarnita con uova sode)”.

Es interesante indicar que la página web no traduce el término, pero efectúa una ampliación añadiendo al original “Il dolce del Lunedì dell’Angelo” y le dedica un párrafo entero:

È il simbolo della pasticceria di Pasqua in quasi tutta la costa mediterranea della Spagna. Presenta due varianti molto popolari. Da una parte, la più tradizionale nella Comunità Valenciana e nella Regione di Murcia, ovvero una pasta di brioche dolce sormontata da uova sode. Dall’altro lato, la più famosa in Catalogna, con uova di cioccolato o curiose figure di cioccolato che decorano il dolce. Infatti, le vetrine di molte

6 Véase <https://www.spain.info/it/cucina/cucina-andalusia/>

7 Véase <https://www.spain.info/it/scoprire-spagna/dolci-ricette-pasqua/>

8 Véase <https://www.spain.info/it/scoprire-spagna/dolci-ricette-pasqua/>



pasticerie si trasformano in veri e propri musei del cioccolato, dove vengono esposte preparazioni sorprendenti.

Questo dolce può essere gustato nel corso di tutta la festa soprattutto nelle Baleari, in Catalogna, nella Comunità Valenciana e nella Regione di Murcia. La tradición, però, vuole che si consumi il Lunedì dell'Angelo. La "mona" rappresenta la fine dell'astinenza dal cibo che accompagna la festa religiosa e ricorre diverse volte nei pasti di famiglia, essendo abitudine dei nonni o dei padrini regalarla a nipoti o figliocci.<sup>9</sup>

Como si no bastara, para atraer más la atención del destinatario, y siempre manteniendo la multifuncionalidad de texto turístico, la descripción está acompañada por la imagen del producto.

Polvorones: el Drae nos dice que se trata de una torta, comúnmente pequeña, de harina, manteca y azúcar, cocida en horno fuerte y que se deshace en polvo al comerla. Para Clave es un dulce elaborado con manteca, harina y azúcar, generalmente de forma redondeada, que se deshace fácilmente al comerlo. En el diccionario bilingüe encontramos la entrada polvorón como "torta soffice". El término no aparece traducido en la página web, sino que está entre comillas, que es una forma utilizada en la traducción para designar que se trata de un elemento "desconocido" en la cultura de llegada, y que forma parte de los dulces navideños. En realidad sabemos que se trata de "biscotti natalizi di pasta frolla, dolce di farina, sugna e zucchero che, appena messo in bocca si riduce in polvere".

Roscas de vino: El Drae dice simplemente que es un pan o bollo en forma de rosca. Clave refiere que todo lo que tiene forma ovalada y deja un espacio vacío en el centro como por ejemplo "una rosca de pan". Para el diccionario bilingüe es una ciambella.

### 3.1. RESULTADOS

En la página web de referencia Portal oficial de turismo de España se encuentran muchas inadecuaciones en la comprensión de los términos casi siempre no traducidos al italiano y acompañados muchas veces por un texto descriptivo que indica cómo elaborar el producto en cuestión. Por lo tanto, se trata de una información aparentemente completa y por supuesto de una traducción opaca.

En palabras de Hurtado Albir (2001: 615) existe "una multiplicidad de soluciones y de técnicas (adaptaciones, paráfrasis, notas, elisiones, etc.) en función del contacto entre dos lenguas". Una solución posible es dejar al culturema en español y añadir una paráfrasis explicativa. Sin embargo, las técnicas son producto de la elección del traductor, pues es siempre funcional según "el tipo textual, la modalidad de traducción, la finalidad de traducción y el método elegido" (2001: 642). En fin, tenemos que:

9 Véase <https://www.spain.info/it/scoprire-spagna/dolci-ricette-pasqua/>

Considerar el tipo de relación entre las dos culturas; el género cultural en el que se inserta; la función del culturema en el texto original; la naturaleza del culturema; las características del destinatario y la finalidad de la traducción (2001: 614-615).

Gorozhankina (2017: 93) afirma que se “establecen una multitud de técnicas y soluciones posibles”. En el ámbito turístico sería deseable unificar los criterios para mejorar la calidad de la traducción; por ejemplo, traducir o no los nombres de las ciudades, de los santos, utilizar el mismo equivalente, etc. Por último, el culturema, cuando su traducción sea “imposible” debería ir acompañado por una explicación o ampliación, conservando su forma original como valor de arraigo de promoción turística. López Santiago (2016: p. 80) opta por mantener el nombre en la lengua original seguido de una traducción o breve explicación (es el caso de caña, tapas, patatas bravas, sangría, paella, etc.).

#### 4. HACIA UN ESBOZO DE CULTUREMAS ITALIANOS Y SU TRADUCCIÓN

En este campo la labor por hacer es muy amplia, de modo que se pasará simplemente revista a una serie de términos, cuyo estudio se llevará a cabo en una investigación posterior al presente trabajo y que se enmarca dentro del proyecto *Noche Europea della Ricerca 2021*. Se trata de una serie de palabras representativas en las que fácilmente se puede incurrir en el error léxico y que se encuentran en un corpus de páginas web que tienen que ver con la promoción del turismo en la región Basilicata. Baste citar como ejemplos los siguientes términos y las posibles traducciones presentadas a continuación.

Peperoni cruschi: Representa uno de los productos más conocidos de la región Basilicata. Una posible traducción podría ser “pimientos crujientes”. En la traducción española encontramos las siguientes informaciones:

El *peperone crusco* (pimiento crujiente en castellano), en plural *peperoni cruschi* y también conocido como pimiento *crusco*, es un producto típico de la gastronomía lucana. Es un pimiento de color rojo intenso, de forma cónica y de sabor dulce. Se cuelga y se seca al aire libre. La variedad más indicada es el *Peperone de Senise*, que ha recibido la indicación geográfica protegida (IGP) en 1996, por su pulpa fina y reducido contenido de agua, cualidades que lo hacen apto para el secado. El *peperone crusco* tiene múltiples usos en la cocina local, hasta el punto de ser apodado ‘el oro rojo de Basilicata’. Generalmente se corta y se priva de semillas para ser frito unos segundos en aceite de oliva. La exposición a altas temperaturas hace que el pimiento se vuelva crujiente (*crusco*). Se puede comer como snack y guarnición. También se muele en polvo y se usa como especia para aliñar carnes, productos de panadería, chocolate y helado.<sup>10</sup>

10 [https://es.wikipedia.org/wiki/Peperone\\_\\_crusco](https://es.wikipedia.org/wiki/Peperone__crusco)

Rafanata: De *rafano* “rábano” encontramos la siguiente descripción: “Tortilla al horno aromatizada con rábano picante, además de patatas y pecorino. Una variante implica la adición de salchicha”.<sup>11</sup>

Cartellate; aunque no se encuentre la entrada en el vocabulario, encontramos una descripción exhaustiva de este culturema:

Las *cartellate* (en dialecto apuliano *carteddàte*), también llamadas *nèvole* y *rose*, o *crispedde* en Basilicata, son junto con los *bocconotti* unos dulces típicos navideños originarios de la provincia de Foggia pero producidos en toda Apulia y la Basilicata. Su nombre deriva del hecho de que la masa debe ser muy delgada, como “la carta oleata” (un tipo de papel de cocina que se utilizaba antiguamente, parecido al moderno papel para horno). Las *cartellate* no son más que trozos de una hoja delgada de masa, elaborada con harina, aceite y vino blanco, unida y envuelta sobre sí misma para formar una especie de ‘rosa’, con alternancia de huecos y aberturas, que luego se fríen en abundante aceite.

La receta regional típica es la que va impregnada en *vincotto* caliente o miel, y espolvoreada luego con canela, azúcar glas o adornos de colores. Sin embargo, hay variantes: por ejemplo, puede sustituirse el *vincotto* por chocolate o simplemente por azúcar glas. Una vez preparadas se mantienen alejadas de la luz y pueden comerse durante varias semanas.<sup>12</sup>

Coppa: No presente en el diccionario con el significado de embutido/insaccato. La única información que encontramos por el momento es la siguiente:

Se trata un ‘salame’ (fiambre italiano) hecho de la cabeza de la paletilla o del cuello de cerdo curado entero. Es parecido a un jamón curado más conocido, el prosciutto, siendo ambos fiambres de cerdo que se usan en recetas parecidas. Sin embargo, el prosciutto es pierna de cerdo (o jabalí), mientras la *capicola* es únicamente carne de la paletilla (hombro) o el cuello.<sup>13</sup>

Scamorza: En el diccionario bilingüe encontramos la siguiente definición: queso fresco de vaca y cabra. Se trata de un queso semiduro, similar a la mozzarella.

Taralli: El diccionario nos dice que se trata de: rosquilla con pimienta o anís. El producto tiene muchas variantes según la zona de producción. Básicamente es una rosquilla con levadura al horno. La masa base se hace con harina, agua, aceite y sal:

Un producto horneado típico de Apulia, Basilicata, Campania y Calabria clasificado por el Ministerio de Agricultura italiano como producto agroalimentario tradicional. Sin embargo, se produce en otras regiones de Italia, como por ejemplo Basilicata, donde

11 [https://es.wikipedia.org/wiki/Gastronom%C3%ADa\\_de\\_Basilicata](https://es.wikipedia.org/wiki/Gastronom%C3%ADa_de_Basilicata)

12 <https://es.wikipedia.org/wiki/Cartellate>

13 [https://es.wikipedia.org/wiki/Gastronom%C3%ADa\\_de\\_Basilicata](https://es.wikipedia.org/wiki/Gastronom%C3%ADa_de_Basilicata)

existe una variante (el *tarallo aviglianese*) hecha pasta de azúcar fundida y aromatizada con anís, que le permite tener un color blanco nieve.<sup>14</sup>

Vemos pues que se trata de un proceso de descodificación y codificación en otra lengua que exige al traductor una competencia elevada de las dos lenguas y culturas. De no cumplirse estos requisitos se cometerán errores que alterarán la comprensión del término.

## 5. CONCLUSIONES

Casi siempre hay una transferencia del culturema, pues el usuario necesita conocerlo para poder utilizarlo. No obstante, encontramos algunos tipos de traducción. La traducción literal es la estrategia preferida para esos términos aparentemente transparentes (roscas de vino, o taralli). En el equivalente cultural o funcional muchas veces ayuda la imagen del producto. Por último, cuando no se encuentra un equivalente se recurre casi siempre a la paráfrasis. En fin “no hay soluciones unívocas ni técnicas características para la traducción de los culturemas” (Hurtado Albir 2001: 615).

En el ámbito del turismo es imprescindible prestar atención a todos los aspectos relacionados con el proceso y fin de la traslación de modo que se cumplan todos los objetivos del texto original, es decir la promoción turística a través del uso de redes sociales con un lenguaje multimodal y funcional integrador de significados, para mejorar la calidad e imagen ofrecida.

14 <https://es.wikipedia.org/wiki/Tarallo>

- Calvi M. V. (2006) *Lengua y comunicación en el español del turismo*, Madrid, Arco/Libros.
- Calvi M. V. (2010) “Los géneros discursivos en la lengua del turismo: una propuesta de clasificación”, en *Ibérica*, 19/2, pp. 9-32.
- Casado Velarde M. (1988) *Lenguaje y cultura*, Madrid, Síntesis.
- Castellano Martínez J. M. (2018) “Análisis de la traducción de textos turísticos promocionales (francés-español): tres ciudades de la Occitania”, en *Gran Tour: Revista de Investigaciones Turísticas*, 18, <https://dialnet.unirioja.es/servlet/articulo?codigo=6984423>, pp. 19-39.
- Coseriu E. (1981) *Lecciones de lingüística general*, Madrid, Gredos.
- Chierichetti L., Garofalo G. y Mapelli G., (2019) (eds.) *Hacia una visión holística del discurso turístico*, *Círculo de Lingüística Aplicada a la Comunicación*, 78, pp. 3-16, <http://webs.ucm.es/chierichetti.pdf>, <http://dx.doi.org/10.5209/CLAC.64368>
- Gamero Pérez S. (2001) *La Traducción de textos técnicos*, Barcelona, Ariel.
- Gorozhankina T. (2017) *Estudio comparativo de la traducción de textos turísticos en español y ruso. Parámetros de calidad y análisis de corpus*, Tesis doctoral, <https://dialnet.unirioja.es/servlet/tesis?codigo=110484>
- Haugen E. (1972) *The Ecology of Language: Essays*, Stanford, Stanford University Press.
- Halliday M. A. K. (1978) *Language as social semiotic. The social interpretation of language and meaning*, London, Edward Arnold.
- Hatim B., Mason I. (1995) *Teoría de la traducción. Una aproximación al discurso*, Barcelona, Ariel.
- Hurtado Albir A. (2001) *Traducción y traductología*, Madrid, Cátedra.
- Katan D. (1999) *Translating Cultures. An Introduction for Translators, Interpreters and Mediators*, Manchester, St. Jerome.
- López González R. C. (2017) “Traducir aspectos culturales para turistas y cruceristas: Vigo, un estudio de caso”, en *Glaucoptis, Boletín del Instituto de Estudios Vigueses*, 22, <https://www.ievigueses.com/storage/publicacions/July2018/mOZoINJHFH8P2srumF.pdf>, pp. 355-372.
- López Santiago M. (2016) “La traducción: aspectos léxicos”, en Artusi A. (coord.), *Guía de Buenas Prácticas Discursivas: Estrategias lingüísticas para la promoción turística en la red*. Normas. *Revista de Estudios Lingüísticos Hispánicos*, Anejo n. 8, pp. 73-91.
- López Santiago M. (2019) “El léxico del turismo religioso: perspectiva contrastiva español-francés”, en Chierichetti L., Garofalo G. y Mapelli G. (eds.), *Hacia una visión holística del discurso turístico*, *Círculo de Lingüística Aplicada a la Comunicación*, 78, <http://webs.ucm.es/info/circulo/78/lopez.pdf>, <http://dx.doi.org/10.5209/CLAC.64370>, pp. 33-48
- Newmark P. (1995) *Manual de traducción*, Madrid, Cátedra.
- Nord C. (2009) “El funcionalismo en la enseñanza de traducción”, en *Mutatis Mutandis, Revista Latinoamericana de Traducción*, 2 (2), pp 209-243.
- Molina L. (2006) *El otoño del pingüino. Análisis descriptivo de la traducción de los culturemas*, Castelló de la Plana, Publicacions de la Universitat Jaume I.
- Pérez Vicente N. (2012) “La traducción de los culturemas en el lenguaje del turismo y su aplicación a la dialéctica del español”, en Cassol A. et alii (eds.), *Metalinguaggi e metatesti. Lingua, letteratura e traduzione*,

XXIV Congreso AISPI, [https://cvc.cervantes.es/literatura/aispi/pdf/23/23\\_\\_629.pdf](https://cvc.cervantes.es/literatura/aispi/pdf/23/23__629.pdf), pp. 629-639.

Ponce Márquez N. (2006) “Metodología en la clase de traducción: Dificultades pragmático-culturales en la traducción de textos turísticos”, en *Tonos Digital*, n. 12, <http://www.tonosdigital.es/ojs/index.php/tonos/article/view/48/47>

Reiss K., Vermeer H. J. (1996) *Fundamentos para una teoría funcional de la traducción*, Madrid, Akal.

Rodríguez Abella R. (2008) “La traducción de los culturemas en el ámbito de la gastronomía (análisis de los folletos de Turespaña)”, en *Rivista Internazionale di Tecnica della Traduzione*, 10, pp.47-69

Sanmartín Sáez J. (2016) “Sobre neología y contextos de uso: análisis pragmalingüístico de lo ecológico y de lo sostenible en normativas y páginas web de promoción turística”, en *Iberica* 31, pp. 175-198.

Sapir E. (1974) *El lenguaje y el medio ambiente. Antología de estudios de Etnolingüística y Sociolingüística*, México, UNAM.

Vermeer H. J. (1983) “Translation theory and linguistics”, en Roinila P., Orfanos R. y Tirkkonen-Conditt S. (eds.), *Näkökohtia käännämisen tutkimuksesta*, Joensuu, University of Joensuu, pp. 1-10.

Vlakhov S., Sider F. (1970) “Neperevodimoe v perevode: realii”, en *Masterstvo perevoda*, Moscú, Sovetskii pisatel, pp. 432-456.

Whorf B. L. (1971) *Lenguaje, pensamiento y realidad*, Barcelona, Seix Barral.

#### WEBGRAFÍA

Clave. *Diccionario de uso del español actual*, SM (ed.), <http://clave.smdiccionarios.com/app.php>

*Diccionario de la Real Academia Española de la lengua*, Real Academia Española, <http://www.rae.es>

Tam L., *Grande Dizionario spagnolo/italiano, italiano/spagnolo*, Hoepli, [https://www.grandidizionari.it/dizionario\\_\\_italiano-spagnolo.aspx](https://www.grandidizionari.it/dizionario__italiano-spagnolo.aspx)

Portal oficial de turismo de España, <https://www.spain.info/es/>  
<https://www.basilicaturistica.it/>  
[https://es.wikipedia.org/wiki/Gastronom%C3%ADa\\_de\\_Basilicata](https://es.wikipedia.org/wiki/Gastronom%C3%ADa_de_Basilicata)



# Studenti di traduzione vs. studenti di lingua: chi traduce meglio? Uno studio greco-italiano

GIACOMO KLEIN  
Università di Trieste  
gklein@units.it

## ABSTRACT

This paper aims at comparing the outcomes of a Greek-Italian translation task between translation students and foreign language students. This study focuses both on a error-analysis in the target texts on different levels. Students at the Advanced School for Interpreters and Translators of the University of Trieste and students attending Greek courses at the Hellenic Foundation for Culture were asked to translate a short informational text using both paper and online dictionaries. The translated texts were then corrected and mistakes in meaning, morphology, lexicon and syntax were pointed out and compared using both a quantitative and a qualitative approach. Results suggest that translation students produced more accurate translations, mainly as far as semantics and lexicon are concerned.

## KEYWORDS

Translation studies, modern Greek, linguistics, Greek Linguistics, translation errors



## 1. INTRODUZIONE

La valutazione della qualità della traduzione è uno degli aspetti più dibattuti nei *translation studies* ed è stata oggetto di numerosi studi sia da un punto di vista accademico sia da un'ottica prettamente professionale (Orlando, 2016). La letteratura ha dimostrato che non è mai un singolo fattore a rendere buona o meno una traduzione. Hatim, per esempio, scrive riguardo all'individuazione degli errori di traduzione:

(...) errors may be ascribed to a diverse variety of causes, ranging from lack of comprehension to misuse of register – factors which may be located on a variety of levels (language, pragmatics, culture). (Hatim 2001: 180)

A definire un errore di traduzione sono un insieme di caratteristiche che variano in base all'equivalenza tra testo di partenza (TP) e testo di arrivo (TA), all'adempimento dello skopos traduttivo ma che dipendono anche da elementi culturali o di percezione individuale (cfr. Gyde Hansen, 2008). Ovviamente questi fattori variano anche tra le diverse tipologie testuali e contesti comunicativi di varia natura. Tra i *translation studies* si incontrano inoltre svariate analisi di cosiddetti errori di traduzione. Magris sottolinea che è da considerarsi errore di traduzione

qualsiasi fattore che incida negativamente sull'efficacia comunicativa della traduzione, ovvero sulla trasposizione delle intenzioni comunicative dell'autore, sia sull'effetto esercitato dal testo sul suo lettore (Magris, 2005: 15).

La letteratura si è poi concentrata sulla classificazione degli errori di traduzione in base alla loro tipologia, cioè errori lessicali, grammaticali, sintattici ecc. (cfr. Stolze, 1997; Hansen, 2009). Un'ulteriore distinzione è quella fra errori di senso ed errori di forma: in genere vengono considerati più gravi i primi in quanto minano la comprensione da parte del lettore (cfr. C. Nord, 1996).

Numerosi studi si concentrano sugli errori commessi dagli studenti di traduzione (cfr. Muñoz 2012, Klepikova 2018, Abbasi&Karimnia 2018). Uno dei pochi studi comparativi su questo tema è il lavoro di Orlando (2016), che presenta un focus particolare sull'ambito giuridico. Da questo studio emerge che gli studenti di traduzione, anche se non specializzati nel campo giuridico, hanno prodotto traduzioni “migliori” rispetto a giuristi con un'ottima conoscenza della L2.

Tuttavia, da un'analisi della bibliografia emerge che non sono stati spesso paragonati gli errori di traduzione degli studenti di traduzione con quelli degli studenti di lingua straniera. Un approccio di questo tipo permetterebbe di identificare possibili abilità e competenze che contraddistinguono gli studenti di traduzione rispetto a chi impara una L2 per interesse personale e senza specializzarsi nella disciplina della traduzione.

Al presente studio hanno partecipato da una parte studenti del corso di lingua e traduzione neogreca della Sezione di Studi in Lingue Moderne per Interpreti e Traduttori (SSLMIT) dell'Università di Trieste e dall'altra alunni di lingua neogreca della Fondazione Ellenica di Cultura.

La lingua neogreca fa parte dell'offerta formativa della SSLMIT come lingua a scelta studente. I partecipanti allo studio hanno portato a termine il corso di livello base e intermedio e due anni di dottorato per un totale di circa 260 ore di lezione. Il livello minimo di conoscenza linguistica dei partecipanti è il livello B1 del Quadro di riferimento europeo per la conoscenza delle lingue. Oltre allo studio della grammatica, gli studenti della SSLMIT si sono cimentati nell'esercizio della traduzione attiva e passiva, soprattutto per quanto concerne tipologie testuali come la scrittura su web, siti turistici, descrizioni commerciali di aziende o di prodotti di vario genere. Inoltre, i grecisti della SSLMIT hanno partecipato ogni anno a una collaborazione con un'azienda o un ente greco per il quale hanno tradotto contenuti web di vario tipo (cfr. Klein 2018; Klein 2019). Ovviamente, oltre alla formazione da traduttori da e verso il neogreco, i partecipanti della SSLMIT hanno anche usufruito di tutti gli apprendimenti nelle loro altre lingue di studio come anche di un inquadramento teorico e competenze informatiche legate all'ambito della traduzione.

D'altro canto, il control group formato dagli studenti di lingua neogreca presso la sede italiana della Fondazione Ellenica di Cultura presenta un background di formazione estremamente diverso e variegato. Tutti i partecipanti sono adulti che studiano il neogreco per passione e interesse personale: non possono dunque contare su nessuna formazione traduttologica. Anche in questo caso si è scelto di selezionare studenti che avessero raggiunto almeno il livello B1: questo accorgimento metodologico è mirato a creare due gruppi di partecipanti paragonabili tra loro dal punto di vista del livello di apprendimento linguistico della L2.

Il testo utilizzato per questo studio è tratto dalla banca dati del KEG - Centro della Lingua Greca di Salonicco (<https://www.greek-language.gr/certification/dbs/teachers/index.html>) e corrisponde a un livello di difficoltà B1. Si tratta di un testo informativo senza particolari difficoltà lessicali, sintattiche o morfologiche.

*Η ιστορία των παπουτσιών είναι τόσο παλιά, όσο και τα βήματα των ανθρώπων στη γη. Τα παπούτσια έγραψαν τη δική τους ξεχωριστή ιστορία. Τα χαμηλά σανδάλια, οι ψηλές μπότες, τα πέδιλα με κορδόνια και πολύτιμες πέτρες δε δείχνουν μόνο τη μόδα. Τα παπούτσια είναι ένα στοιχείο για τον πολιτισμό κάθε εποχής. Πριν από εκατομμύρια χρόνια, οι άνθρωποι στην Ασία κάλυπταν τους αστραγάλους τους με δέρμα για να μην κρυώνουν και χτυπούνε τα πόδια τους στις πέτρες. Αργότερα, στην Αρχαία Αίγυπτο φορούσαν παντόφλες από φύλλα φοίνικα και πάπυρο. Η ποικιλία των παπουτσιών γίνεται ακόμη πλουσιότερη στην Αρχαία Ελλάδα. Οι γυναίκες φορούν μαλακά παπούτσια μέσα στο σπίτι. Έχουν χρώμα λευκό και κόκκινο και τα δένουν με λεπτά σχοινιά μέχρι τα γόνατα. Οι Ρωμαίοι πάλι προτιμούσαν τα σανδάλια που άφηναν γυμνά τα δάχτυλα των ποδιών. Χρησιμοποιούσαν, όμως, και παπούτσια από δέρμα και ξύλο. Τον 15ο αιώνα τα παπούτσια ήταν πολύ στενά*

και οι μύτες τους ήταν πολύ μεγάλες και γύριζαν προς τα πάνω. Ο Βασιλιάς Εδουάρδος έδωσε εντολή οι μύτες των παπουτσιών να μην ξεπερνούν τα 5 εκατοστά! Ο Βασιλιάς Λουδοβίκος πάλι, φορούσε παπούτσια με πολύ ψηλό τακούνι, επειδή είχε πρόβλημα με το ύψος του. Οι δερμάτινες μπότες, τα μποτάκια από ύφασμα, οι γόβες από πλαστικό, όλα τα παπούτσια βρίσκουν μία θέση στην παπουτσοθήκη, αλλά και στη ζωή μας.

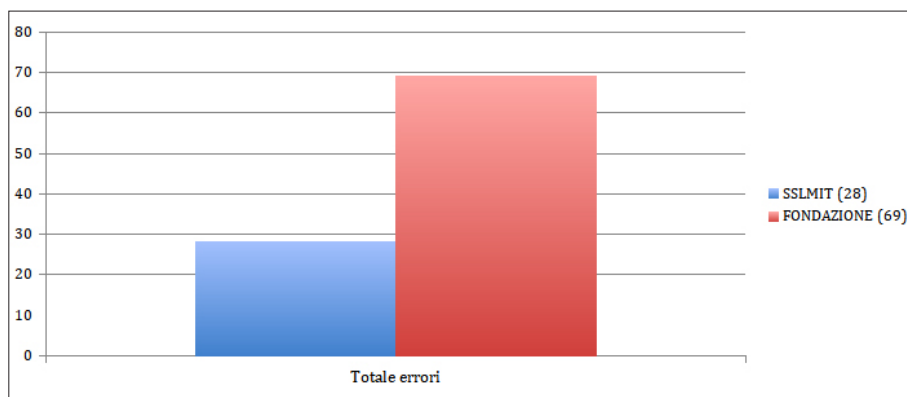
Hanno partecipato allo studio 7 studenti della SSLMIT e sette della Fondazione Ellenica di Cultura che hanno potuto usufruire di un massimo di 60 minuti di tempo per portare a termine la traduzione. Lo scopo di questo studio è stato verificare le abilità traduttive dei partecipanti in una situazione simil-lavorativa. Pertanto è stato consentito l'uso di dizionari sia cartacei sia online. Per quanto concerne la correzione dei TA, sono stati presi in considerazione le seguenti tipologie di errori:

- a. Errori di senso cioè laddove il TA non riporta genericamente il messaggio del TP
- b. Errori di lessico cioè soluzioni problematiche dal punto di vista del registro, della tipologia testuale o, in generale, soluzioni che ostacolano la lettura del TA perché improprie o inadeguate.
- c. Errori di morfologia come, per esempio, soluzioni inadeguate nella scelta dei tempi verbali tra TP e TA
- d. Errori di sintassi: problemi di ordine degli elementi del periodo e scelte di punteggiatura o testualità che ostacolano la leggibilità dell'elaborato.

### 3. ANALISI DEGLI ELABORATI ED ERRORI DI TRADUZIONE

#### 3.1. NUMERO TOTALE DI ERRORI

Per quanto concerne i problemi traduttivi, vale innanzitutto la pena considerare il numero totale di errori commessi dai partecipanti. Le annotazioni di tutte le categorie di errori commessi sono state sommate al fine di capire chi tra i due gruppi di partecipanti sbaglia di più e chi invece traduce “meglio”.



Il numero totale di errori, 28 commessi da parte degli studenti della SSLMIT e 69 da quella della Fondazione Ellenica di Cultura, è stato analizzato con un test t. Questa analisi statistica è mirata a paragonare le medie di due gruppi al fine di capire se sussiste una differenza reale oppure casuale.<sup>1</sup> Il test t ha messo in luce *p-value*<sup>2</sup> inferiore a 0,05 (=0,0202) il quale indica una differenza statisticamente rilevante tra il numero di errori totali commessi dai due gruppi di partecipanti allo studio.

### 3.2. ERRORI DI SENSO

Gli errori di senso sono, secondo l'*American Translators Association*, quegli errori che “compromettono o distorcono la comprensione di fatti o concetti nel testo di partenza.<sup>3</sup> Sostanzialmente, l'errore di senso mina la comprensione del TP a causa di problematiche di omissioni, aggiunte, terminologia o forme verbali. Nel caso del nostro studio, gli studenti della Fondazione Ellenica di Cultura hanno prodotto un numero tre volte maggiore (12) di errori di senso nelle loro traduzioni rispetto agli studenti della SSLMIT (4). Un esempio di errore di senso che è stato individuato in vari elaborati riguarda la seguente frase: *Οι Ρωμαίοι πάλι προτιμούσαν τα σανδάλια που αφήναν γυμνά τα δάχτυλα των ποδιών.* In questo caso, l'avverbio *πάλι* ha il significato di *d'altro canto* o *d'altra parte*, tuttavia molti partecipanti hanno preferito avverbi o locuzioni avverbiali come le seguenti:

- a. Pure i Romani preferivano scarpe che lasciavano scoperte le dita.
- b. I Romani nuovamente preferivano i sandali che lasciavano nude le dita dei piedi.
- c. I romani ancora una volta preferivano i sandali che lasciavano scoperte le dita dei piedi.

Un altro periodo che ha causato non pochi problemi ai partecipanti è il seguente: *Τον 15ο αιώνα τα παπούτσια ήταν πολύ στενά και οι μύτες τους ήταν πολύ μεγάλες και γύριζαν προς τα πάνω.* In italiano questa frase andrebbe tradotta come segue: *Nel XV secolo le scarpe erano molto strette e le punte molto grandi e rivolte all'insù.* Alcuni degli errori che sono stati rinvenuti nei TA sono i seguenti:

- a. Nel XV secolo le scarpe erano molto strette e le dita dei piedi molto larghe e rivolte verso l'alto.
- b. Nel 15° secolo le scarpe erano molto strette e le loro punte erano molto grandi e rivolte all'indietro.

1 [https://www.treccani.it/enciclopedia/test-t\\_%28Dizionario-di-Economia-e-Finanza%29/#:~:text=T%2C%20test%20Test%20\(%E2%9E%94\),%2C%20distribuzione\)%20con%20varianza%20ignota](https://www.treccani.it/enciclopedia/test-t_%28Dizionario-di-Economia-e-Finanza%29/#:~:text=T%2C%20test%20Test%20(%E2%9E%94),%2C%20distribuzione)%20con%20varianza%20ignota).

2 [https://www.treccani.it/enciclopedia/p-value\\_%28Enciclopedia-della-Matematica%29/](https://www.treccani.it/enciclopedia/p-value_%28Enciclopedia-della-Matematica%29/)

3 <https://www.atanet.org/certification/how-the-exam-is-graded/error-categories/>  
Traduzione a cura dell'autore dell'articolo

Nel primo caso si nota un errore lessicale che si trasforma in errore di senso: sono state confuse le punte delle scarpe con le dita dei piedi. Questa scelta è probabilmente dovuta al fatto che la parola greca *μύτες* può significare sia punta dei piedi che punta delle scarpe. Nel secondo, invece, siamo di fronte a un errore semantico poiché il traduttore ha confuso *πάνω* (sopra) con *πίσω* (dietro).

Nei TA sono stati individuati anche alcuni casi di overtranslation, cioè di aggiunte ingiustificate e che non sono presenti nel TA.

Both over-translation and under-translation consist in reproducing in the target language (TL) the non-equivalence of the source language (SL) message, in terms of meaning or style. Over-translation refers to the information that the target language contains is more than that of the source language (Wang, 2012: 130)

Queste soluzioni sono state adottate per aggiungere informazioni al TA rendendolo più specifico: in alcuni casi questo funziona mentre in altri ci si allontana dal messaggio del TP. Per esempio:

- a. *Αργότερα, στην Αρχαία Αίγυπτο φορούσαν παντόφλες από φύλλα φοίνικα και πάπυρο. Più tardi, nell'antico Egitto, portavano babbucce ricavate da foglie di palma e di papiro.*
- b. *Ο Βασιλιάς Εδουάρδος έδωσε εντολή οι μύτες των παπουτσιών να μην ξεπερνούν τα 5 εκατοστά! Il re Edoardo (addirittura) emise un editto (affinché) le punte delle scarpe non superassero i cinque centimetri!*

Mentre nel primo caso la parola *babbuccia* rimanda a una calzatura morbida di origine orientale e può essere un'integrazione fondata al TA, il secondo esempio, parlando di un *editto*, modifica il senso del TP tramite un'aggiunta ingiustificata. La scelta lessicale *editto* rappresenta, tra l'altro, anche un problema di registro e un errore semantico poiché il greco *εντολή* non è niente più che un semplice ordine.

### 3.3. ERRORI LESSICALI

Per quanto riguarda gli errori lessicali, la maggior parte dei problemi riscontrati sono scelte lessicali goffe, poco chiare o inadeguate al contesto. Gli alunni della Fondazione Ellenica di Cultura hanno commesso 5 volte più inesattezze lessicali rispetto agli studenti SSLMIT.

*Milioni di anni fa, gli uomini in Asia coprivano le loro caviglie con la pelle per non sentire freddo e per non battere i piedi sulle pietre.*

In questo primo esempio vediamo come l'uso della parola *pelle* si sia rivelata una scelta lessicale inappropriata: non viene specificato di che tipo di pelle o di calzatura si tratta. In questo caso sarebbe stata auspicabile un'aggiunta da parte del

traduttore che avrebbe potuto parlare di *calzature in pelle* al fine di rendere più intelligibile il messaggio del TP.

Vale poi la pena spendere qualche parola su una frase del TP che ha messo in estrema difficoltà i partecipanti allo studio e che ha portato a varie soluzioni traduttive discutibili. Τα παπούτσια έγραψαν τη δική τους ξεχωριστή ιστορία è una collocazione fissa greca che tradotta letteralmente significa: le scarpe hanno scritto la loro unica storia. Ovviamente era necessario riformulare questo concetto che risulta del resto facilmente comprensibile al lettore/traduttore madrelingua italiano.

*Le calzature hanno scritto la loro storia particolare*  
*Le scarpe hanno scritto una loro storia peculiare*  
*Le scarpe hanno scritto la loro storia separatamente*

In alcuni casi l'aggettivo ξεχωριστός ha portato anche a errori di senso, tuttavia sono stati molti i partecipanti che non hanno saputo "italianizzare" questa frase. Una collocazione italiana paragonabile sarebbe stata *le scarpe hanno una storia tutta loro*.

Dai risultati dello studio sembra dunque trasparire che, anche per quanto riguarda gli aspetti lessicali, una formazione nel campo della traduzione garantisce una maggior sensibilità da parte del traduttore che riesce a capire quando staccarsi dal TP al fine di rendere fruibile al lettore il TA.

#### 3.4. ERRORI DI SINTASSI E TESTUALITÀ

Per quanto riguarda i problemi di sintassi e testualità, i due gruppi di partecipanti allo studio hanno commesso circa lo stesso numero di errori: questa è l'unica tipologia di errori in cui si registra un certo equilibrio tra traduttori specializzati e studenti di lingua. Trattandosi di una traduzione passiva, non si registrano gravi errori nel TA in quanto tutti i partecipanti allo studio sono di madrelingua italiana. Tuttavia, sono stati riscontrate alcune scelte traduttive in cui la testualità, vale a dire il periodare del TA poteva essere getito diversamente.

Il TP presenta periodi molto brevi e un ritmo quasi spezzettato, soprattutto nella sua parte iniziale e centrale. Il traduttore avrebbe potuto agilmente unire questi periodi senza appesantire il TA, facilitando in questo modo la lettura. Un esempio può essere rappresentato dal seguente periodo: se nel primo caso il partecipante ha deciso di seguire la sintassi del TP ricorrendo a una ripetizione del soggetto all'inizio dell'ultimo periodo, la seconda traduzione presenta un periodo unico che non altera il significato della frase e garantisce una lettura più scorrevole.

*Η ιστορία των παπουτσιών είναι τόσο παλιά, όσο και τα βήματα των ανθρώπων στη γη. Τα παπούτσια έγραψαν τη δική τους ξεχωριστή ιστορία. Τα χαμηλά σανδάλια, οι ψηλές μπότες,*

τα πέδιλα με κορδόνια και πολύτιμες πέτρες δε δείχνουν μόνο τη μόδα. Τα παπούτσια είναι ένα στοιχείο για τον πολιτισμό κάθε εποχής.

- a. *La storia delle scarpe è antica quanto i passi dell'uomo sulla terra. Le scarpe hanno scritto la propria particolare storia. I sandali bassi, gli stivali alti, le scarpe col tacco con laccetti e pietre preziose non mostrano solo la moda. Le scarpe sono un elemento della cultura di ogni epoca.*
- b. *Le calzature hanno scritto una propria storia unica, antica quasi quanto i primi passi dell'uomo. I sandali bassi, gli stivali con il tacco, i sandali con lacci impreziositi da pietre non sono semplicemente un risultato della moda, ma rappresentano un elemento importante della cultura di ogni epoca.*

Come già anticipato, in questa sezione non si notano differenze quantitative tra i due gruppi e bisogna tenere in considerazione anche la possibilità e la facoltà del traduttore di restare più fedele al TP, anche per quanto riguarda la sintassi dei periodi e la struttura generale del testo. Questo non rappresenterebbe infatti un errore ma una scelta del traduttore, la quale tuttavia deve tenere in considerazione la tipologia testuale. In questo caso ci troviamo di fronte a un testo informativo che deve presentare il suo messaggio in maniera chiara al lettore, pertanto il mantenimento dello stile dell'autore, anche per quanto riguarda la sintassi, non rappresenta una priorità.

### 3.5. ERRORI DI MORFOLOGIA

Anche in questo caso non ci sono stati veri e propri errori, in quanto l'esercizio di traduzione proposto era dal greco all'italiano. Non sono pertanto stati commessi errori di concordanze o di coniugazione verbale ma si sono riscontrate alcune incongruenze nel TA per quanto concerne l'uso dei tempi verbali. Il TP alterna per tutta la sua durata il tempo aoristo, l'imperfetto e il presente storico. Questa scelta, assolutamente corretta e usuale in greco, non è tuttavia trasferibile all'italiano senza creare qualche ostacolo alla prima lettura del testo tradotto. Molti partecipanti hanno scelto di seguire e si trasferire i tempi verbali del TP nel TA.

*Milioni di anni fa, le persone in Asia si coprivano le caviglie con la pelle per evitare che si raffreddassero e che sbattessero i piedi sulle pietre. Più tardi, nell'antico Egitto, indossavano pantofole fatte di foglie di palma e papiro. La varietà di scarpe diventa ancora più ricca nell'antica Grecia. Le donne indossano scarpe morbide al chiuso. Sono bianche e rosse e sono legate con corde sottili fino alle ginocchia. I romani ancora una volta preferivano i sandali che lasciavano scoperte le dita dei piedi. Ma usavano anche scarpe di cuoio e legno. Nel XV secolo le scarpe erano molto strette e le dita dei piedi molto larghe e rivolte verso l'alto. Re Edoardo ordinò che le dita dei piedi delle scarpe non superassero i 5 cm!*

*Mille anni fa gli uomini in Asia nascondevano le caviglie con la pelle per non prendere freddo e non colpire i piedi con i sassi. Più tardi, nell'antico Egitto indossavano ciabatte di foglie di palma e papiro. La varietà delle scarpe diventa ancora più ricca nell'antica Grecia. Le donne indossano ciabatte morbide in casa. Sono di colore bianco e rosso e vengono legate con cordicelle*

*fino alle ginocchia. I Romani nuovamente preferivano i sandali che lasciavano nude le dita dei piedi. Utilizzavano, però, scarpe di pelle e legno. Nel 15° secolo le scarpe erano molto grandi e puntavano verso l'alto. Il Re Edoardo diede ordine che le punte delle scarpe non superassero i cinque centimetri!*

In questi elaborati notiamo un'alternanza di tempi verbali (imperfetto, presente storico, passato remoto) identica a quella proposta dal TP e che nel TA italiano risulta poco naturale e ostacola parzialmente la lettura. La seguente proposta si distacca invece dal testo di partenza e propone una resa che rende più chiaro il susseguirsi di eventi narrati dal TP e permette una lettura più agevole.

*Migliaia di anni fa, gli uomini in Asia coprivano le caviglie con delle pelli per non prendere freddo e non sbattere i piedi sulle pietre. In seguito, nell'antico Egitto si portavano ciabatte fatte con foglie di palma e papiro. La varietà di scarpe è diventata ancora più ampia nell'antica Grecia. Le donne portavano calzature morbide in casa, bianche e rosse, e le legavano con corde sottili fino alle ginocchia. I romani preferivano sandali che lasciavano scoperte le dita dei piedi, ma portavano anche scarpe fatte di pelle e legno. Nel 15esimo secolo le scarpe erano molto strette, con delle lunghe punte rivolte verso l'alto. Il re Edoardo ordinò che le punte delle scarpe non superassero i 5 centimetri!*

Anche per quanto riguarda la morfologia si registrano più errori di traduzione commessi dagli studenti senza una formazione nel campo della traduzione. In questa, come in tutte le altre tipologie di errori analizzate, traspare una maggiore abilità da parte degli studenti della SSLMIT nel mettere il lettore al centro delle loro scelte traduttive. Sembra che chi studia traduzione traduca per gli altri mentre chi traduce per esercizio non riesca ad avere un approccio *target-oriented*.

#### 4. CONCLUSIONI E LIMITI

I due gruppi di partecipanti hanno prodotto elaborati qualitativamente molto diversi. Questo viene confermato dal numero di errori totale che varia molto tra chi ha goduto di una formazione da traduttore e chi invece studia il greco solamente come lingua straniera. In quasi tutte le categorie di errori analizzati si è riscontrata questa netta discrepanza anche se le differenze più marcate riguardano le scelte lessicali e gli errori di senso in cui le soluzioni adottate dagli studenti della SSLMIT sono nettamente migliori e più accurate rispetto a quelle proposte dagli alunni della Fondazione Ellenica di Cultura. Questi risultati sembrano suggerire che studiare traduzione e studiare lingue straniere siano due cose estremamente diverse e che la pratica della traduzione necessiti di un inquadramento ben preciso, sia dal punto di vista teorico sia da quello pratico.

Questo studio presenta due limiti che non trascurabili: il numero di partecipanti e la loro età. Un numero maggiore di testi raccolti garantirebbe sicuramente una maggiore rappresentatività dei risultati che potrebbero, in quel caso, anche venir analizzati con ulteriori analisi quantitative e statistiche. L'età media



degli studenti della Fondazione (51 anni) è ovviamente di molto superiore a quella degli studenti in corso e da poco laureati alla SSLMIT. Questi limiti verranno presi in considerazione per ricerche future che utilizzeranno questi spunti al fine di produrre un'analisi contrastiva ancora più dettagliata.

- Abbasi M. & Karimnia A. (2011) "An analysis of grammatical errors among Iranian translation students: Insights from interlanguage theory", in *European Journal of Social Sciences*, 25(4), pp. 525-536.
- Hansen G. (2009) "A classification of errors in translation and revision". In *CIUTI-Forum 2008: Enhancing Translation Quality: Ways, Means, Methods*, Bern, Peter Lang, pp. 313-326.
- Hansen G. (2008) "The speck in your brother's eye—the beam in your own". In Hansen, G., Chesterman, A e Gerzymisch-Arbogast, H. (eds.), *Efforts and models in interpreting and translation research*, Amsterdam, John Benjamins, pp. 255-280.
- Hatim B. A. (2014) *Teaching and researching translation*, Londra, Routledge.
- Klein G. & Kourouni K. (2018) "Language and translation of Greek tourism websites: a corpus-based study", in *Rivista Internazionale di Tecnica della Traduzione*, 20, pp. 87-99.
- Klein G. (2019) "Lingua e traduzione neogreca: un'esperienza di insegnamento alla SSLMIT", in *Rivista Internazionale di Tecnica della Traduzione*, 21, pp. 157-170.
- Magris M. (2005) *L'errore in traduzione. Dalla teoria alla pratica*, Udine, Edizioni goliardiche.
- Muñoz I. D. (2012) "Analysing common mistakes in translations of tourist texts (Spanish, English and German)", in *Onomazein*, 26, pp. 335-349.
- Orlando D. (2016) *The Trials of Legal Translation Competence: Triangulating Processes and Products of Translators vs. Lawyers*, Tesi di Dottorato non pubblicata.
- Séguinot C. (1990) "Interpreting errors in translation", in *Meta*, 35(1), pp. 68-73.
- Stolze R. (1997) *Bewertungskriterien für Übersetzungen-Praxis, Didaktik, Qualitätsmanagement. Translationsdidaktik-Grundfragen der Übersetzungswissenschaft*, Tübingen, Narr, pp. 593-602.
- Wang M. (2012) "An analysis of over-translation and under-translation in perspective of cultural connotation", in *Lecture Notes in Information Technology*, 16, pp. 129-133.



# El prefijo *s-* en italiano y su exhaustiva correspondencia en la traducción al español

ROCÍO LUQUE  
Università di Trieste  
[rluque@units.it](mailto:rluque@units.it)

## ABSTRACT

In this article the use of the Italian prefix *s-* is analyzed from the morphosyntactic and semantic point of view, in order to know the main equivalences of translation in Spanish in the various derivations. On being two Latin languages with different cognitive development along the time, we have found various contrastive aspects, that begin in the origins of the tongues and finish in the actual use. Among these aspects that define the different equivalences of this prefix *s-*, we must remark the origin, the cognitive competence of each culture and the high functionality of Italian prefixes.

## KEYWORDS

Derivation, semantics, culture, cognition, contrastivity

## 1. INTRODUCCIÓN

Este estudio de la afijación en la lengua, y concretamente, como es nuestro caso, el de la prefijación,<sup>1</sup> no debe considerarse distinto al de cualquier elemento compositivo, ya que los prefijos presentan una gran productividad tanto en el léxico común (alrededor del 50% de los verbos derivan por prefijación) como en registros cultos y en el lenguaje especializado (Iacobini 2004: 99), sin olvidar tampoco su presencia en lo que respecta a la creación de nuevos neologismos.<sup>2</sup> Su característica principal es la de precisar el significado del lexema al que acompañan,<sup>3</sup> como podemos ver con los prefijos negativos italianos *in-*, *s-*, *dis-* y *a-*, de bases adjetivas, donde la elección de la forma adecuada queda determinada por la tipología del adjetivo. Así, con *in-*, observamos un uso general por aplicarse a una gran variedad de adjetivos como *impossibile*, *incapace*, *incoerente* e *irresponsabile*;<sup>4</sup> con la forma *s-*, por el contrario, debemos pensar en un uso limitado a adjetivos que se refieren principalmente al trato social, como *scontento*, *scortese*, *sgradevole* o *sleale*, y derivados de participios de verbos parasintéticos de la primera conjugación, como *sconfinato*, *sconfortato* y *spezzato*, entre otros muchos; *dis-*, por su parte, lo encontramos aplicado a adjetivos en los que existe la idea de separación, como tenemos en *disattento* o *discontinuo*, y en formaciones derivadas de los participios de pasado de verbos parasintéticos de la primera y segunda conjugación, como *disabitato*, *disintossicato*, *dispiaciuto* o *distorto*; finalmente, *a-*, con valor negativo, lo localizamos por línea general en bases cultas,<sup>5</sup> entre las cuales citamos *acefalo*, *analfabeto* y *apolítico*.

Nuestro objetivo en este trabajo es ahondar en el uso del prefijo italiano *s-*, concretamente con bases verbales, para analizar la variedad de soluciones presentes en la traducción en español, considerando en todo momento que las nuevas formaciones de base verbal y prefijo presentan una polisemia más compleja que la de su origen latino. Pensamos en primer lugar en las coincidencias semánticas con “de-”/“des-” y “dis-” y posteriormente en el distinto uso que diacrónicamente estos han ido desarrollando por tratarse de culturas diversas, en las que las conceptualizaciones han cambiado a lo largo del tiempo según las situaciones pragmáticas en las que aparecían aplicadas estas nuevas formaciones a través de los prefijos.

1 Véase Rainer (1991).

2 Si partimos del italiano, aparte de los clásicos *mono-*, *multi-*, *poli-*, etc., son numerosos los prefijos que aparecen en su formación, como son los casos de *ri-*, *in-*, *s-*, entre otros muchos.

3 Véase Scalise e Zannier (1983).

4 En la formación de *irresponsabile*, además, se observa la asimilación del prefijo *in-* al sonido de la palabra que lo sigue (*in- + r > irr-*).

5 Véase Dardano (1978: 128).

Cuanto más conocemos una determinada cultura, mayor serán los contrastes que percibamos con respecto a la propia. Quizás esta sea la clave para entender la archiconocida afirmación de que italiano y español son dos leguas afines, es decir, existe una falta de conocimiento mutuo entre dos lenguas que han experimentado una evolución distinta a lo largo del tiempo. Este hecho podemos refrendarlo a través de la distinta dimensión pragmática en las que observamos la diversidad de uso de los tiempos verbales, como cuando en italiano decimos *Potresti farmi un favore?*, con el empleo del condicional, que en español encuentra comúnmente su correspondencia en presente como forma más directa preguntando “¿Puedes hacerme un favor?”; con el uso muy hipotético y restringido del subjuntivo en *Fai come se fossi a casa tua*, contrastando con el indicativo, de nuevo más directo, diciendo “Haz lo que quieras, estás en tu casa”; o con la utilización del pretérito perfecto en italiano para referirnos a un hecho pasado en el tiempo, como en *La scuola ha vinto il primo premio nel 2015*, cuyo correspondiente en español se construye con el pretérito indefinido como tenemos en su equivalente “La escuela ganó el primer premio en 2015”. Asimismo, hallamos contrastes en los términos que son típicos de una cultura, como son los casos de los adverbios *proprio* o *assolutamente*, en *È stato proprio lui* y *Siamo assolutamente d'accordo*, con sus correspondientes “Ha sido él” y “Estamos completamente de acuerdo”; o en las complejidad de formas rutinarias<sup>6</sup> que distinguen a cada lengua y que son el resultado de la diversidad de construcciones sociales a la que da lugar nuestra mente, como vemos en frases de exhortación como *Dai, vieni con noi alla festa* o de imposibilidad en *Non ci resta che accettare questa proposta*, donde traduciríamos respectivamente “Venga, ven con nosotros a la fiesta” y “No tenemos más remedio que aceptar esta propuesta”.

Dichos contrastes entre ambas lenguas constituyen el desarrollo de relaciones no intuitivas que, a la postre, tienen lugar en espacios diferentes,<sup>7</sup> con pensamiento, experiencias y sentidos bien diversos. Se trata de frases ligadas a una cultura, en las que la palabra “traducción” es simplemente un término que de forma convencional utilizamos en lugar de decir cómo se expresa esa idea en la otra lengua, o cómo la interpreta el otro pensamiento, ya que se trata realmente de entender los mecanismos mentales de una cultura y de saber si la lengua meta posee las construcciones que expresen esa idea. Pensemos en la fórmula italiana de carácter negativo *neanche ci penso*, que, situada en un contexto como *Andare allo stadio con questo freddo, neanche ci penso!*, podemos expresar como “¡ni se me ocurre!”; o en la que también forma parte del lenguaje cotidiano con valor desiderativo *non vedo l'ora*, que, en *Non vedo l'ora di incontrarti!*, equivaldría a “Tengo unas ganas locas de verte”, donde observamos que mientras que el italiano se sirve respectivamente del pensamiento y del tiempo, el español se apoya en la

6 Seguimos la clasificación de G. Corpas (1996: 270-271).

7 Véase Porzig (1986).

ocurrencia y en el deseo; o, por el contrario, en la española “por la cuenta que me trae”, como riesgo, que solo podremos traducir si la colocamos en un contexto del tipo “Por la cuenta que me trae, tengo que devolverle el libro a Carmen”, que en italiano expresaríamos como *Devo assolutamente restituire il libro a Carmen*, donde observamos una clara diferencia de pensamiento, ya que si en la mente española está presente el riesgo como “perjuicio”, utilizando el sentido figurado de la economía, en italiano para presentar esta situación predomina la idea de obligación; o en “No me trae cuenta”, donde el mismo núcleo en “Comprar un coche usado en esas condiciones no me trae cuenta” aparecerá traducido con la idea de interés como *Non mi interessa comprare una macchina in quelle condizioni*. De ahí que deduzcamos las diferencias existentes en la representación mental de objetos, simples situaciones o acontecimientos entre culturas distintas, por lo que se trata en realidad de entender los mecanismos combinatorios del lenguaje en una cultura (Pinker 2000).

### 3. CONTRASTES EN LA PREFIJACIÓN

La alta frecuencia de *ri-* nos puede servir de muestra para estudiar el distinto funcionamiento de algunos prefijos en ambas lenguas. Efectivamente, *ri-* se relaciona con una mayor tendencia al énfasis expresivo en la realización de las acciones (Luque Toro 2009: 62) o, como señala Plag, con el modo en el que hay que contrapesar la economía de almacenamiento y la de procesamiento en el lexicón mental para lograr la máxima funcionalidad (2003:49). La expresividad del italiano en *ri-* se presenta en usos en los que sería innecesario por la presencia de cuantificadores, como es el caso de *richiamare* en *Ti ho richiamato cinque volte senza ottenere una risposta*, que en español correspondería a “Te he llamado cinco veces sin ninguna respuesta”; mientras que hablaremos del sentido funcional cuando *ri-* está presente con verbos como *ridare*, *rientrare*, *rimandare*, etc., cuyo uso significa entender la acción del verbo desde una óptica distinta según el valor del prefijo. Por lo tanto, si con *rientrare*, en la frase *Ho pensato che rientrerò dopo le otto*, entendemos que marcamos el hecho de “entrar de nuevo”, con “volver”, en “He pensado que volveré después de las ocho”, nos centramos en la vuelta, pero no precisamos si nos referimos al punto de partida o al destino.

La contrastividad en el uso de la prefijación, pues, alcanza una compleja dimensión tanto morfológica como semántica entre estas dos lenguas, que desde nuestro punto de vista podemos analizar, tomando como base el italiano, del siguiente modo:

- Contraste en la prefijación. El uso de uno u otro prefijo ha ido cambiando diacrónicamente y un prefijo en italiano puede no tener su correspondiente en español: *arrabbiato* / “enfadado”, *ingigantire* / “agigantar”, *inoccupato* / “desocupado”, *inoltrare* / “reenviar”, *ultrapiatto* / “extraligero”.

- Contraste con la ausencia de prefijos. Quizás sea el caso más frecuente debido a la tendencia del italiano al sentido relacional de estos morfemas: *incivilire* / “civilizar”, *incoraggiamento* / “estímulo”, *rientrare* / “volver”.
- Uso de paráfrasis en ausencia de prefijos. Nos referimos a formaciones como las siguientes: *infarinare* / “cubrir de harina”, *sparecchiare* / “quitar la mesa”, *spolverare* / “quitar el polvo”.
- La amplia polisemia de las palabras prefijadas. Desde el punto de vista de la traducción, tenemos que señalar el valor polisémico de las palabras prefijadas. Así, con la palabra *rimandare* tenemos, entre otros, los siguientes significados: “volver a mandar”, “devolver”, “aplazar”, “remitir”, “despedir”.<sup>8</sup>
- La formación de numerosos falsos amigos por las relaciones que definen estos prefijos. Señalamos por su asimetría los siguientes: *incassare* / “encajar”, *rimuovere* / “remover”, *scomposto* / “descompuesto”, *stendere* / “extender”, etc.

#### 4. SOBRE LA PRESENCIA Y LOS VALORES DEL PREFIJO S-

El origen latino del prefijo *ex-*,<sup>9</sup> del que deriva *s-*, nos indica la idea de ‘fuera’ o ‘más allá’, lo que está situado en el exterior, por lo que hemos de interpretar cognitivamente un alto índice de negatividad, entendiendo esta desde el punto de vista de lo contrario o de lo que se aparta de algo, física o figurativamente. De ahí que como valores principales<sup>10</sup> citemos el de intensidad, en verbos como *scambiare*,<sup>11</sup> *spremere* o *sterminare*, y el privativo, en palabras como *eccentrico*,<sup>12</sup> *scorrevole* o *sparecchiare*, donde se marca el contrario del término o su inversión o reversión sin prefijo. Señala igualmente De Mauro su valor despectivo con bases sustantivas (2000: 2263), como son los casos de *sfiducia*, *sfortuna*, *sproporzione* o *sventura*.

Este prefijo, como otros muchos, además de los cambios semánticos que implica la prefijación, posee frecuentemente la característica de cambiar la sin-

8 Véase Arqués y Padoan (2012: 2443).

9 Indicamos en contraste con el español, donde se mantiene este origen, la transformación en *s-* característica del italiano. Entre otros casos, citamos: *esalare* / “exalar”, *esanime* / “exánime”, *stendere* / “extender”, etc.

10 Véase Guilbert (1975).

11 Esta formación, a diferencia de *cambiare* donde se hace referencia a la sustitución o a la conversión principalmente, presenta matices que entendemos que son igualmente el resultado del valor intensivo del prefijo, como la confusión en *Al guardarropa mi hanno scambiato la giacca* (“En el guardarropa me han cambiado la chaqueta”), el intercambio comercial *Sono riuscito a scambiare i dollari in euro* (“He conseguido cambiar los dólares en euros”) o el intercambio de opiniones *Nel dibattito gli opinionisti scambiavano le loro impressioni sull'arte* (“En el debate los opinionistas intercambiaban sus impresiones sobre el arte”).

12 Este es un caso de transformación del prefijo en *ec-*.



taxis<sup>13</sup> de las acepciones de algunos verbos, transformándolas de inacusativas en acusativas, o viceversa. A título de ejemplo, pensamos en el uso intransitivo de *campare*<sup>14</sup> en frases como *Rimarrà con i suoi genitori finché campa* (“Vivirá con sus padres mientras viva”) o *Continua a campare di rendita* (“Sigue viviendo de la renta”), mientras que *scampare* en su función transitiva nos aparece traducido como “evitar” en *Ha scampato la prigione perché ha trovato un buon avvocato* (“Ha evitado la cárcel porque ha encontrado un buen abogado”). Del mismo modo, la acepción acusativa de *fumare* en *Fuma la pipa tutti i giorni* (“Fuma la pipa todos los días”) se convierte, con *sfumare*, en inacusativa en frases como *Il nostro progetto è sfumato* (“Nuestro proyecto se ha frustrado”) o *Le nostre speranze sono sfumate* (“Nuestras esperanzas se han desvanecido”); o el uso transitivo de *frigere*, en *Mia mamma frigge delle patate a volte* (“A veces mi mamá fríe patatas”), se transforma en intransitivo en *Il burro sfrigge anche a fuoco lento* (“La mantequilla chisporrotea también con el fuego lento”).

Recordamos también que este prefijo, debido a su alta frecuencia, ha llegado a sustituir a formas que ahora se consideran en desuso. Citamos como ejemplo el verbo *sprofondare*, en lugar de *profondare*,<sup>15</sup> en construcciones que presentan un uso transitivo, como en *Il terremoto ha sprofondato il paese* (“El terremoto ha sepultado el pueblo”) (Arquésy Padoan 2012: 2622), e intransitivo, como podemos ver en *La barca della baia sprofondò dopo qualche giorno* (“La barca de la bahía se hundió después de algunos días”); o bien formas típicas de un determinado lenguaje especializado, como puede ser, entre otros, el término marinero *largare*, que encontramos con el sentido de «allontanare un'imbarcazione dalla banchina, dalla costa e sim.» (De Mauro 2000: 1340), mientras que la forma prefijada nos aparece con frecuencia en construcciones como *slargare un paesaggio* (“ampliar un paisaje”) o *slargare uno spazio* (“ensanchar un espacio”), donde la presencia de *s-* hay que interpretarla como un indicador intensivo del valor original de la base.

A nivel lexicográfico debemos señalar que los verbos a los que se les agrega *s-* presentan una fraseología en función del nuevo significado de la base verbal. Este es el caso de los verbos citados anteriormente, donde observamos que *campare* da lugar a formaciones como *campare alla giornata*, *campare alla meglio*, *campare cent'anni* o *campare d'aria*, en las que la idea de “vivir” es el núcleo generador, en cambio, con *scampare* prevalece la idea de “evitar” o “salvar”, como tenemos en las construcciones pronominales *samparla* o *scamparla bella*.

13 Pensamos igualmente en los numerosos casos de pronominalización.

14 Su uso transitivo resulta obsoleto o se limita a lenguajes especializados, como tenemos en *Campa una figura sulla tela* (De Mauro 2000: 365).

15 Señala De Mauro el bajo uso de la forma transitiva (2000: 1981).

## 5. LA FORMACIÓN DE FALSOS AMIGOS

Como hemos comentado *supra* en los contrastes con la prefijación, hemos de destacar los distintos valores que adquiere el verbo base. En el caso de la prefijación con *s-*, son numerosos los términos que presentan una amplia asimetría en su traducción al español, debido a que los conceptos se forman por la acción interactiva en la cultura y en la sociedad, por lo cual es compleja la variedad generada de rasgos semánticos en función del contexto (Labov 1973: 347). Estos nuevos valores llegan a crear un vasto corpus de falsos amigos,<sup>16</sup> entre los que destacamos los siguientes verbos con *s-*:

### *Scansare*/"Descansar"

1. Apartar  
*Hanno scansato tutte le sedie del salotto per aver spazio per il break.*
2. Evitar  
*Partendo prima abbiamo scansato il temporale.*
3. Sortear  
*Il cognato è riuscito a scansare i problemi familiari con un atteggiamento positivo.*

El uso del prefijo *des-* en español con el sema de negación con sentido privativo, en contraste con el italiano *ex-* con la idea de 'más allá', a pesar de partir de la misma base *camp̄s̄are*, debemos entenderlo como la causa de las desviaciones semánticas presentes. De este modo, en italiano, como vemos en las acepciones, la idea en la que aparece estructurado este concepto con la implicación de una cierta subjetivación (Langacker 1991: 57) queda reflejada en los correspondientes significados de "apartar", "evitar" y "sortear".

### *Scorrere*/"Ecurrir"

1. Pasar  
*Scorse tutto l'elenco fino a trovare il suo nome.*
2. Fluir  
*Nelle ore di punta il traffico non scorre.*
3. Surcar  
*Sentendo quelle parole, le lacrime iniziarono a scorrere lungo il viso.*
4. Transcurrir  
*Durante le vacanze le giornate scorrono troppo velocemente.*
5. Pasar por  
*Al cinema le immagini scorrevano sullo schermo, ma lei aveva la testa da un'altra parte.*
6. Deslizarse por  
*I suoi occhi scorrevano sulle pagine del libro alla ricerca della verità del personaggio.*

<sup>16</sup> Parte del muestrario de ejemplos aquí presente procede de Luque Toro, L. y Luque Colautti, R. (2019).

Encontramos en este grupo dos verbos con el mismo origen latino *excurrere*, con un desarrollo bien distinto, sobre todo a nivel de uso. Mientras que en español la idea de “escurrir” en sentido físico coincide parcialmente<sup>17</sup> con respecto al italiano, por el contrario, en su uso figurativo contrasta y aparece pronominalizado, como vemos en las acepciones siguientes del DLE: «7. Salir huyendo; 8. Esquivar algún riesgo, dificultad, etc.; 9. Decir algo más de lo que se debe o quiere decir; 10. Excederse, por lo común inadvertidamente, al ofrecer o dar por una cosa más de lo debido» (201423: s. p.). Si en estas acepciones figurativas podemos hablar de una cierta equivalencia, salvo en la temporal, hacemos notar que en las dos últimas, como bien señala De Mauro, existe un uso obsoleto (2000: 2368), que igualmente se da en español, pero que no aparece en el Diccionario académico.

#### *Sfamare* / “Desfamar”

1. Quitar el hambre  
*Gli antipasti mi hanno sfamato.*
2. Alimentar a  
*Col suo stipendio riesce a malapena a sfamare la sua famiglia.*

Es este uno de los casos en que se da un vacío de base entre las lenguas, por lo que, si en italiano tenemos *fame*, en español, por el contrario, se parte de *fama*, con lo cual el falso amigo se genera por razones fonéticas. Ante esta falta de identidad de bases en los respectivos orígenes, debemos señalar que el papel del prefijo *s-* se aplica tanto al hecho de “quitar el hambre” como al de “alimentar” o “sustentar”, donde se refleja el valor intensivo del mismo.

#### *Sfigurare* / “Desfigurar”

1. Estropear  
*Durante l'ultimo incontro di box gli hanno sfigurato il viso.*
2. Quedar mal  
*Avevo paura di sfigurare in confronto agli altri colleghi.*

El diferente origen de estas formaciones, *exfigurare* y *defigurare* respectivamente, nos explica la diferencia de uso, ya que en español predomina la idea de privación, entre otros valores (Varela 2005:66), como vemos en sus acepciones no pronominales: «1. Desemejar, afear, ajar la composición, orden y hermosura del semblante y de las facciones; 2. Disfrazar y encubrir con apariencias diferentes el propio semblante, la intención u otra cosa; 3. Oscurecer e impedir que se perciba la forma o figura de algo; 4. Alterar las verdaderas circunstancias de algo» (DLE, 201423: s. p.). En italiano, por el contrario, existe una mayor fre-

17 Efectivamente en todas las acepciones con sentido físico, menos en la temporal, observamos que no es común el uso de “escurrir”.

cuencia determinada por la funcionalidad, como ya hemos observado en los otros casos, y por el valor semántico de carácter intensivo que infiere s- a la base verbal en la acepción de “quedar mal”.

#### *Stendere* / “Extender”

1. Estirar  
*Dopo la gran camminata ha sentito il bisogno di stendere le gambe.*
2. Tender  
*Se stendi le camicie al sole si asciugheranno rapidamente.*
3. Untar  
*A colazione ci piace stendere un po' di burro sul pane.*
4. Derribar a  
*Il turista stese il delinquente con un pugno in faccia.*
5. Derrotar a  
*Il Milan stese la squadra tedesca nel primo confronto.*
6. Redactar  
*Il segretario ha steso il verbale della riunione.*

El mismo origen latino *extendere* de los dos términos nos explica, en sentido físico, una mayor congruencia en el uso entre ambas lenguas, con la particularidad de una significativa pronominalización en español,<sup>18</sup> si bien hemos de señalar que en este último, a nivel traductor, predominan las acepciones diversas del verbo origen, como son “estirar”, “tender”, “untar” y “redactar”. Por lo contrario, en sentido figurado, en italiano aparece el valor intensivo máximo de la extensión, entendido como “derribar” o “derrotar”, que encontramos en las acepciones 4 y 5.

#### 6. SOBRE LA COMPLEJIDAD TRADUCTORA DEL PREFIJO S- EN ESPAÑOL

En italiano, s- da lugar por línea general a formaciones parasintéticas con un sentido intensificativo a partir de nombres o adjetivos, como *sbianchire* (de *bianco*). El español, contrariamente, construye estos verbos con valor intensivo sin recurrir a la prefijación derivada de “ex-”<sup>19</sup> o recurriendo a otros prefijos como “a-” (véase “avergonzar”). A continuación, presentamos algunos ejemplos significativos:

18 De las diez acepciones presentes en el DLE son ocho la que aparecen en forma pronominal.

19 Prueba de esta distinta evolución del prefijo es la tercera acepción que el DLE presenta para “ex-”, con la idea de «sin significación precisa» (201423: s. p.).

<i>sbianchire</i>	“blanquear”
<i>sbrigare</i>	“tramitar”
<i>scaldare</i>	“calentar”
<i>scampare</i>	“evitar”
<i>scoglionare</i> <sup>20</sup>	“joder”
<i>scongiurare</i>	“conjurar”
<i>scuotere</i>	“sacudir”
<i>sgraffiare</i>	“arañar”
<i>smerciare</i>	“vender”
<i>spelare</i>	“pelar”
<i>svergonnare</i>	“avergonzar”

Además, a diferencia del italiano, en español encontramos pocos términos con el prefijo “ex-” con el valor de ‘fuera’ o ‘más allá’ en sentido físico o figurado. Entre los verbos con esta prefijación, señalamos los siguientes:

“exacerbar”, “exaltar”, “exasperar”, “excarcelar”, “excavar”, “exceder”, “excitar”, “exclamar”, “excluir”, “exfoliar”, “eximir”, “existir”, “exorbitar”, “exornar”, “exorcizar”, “expandir”, “expansionar”, “expectorar”, “expedientar”,<sup>21</sup> “expedir”, “expeler”, “expender”, “experimentar”, “expiar”, “expirar”, “explayar”, “explicar”, “explorar”, “explosionar”, “explotar”, “expoliar”, “exponer”, “exportar”, “expresar”, “exprimir”, “expurgar”, “expulsar”, “extender”, “exterminar”, “extraer”, “extremar”, “exultar”.

Estos verbos encuentran correspondencia en italiano con *s-* (*scarcerare, scavare, spedire, spacciare, sperimentare, spiare, spiegare, scoppiare, spogliare, spremere, spurgare, sterminare, stremare*) o con el prefijo latino derivado en *es-* o *ec-* (*esacerbare, esaltare, esasperare, eccedere, eccitare, esclamare, escludere, esfoliare, esimere, esistere, esagerare, esornare, esorcizzare, espandere, espettorare, espellere, espiare, estendere, esplorare, esplodere, esporre, esportare, esprimere, estendere, estrarre, esultare*), salvo “expedientar”, como indicamos en la nota a pie de página.

Con frecuencia, los verbos con *s-* en italiano encuentran correspondencia, por equivalencia semántica, con verbos con el prefijo “des-” en español. Entre ellos, señalaremos los siguientes:

20 Existe en español la forma pronominal “descojonarse”, que constituye un falso amigo dado que significa «desternillarse, troncharse de risa» (201423: s. p.).

21 Anotamos que, partiendo del español, sería este un caso en el que tendríamos que parafrasear en italiano, exactamente, *Aprire un provvedimento disciplinario nei confronti di qualcuno*.

*sbarazzare, sbarcare, sbattezzare, sbiadire, sbilanciare, sbloccare, sbrigliare, sbrogliare, sbucare, scardinare, scaricare, scollare, scollegare, scomparire, scomporre, sconcertare, scongelare, sconvolgere, scontare, sdegnare, sfasare, sfavorire, sferrare, sfilare, sfrattare, sfrenare, sganciare, sgranare, sgrassare, sgravare, slegare, smentire, spacciare, sparare, spargere, sparpagliare, spellare, spennare, spiccare, spiazzare, spogliare, staccare strappare, straripare, svanire, svegliare, svolgere, svoltare.*<sup>22</sup>

La alta funcionalidad del italiano en la creación de derivados con este prefijo contrasta con el español tanto morfológica como semánticamente, de ahí que encontremos un buen número de formaciones que recurren a la paráfrasis para encontrar su correspondencia. Entre otras, citamos:

<i>sbroccare</i>	“perder los papeles”
<i>sconfinare</i>	“traspasar la frontera”
<i>sfamare</i>	“quitar el hambre”
<i>sfebrare</i>	“bajar la fiebre”
<i>sfrecciare</i>	“pasar zumbando”
<i>smacchiare</i>	“quitar las manchas”
<i>smezzare</i>	“partir por la mitad”
<i>smontare</i> <sup>23</sup>	“terminar de trabajar”
<i>snodare</i>	“quitar el nudo”
<i>spagliare</i>	“quitar la paja”
<i>spanare</i>	“pasarse de rosca”
<i>spanciare</i> <sup>24</sup>	“darse un planchazo”
<i>sparecchiare</i>	“quitar la mesa”
<i>spolverare</i>	“quitar el polvo”
<i>svasare</i>	“cambiar de maceta”
<i>svenare</i>	“cortar las venas”
<i>svernare</i>	“hacer pasar el invierno”

22 La sobreposición de *s-* en la base verbal llega a ser frecuente en la morfología de estos verbos en formaciones como *sbarazzare* procedente de la base *imbarazzare*, *sbarcare* de *imbarcare*, *sbrogliare* de *imbrogliare*, etc.

23 Nos referimos al uso intransitivo del verbo, ya que, en su forma transitiva, equivale, entre otras formas, a “desmontar”, “desmantelar”, “desanimar”, “desinflar”, etc. (Arqués y Padoan 2012: 2575).

24 Hacemos referencia al uso intransitivo del verbo, puesto que, en su forma transitiva, corresponde a “encabritar” con referencia a un avión (Arqués y Padoan 2012: 2603).

Donde podemos observar una serie de paráfrasis en español que se configuran en torno a los valores originales del prefijo, como el de intensidad, en “perder los papeles”, “pasar zumbando”, “partir por la mitad”, “terminar de trabajar”, “pasarse de rosca”, “darse un planchazo”, “cortar las venas” y “hacer pasar el invierno”; y el privativo en “traspasar la frontera”, “quitar el hambre”, “bajar la fiebre”, “quitar las manchas”, “quitar el nudo”, “quitar la paja”, “quitar la mesa”, “quitar el polvo” y “cambiar de maceta”.

Equivalencias traductoras a las que añadiremos verbos especiales como, entre otros, *smettere*, *smuovere* o *svenire*, que son proyecciones metafóricas de la idea original de este prefijo como encontramos en sus respectivas acepciones. Con *smettere*, la idea base de ‘más allá’ nos aparece con carácter intensivo como abandono, en *Ho smesso il lavoro dopo venti anni* (“He dejado el trabajo después de veinte años”), y como proyección máxima de la intensidad, en *Ha smesso un paio di scarpe che aveva da anni* (“Ha tirado un par de zapatos que tenía desde hace años”), donde entenderemos el abandono como una terminación. Con *smuovere*, esta idea se concibe como ir ‘más allá’ de la propia idea, es decir, cambiar de opinión, en *Pietro ha deciso di andare in Afganistan e nessuno riesce a smuoverlo* (“Pietro ha decidido ir a Afganistán y nadie consigue que abandone esa idea”). Finalmente, con *svenire* esta idea aparece interpretada como un ir ‘más allá’ del movimiento, es decir, “desvanecerse” o “desmayarse”, en *È svenuto per la stanchezza* (“Se ha desmayado por el cansancio”).

## 7. CONCLUSIONES

Analizando las diferencias etimológicas y las competencias cognitivas entre ambas lenguas hemos intentado arrojar un poco de luz a un tema de difícil sistematización, ya que la palabra presenta en cada cultura un desarrollo bien distinto. En nuestro caso hablamos incluso de palabras que tienen un mismo origen, lo cual tantas veces representa una mayor divergencia (piénsese en el caso de los falsos amigos *scansare* / “descansar”, *scorrere* / “escurrir”, *sfamare* / “desfamar”, *sfigurare* / “desfigurar” y *stendere* / “extender”), visto el distinto desarrollo que presentan las lenguas a lo largo del tiempo.

Por otra parte, podemos afirmar que, si bien existen con *s-* unos principios morfológicos que en cierta medida facilitan algunas equivalencias (por ejemplo, las correspondencias con las formas españolas “*ex*” o “*des-*”), como hemos podido ver a lo largo de este trabajo, queda por resolver el estudio de los rasgos semánticos inherentes a cada palabra, tanto en una lengua como en otra. No se trataría, por consiguiente, de una investigación solo a nivel de equivalencias de prefijos, sino de analizar esa nueva base (prefijo-base verbal) que nos permita estudiar todo el perfil posible que se crea en el dominio pragmático.

Nuestra conclusión, pues, desde el punto de vista de la traducción, es la variedad de soluciones encontradas en los lemas estudiados (recordemos los casos de ausencia de prefijo o de paráfrasis en español), así como la alta funcionalidad de *s-* en italiano que nos permite deducir que existe una exhaustiva correspondencia en el estudio de cada una de las palabras derivadas.

- Arqués R. y Padoan A. (2012) *Il Grande Dizionario di Spagnolo*, Bologna, Zanichelli.
- Corpas G. (1996) *Manual de fraseología española*, Madrid, Gredos.
- Dardano M. (1978) *La formazione delle parole nell'italiano di oggi*, Roma, Bulzoni.
- De Mauro T. (2000) *De Mauro. Il dizionario della lingua italiana*, Milano, Paravia.
- Guilbert L. (1975) *La créativité lexicale*, Paris, Larousse.
- Iacobini C. (2004) "Prefissazione", in *La formazione di parole in italiano*. Ed. by M. Grossman & F. Rainer, Tübingen, Max Niemeyer Verlag, pp. 97-163.
- Labov W. (1973) "The boundaries of words and their meaning", in *New ways of analyzing variation in English*. Ed. by J. Fishman, Washington D. C., Georgetown University Press, pp. 340-373.
- Langacker R. W. (1991) *Concept, Image and Symbol: Basis of Grammar*, Berlin, Mouton de Gruyter.
- Luque Toro L. (2009) "Una distinta conceptualización en el uso de los prefijos entre italiano y español", in *Didáctica del español como L2. Enseñanza, contrastividad y registros lingüísticos*. Ed. by M. C. Muñoz Medrano, Napoli, Flavius Edizioni Pompei, pp. 59-72.
- Luque Toro, L. y Luque Colautti R. (2019) *Diccionario contextual italiano-español de parónimos*, Madrid, Arco/Libros.
- Pinker S. (2000) *Come funziona la mente*, Milano, Mondadori.
- Plag I. (2003) *Word-Formation in English*, Cambridge, CUP.
- Porzig W. (1986) *El mundo maravilloso del lenguaje*, Madrid, Gredos.
- RAE (2014) *Diccionario de la Lengua Española*, <https://dle.rae.es/>, consultado el 03/09/2021.
- Rainer, F (1991) "On the Nature of Word-Formation. Processes: Evidence from Spanish", in *Manuscriptos*, Madrid, Fundación Ortega y Gasset, pp. 14-52.
- Scalise S. y Zannier I. (1983) "Restrizioni alle regole di formazione di parole: le condizioni di adiacenza e la condizione atomo", *Quaderno Paravini di Linguistica*, 3, pp. 159-210.
- Varela S. (2005) *Morfología léxica: La formación de palabras*, Madrid, Gredos.





# La traducción automática del género (español-italiano): análisis de ejemplos traducidos con DeepL y Google Traductor

ROSARIA MINERVINI  
Università di Salerno  
[rminervini@unisa.it](mailto:rminervini@unisa.it)

## ABSTRACT

In an increasingly globalised world, with the need to communicate and trade quickly, machine translation has become more and more important. It helps both language professionals in their daily work, translators, and non-professionals to understand and create foreign language texts. The question of gender in language is still very topical and although it has been discussed a lot in translation theory, it still needs to be studied in the field of machine translation. In this paper we will observe through some examples deriving from political and journalistic discourse how two machine translation engines available on the market today, DeepL and Google Translator, translate gender in the Spanish-Italian linguistic combination.

## KEYWORDS

Machine translation, Spanish-Italian translation, Gender, DeepL, Google Translator

## 1. INTRODUCCIÓN

El mundo de la traducción ha experimentado en los últimos años cambios sustanciales, debidos en especial al gran avance de las tecnologías y de la inteligencia artificial, que han permitido el desarrollo de la traducción automática (en adelante, TA), tanto de tipo estadístico como neuronal. De hecho, la TA ha entrado a formar parte de la vida de un gran número de personas, que la emplean en el día a día para fines personales. También ha llegado a convertirse en un instrumento imprescindible para los traductores, que suelen utilizarla a nivel profesional y académico. Así, los sistemas de TA han cambiado rápidamente, adaptándose tanto a las exigencias del mundo globalizado en el que vivimos como a las del sector profesional, que requiere –entre otras demandas–, traducir un número cada vez mayor de palabras en el menor tiempo posible, reducir los costes y conseguir traducciones de alta calidad.

El objetivo de este estudio es, por un lado, averiguar la forma en que dos de los principales sistemas de TA disponibles hoy en día en el mercado –DeepL y Google Traductor (en adelante, DL y GT)– traducen el género para así explorar eventuales semejanzas y/o diferencias entre ambos sistemas y, por otro, ampliar la reflexión sobre un tema tan actual y complejo como el lenguaje inclusivo. Como es bien sabido, el empleo de las herramientas de TA permite llevar a cabo la traducción de una gran cantidad de texto en poco tiempo, pero estos sistemas no siempre consiguen transmitir los elementos culturales ni sus diferentes connotaciones lingüísticas. Por este motivo, cuando se emplean estos instrumentos, suele resultar imprescindible una labor de posesición, más o menos profunda. En nuestro caso, proponemos, además, una reflexión sobre elementos muy concretos, como los vinculados al género. Después de un breve recorrido a través de los principales sistemas de TA,<sup>1</sup> se analiza la traducción del género en algunos textos periodísticos y políticos. Se han elegido estas tipologías textuales no solo por la gran cantidad de personas que a diario están expuestas a este tipo de textos sino también porque a menudo se emplean en el aula de lengua extranjera y, por lo tanto, podrían servir de inspiración para desarrollar actividades en el aula.<sup>2</sup>

1 Se proporciona aquí solo un resumen; para profundizar en los diferentes sistemas de TA y en aspectos como la evaluación, la preedición y la posesición, véase Sánchez Ramos y Celia Rico Pérez (2020).

2 Los sistemas de TA pueden emplearse incluso en contextos de enseñanza/aprendizaje de lenguas extranjeras, en especial, en niveles altos, ya que en el posterior y necesario análisis de las traducciones efectuadas por las máquinas, la posesición, representa un estímulo muy útil para reflexionar sobre la lengua extranjera que se está aprendiendo y para ampliar el conocimiento sobre ella. Véanse Zang y Torres (2019), Niño (2009) y Conde Noguero (2018) entre otros.

A pesar de ser un sistema muy innovador, cabría datar los orígenes de la TA en el siglo XVII, cuando Leibniz y Descartes vislumbraron una herramienta para comunicarse en el uso de diccionarios mecánicos basados en códigos numéricos universales (Díaz Prieto 2012: 142). Sin embargo, hubo que esperar hasta la Segunda Guerra Mundial, cuando los ordenadores comenzaron a ser realidades tangibles, para que se empezara a investigar de forma más sistemática en este ámbito. De hecho, la TA indaga el uso de software para que los textos de un lenguaje natural sean traducidos a otro; Sánchez Ramos y Rico Pérez (2020: 1) dicen que “[l]a TA se concibe como la aplicación de la tecnología informática a la traducción de textos de una lengua a otra sin intervención humana”.

Puede afirmarse que el estudio de la TA empezó hacia mediados del siglo XX, con Warren Weaver, quien, en su célebre memorándum, propuso emplear los ordenadores para traducir de una lengua a otra.<sup>3</sup> La primera traducción llevada a cabo con uno de estos sistemas se hizo en 1954, con un texto ruso traducido al inglés, en la que pudieron observarse muchas debilidades del sistema, pero también que la TA era posible y que se podían alcanzar niveles adecuados y útiles (Hutchins 1986). Sin embargo, la TA vivió una serie de altibajos y experimentó una verdadera debacle en los años sesenta del siglo pasado, cuando se empezó a dudar de sus posibilidades para conseguir resultados satisfactorios. En EE. UU., se creó el ALPAC (*Automatic Language Processing Advisory Committee*), cuyo objetivo era averiguar la evolución de la TA. El ALPAC publicó un informe en el que se manifestaba la dificultad para conseguir resultados cualitativamente aceptables en plazos breves (Hutchins 1986). Este informe frenó la investigación, pero no consiguió pararla del todo. Además, en los años setenta, la TA fue impulsada por la situación de multilingüismo y las consiguientes necesidades de traducción de la Comunidad Europea. En Luxemburgo, la entonces Comunidad Económica Europea empleó desde 1976 el sistema Systran, y en Canadá, se utilizó MÉTEO (Hutchins 1986).

En los años ochenta del siglo pasado, la TA vivió un nuevo empuje, que fue amplificado en los años noventa con la llegada de Internet. En estos años, de hecho, se crearon nuevas aplicaciones informáticas y se trabajó principalmente con sistemas basados en ejemplos, que tomaban frases y ejemplos traducidos de corpus bilingües (Hutchins 1986). Desde entonces, la investigación ha seguido adelante reconociendo el papel fundamental que los sistemas de TA pueden desempeñar en un mundo cada vez más globalizado.

Hoy en día existen diferentes sistemas de TA, que codifican una lengua de partida y generan textos en una lengua meta. A pesar de que estos sistemas hayan sido clasificados de forma distinta, podríamos dividirlos en tres tipos principales: los sistemas basados en reglas, los basados en corpus y los neuronales.

3 Para profundizar en los aspectos del memorándum, véase Hutchins (1986 y 1997).

Sintetizando, los primeros sistemas realizan cambios sustituyendo las palabras de la lengua de partida por equivalentes de la lengua meta (Díaz Prieto 2012: 160). En este patrón entran los sistemas directos, los indirectos y la interlingua. Cuando se emplea el sistema directo, el texto que se necesita traducir se introduce en el sistema, que sustituye las palabras buscando equivalentes en el texto meta a través de un glosario. También se realiza un análisis de tipo morfológico, pero los resultados suelen ser poco satisfactorios, en especial, cuando la combinación lingüística no es de lenguas próximas.<sup>4</sup>

Emplear listas o grupos de palabras hace que la traducción sea literal, y por eso mismo que se produzcan muchos errores, ya que no se realiza un análisis sintáctico como tampoco se elige siempre el sentido correcto de las palabras que tienen más acepciones (Díaz Prieto 2012: 160). Sin embargo,

[...] a pesar de las limitaciones, algunos de estos sistemas han demostrado ser rentables y útiles, especialmente cuando nos hallamos ante un texto con un vocabulario y un estilo que están bien definidos y limitados. La calidad de la traducción está en relación directa con la calidad de la información que contiene el diccionario y en general, producen traducciones aceptables y legibles, aunque con elevada dosis de post-edición (Díaz Prieto 2012: 161).

Para intentar solventar los problemas de estos sistemas, en un segundo momento, se desarrollaron los sistemas indirectos, que integran reglas lingüísticas (morfológicas, sintácticas, semánticas) a la hora de analizar el texto de salida, y lo examinan en módulos diferentes antes de producir el nuevo texto. Se diversifican dos tipos de estrategias indirectas: la transferencia y la interlingua. Los sistemas basados en la transferencia hacen un primer análisis del texto de partida, crean una representación abstracta de él, para luego hacer una representación abstracta del texto de llegada, que permite establecer equivalencias de tipo sintáctico, léxico, semántico, para finalmente, gracias a esta información, generar el texto final (Tertoolen 2012: 142). Se trata, por tanto, de un sistema que crea una representación intermedia de las dos lenguas y que necesita glosarios tanto monolingües (para los aspectos gramatical, morfológico y semántico) como bilingües (para la equivalencia léxica entre lenguas). También el sistema de la interlingua se basa en reglas lingüísticas como los anteriores; la diferencia con respecto al sistema basado en la transferencia es que aquí el texto se convierte a un lenguaje intermedio, que no depende ni de la lengua de partida ni de la lengua de llegada. El texto intermedio que se crea es independiente, de manera que se pueden añadir módulos de distintas lenguas e incluso se puede traducir de la lengua origen a la lengua meta y viceversa.

En los años ochenta del siglo pasado, aparecieron otros sistemas, caracterizados por tener enfoques más empíricos hacia el problema de la traducción, y que se desarrollaron gracias al avance de Internet en los años noventa del siglo pasa-

4 Los sistemas Méteo y el antiguo Systran son sistemas de TA directa.

do. En concreto, los enfoques fueron tres: el basado en ejemplos, el estadístico y, finalmente, el neuronal.

El primer tipo, basado en ejemplos, emplea corpus de textos bilingües alineados al nivel de la frase. Este sistema busca equivalencias y relaciona la frase del texto origen con los ejemplos/segmentos que aparecen en el corpus para luego traducirlos. Este tipo de construcción no hace otra cosa que recuperar de una base de datos bilingüe traducciones realizadas anteriormente, pero no le permite al traductor (humano) intervenir en el proceso de traducción, a diferencia de lo que sucede con las herramientas de traducción asistida. Por lo tanto, el traductor no puede elegir entre diferentes traducciones y se invalida la relación persona-máquina.

Los sistemas estadísticos han representado durante muchos años el modelo dominante en TA. Ejemplos de ello son Bing, Google Traductor o Moses. Estos sistemas emplean corpus de textos que ya han sido traducidos y que están disponibles en Internet. Las frases o los grupos de palabras de los textos en la lengua de salida son alineados con sus equivalencias traducidas, para luego calcular la probabilidad de que una palabra determinada sea la equivalente de otra en el texto original. Se trata de un enfoque basado en datos (*data-driven*) y técnicas de aprendizaje (*machine learning*), así como en el principio de la probabilidad de distribución  $p(e|f)$ , donde  $e$  en la lengua de llegada es la traducción estadísticamente más frecuente de un segmento textual  $f$  en la lengua origen (Monti 2019). Esto significa que es necesario adiestrar estos sistemas con datos de entrenamiento como corpus paralelos y monolingües voluminosos. Los corpus paralelos operan como modelo traductivo y se utilizan para computar en la lengua de llegada las traducciones más frecuentes de términos o secuencias de términos en la lengua origen. Los corpus monolingües, por su parte, sirven de modelo lingüístico, ya que operan en la lengua de llegada con el objetivo de medir que la palabra o la secuencia de palabras traducidas se corresponda realmente, tanto a nivel morfológico y sintáctico como en el uso real en la lengua de llegada (Monti 2019).

En los últimos años el panorama de la traducción, en general, y de la TA, en particular, han cambiado radicalmente; en especial, tras la aparición de un nuevo modelo en el mercado: la traducción automática neuronal (TAN). Hoy en día, los sistemas de tipo neuronal representan un nuevo estado de la técnica en la investigación de la TA (Castilho et al. 2017). Estudios recientes (Hasler et al. 2018; Michon et al. 2020) destinados a fomentar la investigación en el campo de la TA muestran, de hecho, que los sistemas de traducción neuronal suelen ser capaces de producir mejores traducciones que los basados en la estadística. Sin embargo, a pesar de la gran calidad alcanzada por la TAN en la actualidad, sus resultados aún no cubren muchos dominios específicos que se manejan a diario en el sector de la traducción. No en vano, es interesante observar que la traducción de la terminología específica de un dominio, además de desempeñar un papel crucial en los flujos de trabajo de proyectos de traducción especializada, aún constituye una de las áreas menos exploradas, dados los numerosos retos y dificultades

que plantea. En este contexto, algunos estudios (Macketanz et al. 2017; Haque et al. 2019) han argumentado que la TAN presenta menor eficacia en la traducción de los términos específicos de un dominio en comparación con la TAE (Traducción Automática Estadística). Aunque se ha demostrado que la TAN, al igual que la TAE, se beneficia de la disponibilidad de corpus paralelos o monolingües específicos del dominio para aprender términos específicos del mismo (Farajian et al. 2018), no es una solución universalmente aplicable, ya que a menudo un dominio puede ser demasiado estrecho y carecer de datos útiles y necesarios para entrenar un motor neuronal. Además, mientras que las memorias de traducción representan datos de entrenamiento listos para ser usados para la adaptación del sistema neuronal a un dominio determinado, las bases de datos terminológicas son más difíciles de manejar, como atestigua el considerable número de métodos propuestos para integrar la terminología perteneciente a un dominio específico en la TAN en tiempo de ejecución (Dinu et al. 2019; Hasler et al. 2018; Michon et al. 2020). A pesar de estas limitaciones, la TAN ha revolucionado el mundo de la traducción, ampliando el nivel de calidad y acortando el tiempo de posesión, con una consiguiente reducción drástica de los costes.

### 3. LA TRADUCCIÓN AUTOMÁTICA DEL GÉNERO

El debate sobre el género y, más específicamente, sobre el lenguaje inclusivo está viviendo en los últimos años un gran vigor en España. Se pretende impulsar el empleo del género femenino en el idioma, con el fin de fomentar un uso del lenguaje que sea inclusivo y no sexista. Se trata de un debate amplio y complejo, que se ha instrumentalizado a menudo con fines ideológicos y, por eso mismo, a veces muy polarizado, al cual no se pretende ofrecer ningún tipo de propuesta/solución en este trabajo.<sup>5</sup> A propósito de lenguaje inclusivo en cuanto al género, muchos organismos, como por ejemplo las Naciones Unidas, han proporcionado unas orientaciones para su uso, que atestiguan la importancia del debate.<sup>6</sup>

Cuando se habla de género, hay que distinguir entre género natural (o sexo) y género gramatical. Esto significa que un atributo del mundo, el sexo (género natural, biológico y también concepto sociocultural) suele expresarse a través de una categoría lingüística, el género, y muchas veces estas categorías no coinciden. El género como categoría lingüística se manifiesta a través de los sustan-

5 El lector interesado encontrará, en Álex Grijelmo (2019), una propuesta para acercar posiciones distantes sobre lenguaje inclusivo y, en Álvarez de Miranda (2018), una reflexión sobre cómo funcionan el género y la lengua española.

6 La orientación se resume en tres puntos principales: 1. Evitar expresiones discriminatorias; 2. Visibilizar el género cuando lo exija la situación comunicativa; 3. No visibilizar el género cuando no lo exija la situación comunicativa (por ejemplo, preferir la expresión *Acudieron periodistas de distintos medios* a *Acudieron algunos periodistas de distintos medios*) [<https://www.un.org/es/gender-inclusive-language/guidelines.shtml>].

tivos, los artículos, los pronombres, los adjetivos y algunas formas verbales. En lenguas como el español y el italiano, existen solo los morfemas de género masculino y femenino por lo que concierne a los sustantivos y adjetivos. Ambas lenguas, además, marcan el género gramatical y el sexo biológico mediante la terminación (*chica/-o, ragazza/-o*), la oposición de palabras (*padre-madre*), los artículos (*el/la estudiante, il/la badante*), los demostrativos (*este/esta representante, questo/questa rappresentante*), los sustantivos epicenos (*la persona*). Sin embargo, en ambas lenguas, se proponen interpretaciones por defecto del género.

Como escriben Rescigno et al. (2020) en su estudio sobre los fenómenos naturales de género en la traducción,

[g]ender manifests itself in a language in many ways, and different languages use different linguistic devices to mark (or sometimes 'not mark') gender. When dealing with language, three types of gender come into play: natural gender, grammatical gender and social gender. Natural gender is generally based on the sex of a person or an animal realised by means of the male/female polarity or on the absence of sex for neutral nouns. Grammatical gender, instead, is not always coherent with the semantic categorization of a word and can vary from language to language since it depends on the representation of objects in the world on the basis of specific properties attributed to them in a specific cultural context. Social gender is used in relation to the properties of a word on the basis of which the speakers of a language associate the natural gender of a person to a word [...] this mainly happens with names of professions, such as for instance *doctor* or *nurse* which are interpreted according to social stereotypes concerning the roles of males and females in the society.

Monti (2017: 412) señala que la cuestión del género –ampliamente debatida por la traductología– sigue necesitando estudios cuando se trata de TA. Sin embargo, el debate político, social y, por supuesto, lingüístico que se ha avivado en los últimos años ha acarreado que se estudien toda una serie de cuestiones relacionadas con los estereotipos en el lenguaje natural y también en los sistemas de TA. De hecho, varios son los estudiosos que están investigando la relación entre TA y sesgos de género (Prates et al. 2019; Escudé Font e Costa-jussà 2019). Sin embargo, se requieren más investigaciones al respecto. Hablando de la combinación lingüística inglés-italiano, Monti (2017: 412) resume los problemas para traducir el género en la TA –que podrían aplicarse incluso a la pareja de lenguas que aquí nos interesa– de la siguiente manera:

sono molteplici e di diversa natura, come la riproduzione di stereotipi di genere principalmente legati ai ruoli maschili e femminili nella società, la scorretta attribuzione del genere a proforme (pronomi personali, relativi, etc.), aggettivi, participi passati e apposizioni, oltre a un uso più frequente del genere maschile per pronomi e nomi anche quando questi si riferiscono al genere femminile.

De hecho, como señalan Rescigno et al. (2020), el género está presente en los datos que se utilizan para entrenar los sistemas de TA



due to the demographic features of the human training data and because of the nature of stereotypes and biases we communicate in our day-to-day communications. As most state-of-the-art MT systems handle translations on the sentence-level, gender phenomena are, usually, resolved on statistics inferred from the training data. Mistranslations of gender information occur more frequently when translating from gender-neutral languages, such as English, into morphological-rich languages, such as Italian or French, which explicitly mark gender and require additional information to correctly translate gender phenomena. When such additional information or context is not provided, the system will pick the most likely variant.

Son conocidos los esfuerzos para crear traducciones equitativas y no sexistas, eliminando incluso los estereotipos de género, gracias también a la disponibilidad de una cantidad cada vez más copiosa de datos para entrenar las máquinas. Sin embargo, se hace necesario seguir investigando para disminuir o incluso llegar a eliminar estereotipos y sesgos de género. En este sentido, es importante incluso que los programadores que eligen los datos de los que se nutre la inteligencia artificial tengan una visión informada y equilibrada –neutral– del mundo, e incluso que sean mujeres, ya que los datos reflejan la visión del mundo del programador.<sup>7</sup>

Como escriben Vanmassenhove et al. (2019), “[...] gender-related errors are not just harming the quality of the translation as getting the gender right is also a matter of basic politeness”, y añaden que la actual tendencia de los sistemas a perpetuar un prejuicio masculino equivale a una discriminación negativa contra la mitad de la población.

En 2018, Google anunció que iba a introducir en su sistema las posibles alternativas de género, con la traducción del masculino y femenino para algunas lenguas. Sin embargo, es necesario seguir perfeccionando el sistema, ya que, como han puesto de manifiesto varios estudiosos, como Sanz Romero (2021), cuando una lengua tiene pronombres neutros, GT sigue aplicando un sesgo machista, ya que emplea el género masculino para muchas profesiones y el femenino para tareas relacionadas con el cuidado de la casa y de los niños. Otra cosa ocurre con lenguas como el inglés, para cuya traducción del género hacia el español GT emplea los dos géneros a la vez.

En el estudio de Prates et al. (2019), en el que se traducen frases de 12 diferentes lenguas naturales al inglés empleando la API de GT, se ha observado que este sistema tiene una marcada tendencia hacia el masculino, en particular, para aquellos sectores que suelen asociarse a una distribución poco equitativa del género o a estereotipos como los trabajadores de los ámbitos CTIM (Cien-

7 A este respecto, Vanmassenhove et al. (2019) han experimentado con la incorporación de etiquetas de género del hablante durante el entrenamiento de sistemas de TAN con el objetivo de mejorar la concordancia morfológica, centrándose en especial –aunque no solo– en pares de lenguas que expresan el género gramatical. Y han llegado a la conclusión de que informar al sistema TAN proporcionándole etiquetas que indiquen el género del hablante puede acarrear mejoras significativas respecto a los sistemas de referencia, especialmente en el caso de las lenguas que expresan acuerdo gramatical de género, mientras que estas mejoras no siempre son consistentes para traducciones EN-FR.

cia, Tecnología, Ingeniería y Matemáticas). Estos datos han sido comparados en un segundo momento con los de la Oficina de Estadísticas Laborales de EE. UU. (U.S. Bureau of Labor Statistics, BLS) para cotejar las estadísticas obtenidas con los datos del *bureau* relativos a la incidencia de la participación femenina en cada puesto de trabajo. Se ha mostrado que GT no consigue reproducir la distribución laboral de los trabajadores de sexo femenino de forma que refleje el mundo real. A pesar de que los investigadores no esperaban encontrar una distribución de género pronominal 50:50, GT produce valores masculinos por defecto mucho más a menudo de los que se esperaría de los datos demográficos.

Frank et al. (2004) diseñaron para su investigación sobre la traducción alemán-francés 20 oraciones de prueba para determinar hasta qué punto los sistemas de traducción (Translate, T1 y Systran) cubrían los fenómenos lingüísticos relacionados con el género. También en este caso, “[t]he results demonstrate that most of the systems show severe shortcomings in their handling of gender-appropriate translations and in the treatment of gender-neutral language”. Además, los resultados arrojados señalan la necesidad de ampliar el número de sustantivos del diccionario que denominan a las mujeres en el trabajo, el deporte, la ciencia, etc.

Una vez descritos los principales sistemas de TA y comentados algunos trabajos relacionados con la traducción del género, nos parece oportuno describir la metodología empleada en el presente estudio y las principales características de los dos sistemas que se ha decidido utilizar, DL y GT, para luego pasar a analizar algunos ejemplos concretos.

#### 4. METODOLOGÍA

El objetivo de este trabajo es analizar cómo traducen el género dos sistemas de TA, DL y GT, en textos periodísticos y políticos. Se analiza, por lo tanto, el tratamiento de fenómenos relacionados con el género, como las frases sustantivas, el género gramatical y las referencias anafóricas. El motivo por el que se han elegido estos dos sistemas es porque son gratuitos y están entre los más empleados. Además, el primero nació como sistema de base estadística y luego se ha reconvertido en neuronal, mientras que el segundo surgió como sistema de base neuronal, características que permiten observar diferencias entre los dos sistemas. En ambos casos, se trata de sistemas sencillos, con interfaces simples e intuitivas, que se pueden emplear sin dificultades. Una vez establecido el objetivo del trabajo, se buscaron y seleccionaron unos textos para luego introducirlos en los dos sistemas y finalmente examinarlos. A pesar de que era posible introducir los textos directamente en los sistemas (ya que GT admite .pdf y DL .html) se prefirió convertirlos en archivos de Word, para que la traducción resultase más limpia en cuanto a formato y se facilitara así el posterior análisis de los textos. A continuación, se describen las principales características de ambos sistemas.

#### 4.1. DEEPL

DL es un sistema de traducción automática y fue lanzado en su versión gratuita por la empresa alemana DeepL GmbH en agosto de 2017, es decir, varios años después de que apareciesen en Internet otros motores de traducción, como GT. La plataforma también cuenta con una versión de pago, DeepL Pro, estrenada en marzo de 2018, que les permite a los abonados beneficiarse de un traductor web optimizado, integrar los algoritmos en su software de traducción o desarrollar nuevas aplicaciones y servicios con la API de DeepL. En un principio, el sistema solo estaba disponible para algunos idiomas, pero en la actualidad es capaz de traducir 26 lenguas. DL se apoya en los cimientos de la experiencia de Linguee – diccionario en línea que funciona a través de corpus paralelos multilingües integrados por traducciones hechas a partir de textos alineados en la web– y de los algoritmos de inteligencia artificial y aprendizaje automático ya utilizados. El sistema de inteligencia artificial, también llamando aprendizaje profundo, dio el nombre al sistema, *deep learning*. A esto hay que añadir nuevos algoritmos de traducción y un sistema de redes neuronales capaz de ofrecer un alto rendimiento en cualquier campo; no solo guarda los datos, sino que sigue aprendiendo, de manera que los mismos errores sirven para aumentar las redes neuronales artificiales con el resultado de mejorar las traducciones. Todo esto se apoya en un superordenador muy potente alojado en Islandia, capaz de traducir más de un millón de palabras por segundo, con una base de datos de búsquedas y traducciones de idiomas casi infinita. Las traducciones en línea de DL son el resultado de una combinación de algoritmos de aprendizaje automático altamente especializados y de redes neuronales convolucionales, así como de una enorme base de datos de búsquedas en línea y traducciones antiguas que permiten “entrenar” el motor de traducción y mejorar constantemente su rendimiento. Finalmente, a la estructura informática se suma un equipo de especialistas que controlan y corrigen los errores en el código, mejorando las traducciones. Uno de los límites del sistema es que una vez efectuadas unas cuantas traducciones gratuitas, deja de funcionar y es necesario abonarse. Además, solo gestiona los formatos .docx, .pptx, .txt y .html. El documento traducido puede descargarse directamente en el ordenador y el sistema ofrece sugerencias de lenguaje formal/informal. Haciendo clic sobre una palabra, aparece un glosario que sugiere otras posibles traducciones y que le permite al traductor intervenir en el proceso traductivo. Dependiendo de contexto y co-texto, el traductor puede elegir la que considere más apropiada entre las opciones propuestas.

#### 4.2. GOOGLE TRADUCTOR

Este sistema se lanzó en abril de 2006 y es una herramienta de referencia para traducciones en línea. En un principio, surgió como motor de Traducción Automática Estadística (TAE) en el que la traducción del texto requerido se realizaba antes al

inglés y luego al idioma seleccionado. En 2016, Google cambió su método de traducción, adoptando la TAN, que utiliza técnicas de aprendizaje profundo (*deep learning*) para traducir toda la frase y garantizar así una mayor precisión contextual. El servicio ofrece a día de hoy traducciones en 103 idiomas y cuenta con más de 500 millones de usuarios diarios. Sin embargo, no todas las funciones pueden utilizarse en todos los idiomas: la traducción instantánea de la cámara está disponible en alrededor de 88 idiomas, el modo de conversación se ofrece para 32 idiomas y la traducción sin conexión a Internet es posible para 59 idiomas. GT utiliza una combinación de sistemas de aprendizaje automático e inteligencia artificial. La traducción automática analiza millones de documentos de la web que ya han sido traducidos e identifica las muestras estadísticas más significativas que relacionan la traducción con el texto original. Se trata de muestras empleadas por los algoritmos para crear patrones para traducir textos similares. No es difícil emplearlo y gestiona archivos en diferentes formatos (.doc, .docx, .odf, .pdf, .ppt, .pptx, .ps, .rtf, .txt, .xls o .xlsx), pero cuando se introducen textos muy voluminosos, el sistema no consigue respetar la estructura original de los documentos.

Una vez descritos los dos sistemas, a continuación se esbozan las características de los tipos de textos empleados.

#### 4.3. LOS TEXTOS: PERIODÍSTICOS Y POLÍTICOS

Como es bien sabido, la política y los medios de comunicación son interdependientes a muchos niveles, tanto que casi no sería posible entender la política sin los medios de comunicación y viceversa, en un influjo recíproco del lenguaje empleado. Este estrecho vínculo entre ambos hace que el lenguaje político y el periodístico compartan muchas características, aunque existen entre ellos analogías basadas en su naturaleza interna (Núñez Cabezas y Guerrero Salazar 2002: 92).

Entre otras, comparten características tales como las siguientes (Núñez Cabezas y Guerrero Salazar 2002: 92-94): 1. Se trata de lenguajes profesionales dentro de la lengua común, es decir, formas de emplear el lenguaje propias de políticos y periodistas y, por tanto, son lenguajes sectoriales. 2. A menudo, surgen vocablos nuevos, tecnicismos y extranjerismos en ambos. 3. La persuasión es una característica típica del lenguaje político, pero también se da en el periodístico. 4. El lenguaje periodístico puede tener un valor político, incluso decidiendo simplemente si tratar una noticia o evitarla. 5. En ambos, aparecen procedimientos de generación léxica semejantes, como la prefijación en el periodístico y la sufijación en el político. 6. Los dos emplean la técnica del alargamiento de las palabras. 7. Emplean eufemismos. 8. Ambos son cada vez más permeables al lenguaje coloquial e incluso vulgar. 9. Utilizan rasgos formales parecidos, como las expresiones más largas, la adjetivación sinonímica, los extranjerismos (en especial, anglicismos) y los neologismos, léxico impropio (uso incorrecto de preposiciones, expresiones redundantes, etc.).

Como señalan Núñez Cabezas y Guerrero Salazar (2002: 96), el influjo entre los dos lenguajes es bidireccional aunque el político parece ser el que ejerce una mayor influencia, proporcionando las pautas para la conducta lingüística. Las palabras pertenecientes al lenguaje burocrático y administrativo, propias del lenguaje político, empleadas por los especialistas, aparecen en los medios de comunicación, haciendo que sea el propio periodista el responsable de su expansión y difusión. Los estudiosos señalan tres puntos en los que el lenguaje político penetra en el periodístico: 1. en el uso perifrástico y esotérico; 2. en la oscuridad, con léxico extraño, onomasiología ambigua y construcciones difíciles, y la aparición de palabras clave que el lector medio no conoce; 3. en el uso figurado y metafórico.

Teniendo en cuenta estos elementos, resulta evidente que estas tipologías textuales pueden analizarse juntas. Por su parte, el lenguaje periodístico influye sobre la población, que lo absorbe y lo hace propio, representando desde siempre el modelo de lengua basado en la corrección.

Para nuestro estudio, se han elegido tres textos periodísticos y tres textos políticos. De los textos periodísticos, uno trata del nombramiento de una nueva embajadora en EE.UU por parte de Biden, otro aborda el nombramiento de la periodista Pepa Bueno como directora del periódico *El País* y el tercero es una entrevista a las ministras de Igualdad de Francia y España, Élisabeth Moreno e Irene Montero, sobre cómo frenar la violencia machista. Los discursos políticos también son tres: uno lo pronunció la Presidenta de la República de Chile, Michelle Bachelet Jeria, al asumir el mando presidencial; el segundo es de la Presidenta de Argentina, Cristina Fernández, que pronunció para las trabajadoras sociales de Avellaneda en el Día Internacional de la Mujer y, finalmente, el último corresponde a una comparecencia del Presidente del Gobierno de España, Pedro Sánchez, también pronunciado con ocasión de la misma conmemoración. Pensamos que, a pesar de no ser muy extensa, la muestra es suficiente y significativa para ver cómo tratan los dos sistemas de traducción la cuestión del género. Una vez elegidos los textos, se han introducido en DL y GT respectivamente para luego proceder a su análisis, centrando nuestra atención en los elementos objeto de estudio.

## 5. ANÁLISIS DE LOS TEXTOS

Al analizar los textos y como se podrá apreciar al leer algunos de los fragmentos seleccionados, ambos sistemas traducen, en la mayoría de los casos, los nombres de profesiones con una preferencia hacia el masculino (*ambasciatore, sottosegratario*). En el primer texto analizado, se observa algo que volverá a presentarse, es decir, que ambos sistemas emplean, en un caso, el masculino y, en otro, el femenino para la misma profesión, como con *ambasciatore* y *ambasciatrice*. DL, además, se equivoca del todo con el género (*educato, finito*) pero, a diferencia de GT, emplea el

femenino para ‘senadora’. También puede apreciarse que, con respecto a la profesión de abogado, DL emplea el artículo femenino con el sustantivo masculino, en lugar de usar el poco frecuente *avvocatessa*, mientras que GT opta por el epiceno *forense*, aconsejado para el empleo de un lenguaje inclusivo. Además, a pesar de emplear el masculino para la función de *sottosegretario*, el participio *stata* ayuda a entender que se trata de una mujer, cosa que no ocurre en DL, que traduce con *ha anche servito come sottosegretario*.

<a href="https://elpais.com/internacional/2021-07-27/biden-elige-a-una-mujer-de-origen-dominicano-como-nueva-embajadora-de-ee-uu-en-espana.html">https://elpais.com/internacional/2021-07-27/biden-elige-a-una-mujer-de-origen-dominicano-como-nueva-embajadora-de-ee-uu-en-espana.html</a>	DL	GT Traductor
<p>Biden elige a una mujer de origen dominicano como <b>nueva embajadora</b> de EEUU en España [...]</p> <p>El presidente de Estados Unidos, Joe Biden, ha propuesto este martes a Julissa Reynoso como <b>nueva embajadora</b> de Estados Unidos en España y Andorra [...]</p> <p>Nacida en República Dominicana, <b>la diplomática</b> emigró a Estados Unidos cuando tenía siete años y su juventud transcurrió en la zona sur del Bronx (en la ciudad de Nueva York). <b>Formada</b> en la Universidad de Harvard (Massachusetts), <b>esta mujer de 46 años</b> pasó también por las aulas de la británica Cambridge y <b>acabó por perfilar</b> su currículum en Columbia (Nueva York). Su carrera en la política se inició en 2008 cuando participó en la campaña para la Casa Blanca de <b>la exsenadora Hillary Clinton</b>. Tras la derrota de la ex primera dama y la victoria de Barack Obama, Reynoso <b>ejerció la abogacía</b> en un gabinete jurídico. En marzo de 2012, <b>se convirtió en embajadora extraordinaria y plenipotenciaria</b> de EE UU en Uruguay. Dentro de la Administración de Obama, <b>también fue subsecretaria</b> de la Oficina de Asuntos del Hemisferio Occidental en el Departamento de Estado. [...]</p>	<p>Biden sceglie una donna di origine dominicana come <i>nuovo ambasciatore</i> degli Stati Uniti in Spagna [...]</p> <p>Il presidente degli Stati Uniti Joe Biden ha nominato martedì Julissa Reynoso come <i>nuovo ambasciatore</i> degli Stati Uniti in Spagna e Andorra [...]</p> <p>Nata nella Repubblica Dominicana, <i>la diplomatica</i> è emigrata negli Stati Uniti quando aveva sette anni e ha trascorso la sua giovinezza nel South Bronx (New York City). <i>Educato</i> all'Università di Harvard (Massachusetts), <i>il 46enne</i> ha studiato anche all'Università di Cambridge (Regno Unito) <i>ed è finito alla Columbia</i> (New York). La sua carriera in politica è iniziata nel 2008 quando ha partecipato alla campagna per la Casa Bianca <i>dell'ex senatrice Hillary Clinton</i>, e dopo la sconfitta dell'ex first lady e la vittoria di Barack Obama, Reynoso è <i>diventata avvocato</i> in uno studio legale e nel marzo 2012 è <i>diventata ambasciatrice straordinaria e plenipotenziaria</i> degli Stati Uniti in Uruguay. All'interno dell'amministrazione Obama, <i>ha anche servito come sottosegretario</i> dell'ufficio degli affari dell'emisfero occidentale al Dipartimento di Stato. [...]</p>	<p>Biden sceglie una donna di origine dominicana come <i>nuovo ambasciatore</i> degli Stati Uniti in Spagna [...]</p> <p>Il presidente degli Stati Uniti, Joe Biden, ha proposto questo martedì Julissa Reynoso come <i>nuovo ambasciatore</i> degli Stati Uniti in Spagna e Andorra [...]</p> <p>Nata nella Repubblica Dominicana, <i>la diplomatica</i> emigrò negli Stati Uniti quando aveva sette anni e trascorse la sua giovinezza nel South Bronx (a New York City). <i>Formatasi</i> alla Harvard University (Massachusetts), <i>questa donna di 46 anni</i> è passata anche per le aule della Cambridge britannica e <i>ha finito per plasma-re</i> il suo curriculum in Columbia (New York). La sua carriera in politica è iniziata nel 2008 quando ha partecipato alla campagna <i>dell'ex senatore Hillary Clinton</i> per la Casa Bianca. Dopo la sconfitta dell'ex first lady e la vittoria di Barack Obama, Reynoso <i>ha esercitato la professione forense</i> in un gabinetto legale. Nel marzo 2012 è <i>diventata ambasciatrice straordinaria e plenipotenziaria</i> degli Stati Uniti in Uruguay. All'interno dell'amministrazione Obama, <i>è stata anche sottosegretario</i> dell'Ufficio per gli affari dell'emisfero occidentale presso il Dipartimento di Stato. [...]</p>

También en el fragmento siguiente, se asiste a cierta oscilación en ambos traductores automáticos, ya que hay una alternancia de los términos *direttore/direttrice* dentro del mismo texto de llegada, con DL que traduce dos veces el término ‘directora’ en femenino (incluso con *vice*) y una vez en masculino, y GT que hace justo lo contrario, es decir, que emplea dos veces el masculino y una vez el femenino.

Por otro lado, la palabra ‘jefa’, que en italiano no tiene el femenino ‘capa’ o que como mucho se emplea en contextos informales, se convierte en otro epiceno en GT (*responsabile*) y se mantiene en masculino precedido por el participio femenino en DL (*è diventata capo*). En ningún caso se traduce con otro tipo de expresión, como podría ser *direttrice del servizio di informazione*.

<a href="https://elpais.com/comunicacion/2021-07-27/pepa-bueno-nueva-directora-de-el-pais.html">https://elpais.com/comunicacion/2021-07-27/pepa-bueno-nueva-directora-de-el-pais.html</a>	DL	GT Traductor
<p>Pepa Bueno, <b>nueva directora</b> de EL PAÍS</p> <p>Pepa Bueno (Badajoz, 1963), es licenciada en Periodismo por la Universidad Complutense de Madrid. Comenzó su carrera en los servicios informativos de Radio Nacional de España en Extremadura y, posteriormente, <b>fue jefa</b> de informativos en RNE en Aragón y en Madrid.</p> <p>En 1991 se incorpora a RTVE para dirigir el informativo regional, primero en Andalucía y luego en Madrid. En septiembre de 1996 <b>es nombrada subdirectora</b> del programa de actualidad <i>Gente</i> [...]</p> <p>En 2019 <b>es nombrada directora del informativo nocturno Hora 25</b>. Durante estos años, Pepa Bueno también ha colaborado en programas de Telecinco y Cuatro. [...]</p>	<p>Pepa Bueno, <i>nuova direttrice</i> di EL PAÍS</p> <p>Pepa Bueno (Badajoz, 1963) si è laureata in giornalismo all'Università Complutense di Madrid. Ha iniziato la sua carriera nei servizi giornalistici di Radio Nacional de España in Extremadura e poi è <i>diventata capo</i> delle notizie di RNE in Aragona e Madrid.</p> <p>Nel 1991 è entrata in RTVE per dirigere le notizie regionali, prima in Andalusia e poi a Madrid. Nel settembre 1996 è stata <i>nominata vicedirettrice</i> del programma di attualità <i>Gente</i> [...]</p> <p>Nel 2019 è stata <i>nominata direttore del programma di notizie serali</i> <i>Hora 25</i>. Durante questi anni, Pepa Bueno ha anche collaborato in programmi su Telecinco e Cuatro. [...]</p>	<p>Pepa Bueno, <i>nuovo direttore</i> di EL PAÍS</p> <p>Pepa Bueno (Badajoz, 1963), è laureata in giornalismo all'Università Complutense di Madrid. Ha iniziato la sua carriera nei servizi di informazione di Radio Nacional de España in Extremadura e, in seguito, è stata <i>responsabile</i> delle notizie di RNE in Aragón e Madrid.</p> <p>Nel 1991 entra in RTVE per dirigere il telegiornale regionale, prima in Andalusia e poi a Madrid. Nel settembre 1996 è stata <i>nominata vicedirettore</i> del programma di attualità <i>Gente</i> [...]</p> <p>Nel 2019 è stata <i>nominata direttrice del tg serale</i> <i>Hora 25</i>. In questi anni Pepa Bueno ha collaborato anche ai programmi Telecinco e Cuatro. [...]</p>

También en los últimos fragmentos del artículo sobre violencia contra las mujeres, se puede ver que hay una inclinación por defecto hacia el masculino para los sustantivos de profesión. En concreto, el término ‘ministras’ es declinado en femenino por DL (*le ministre*) mientras que GT opta por el masculino (*i ministri*). Esto no quita que incluso DL en el resto del documento traduzca siempre *ministro* cuando el término aparece en singular. Lo mismo ocurre con otras profesiones, como ‘jurista’ y ‘directora’, que DL traduce empleando el masculino (*avvocato, direttore*), mientras que GT opta una vez más por el epiceno, ya presente en el texto origen, *giurista*, y por el masculino, *direttore*. Además, GT mantiene el masculino y se equivoca con el pronombre (*riceverlo*), y DL reconoce el femenino (*accolta*) a pesar de emplear –una vez más, fluctuando entre masculino y femenino– el genérico *ministro* en la misma frase, si bien anteriormente había traducido empleando el plural femenino. El masculino aparece incluso en otros fragmentos y con verbos: ‘si estás segura’ se traduce como *se sei sicuro* en ambos sistemas; ‘una mujer ...se le va a informar’ como *sarà informata* en DL y *sarai informato* en GT.



<p><a href="https://elpais.com/sociedad/2021-07-27/nos-inspiramos-en-espana-para-combatir-la-violencia-contra-las-mujeres.html">https://elpais.com/sociedad/2021-07-27/nos-inspiramos-en-espana-para-combatir-la-violencia-contra-las-mujeres.html</a></p>	DL	GT Traductor
<p>“Nos inspiramos en España para combatir la violencia contra las mujeres” Las ministras de Igualdad de Francia y España, Elisabeth Moreno e Irene Montero, debaten en una entrevista en Madrid sobre cómo frenar la violencia machista [...] <b>La recibe la ministra de Igualdad española, Irene Montero</b> (Madrid, 33 años). [...] Moreno cree que la preparación de esos profesionales es crucial: “La primera toma de contacto es la más importante. Imagina que vas a una comisaría y te preguntan <b>si estás seguras</b> [sic] de que quieres denunciar, o te dicen que puedes desestabilizar a tu familiar o que es solo una bofetada... todavía se escuchan estas cosas. [...] <b>Moreno, jurista y directora</b> de varias empresas tecnológicas antes de desembarcar en el ministerio de Emmanuel Macron. [...] “No se trata tanto de que se pueda denunciar en una panadería, como de que cada vez más establecimientos como farmacias, centros comerciales o lugares privados puedan estar identificados como lugares seguros y que una mujer pueda saber en una situación de emergencia que en ese espacio <b>se le va a informar</b> de sus derechos”, explica Montero. [...]</p>	<p>“Ci ispiriamo alla Spagna per combattere la violenza contro le donne” Le ministre francese e spagnola per l’uguaglianza, Elisabeth Moreno e Irene Montero, discutono in un’intervista a Madrid su come arginare la violenza di genere [...] <i>Viene accolta dal ministro spagnolo per l’uguaglianza, Irene Montero</i> (Madrid, 33 anni). [...] Moreno cree que la preparación de questi professionisti sia cruciale: “Il primo contatto è il più importante. Immagina di andare in una stazione di polizia e ti chiedono <i>se sei sicuro</i> di voler sporgere denuncia, o ti dicono che potresti destabilizzare il tuo familiare o che è solo uno schiaffo in faccia... senti ancora queste cose. [...] Moreno, <i>avvocato e direttore</i> di diverse aziende tecnologiche prima di approdare al ministero di Emmanuel Macron [...] “Non si tratta tanto di poter denunciare una donna in una panetteria, ma piuttosto che sempre più stabilimenti come farmacie, centri commerciali o luoghi privati possano essere identificati come luoghi sicuri e che una donna possa sapere in una situazione di emergenza che <i>sarà informata</i> dei suoi diritti in quel luogo”, spiega Montero. [...]</p>	<p>“Ci ispiriamo alla Spagna per combattere la violenza sulle donne” <i>I ministri dell’uguaglianza di Francia e Spagna, Elisabeth Moreno e Irene Montero, discutono in un’intervista a Madrid su come fermare la violenza sessista [...]</i> <i>Ad riceverlo è il ministro spagnolo dell’Uguaglianza, Irene Montero</i> (Madrid, 33 anni). [...] Moreno ritiene fondamentale la preparazione di questi professionisti: “Il primo contatto è il più importante. Immagina di andare in una stazione di polizia e ti chiedono <i>se sei sicuro</i> di voler denunciare, oppure ti dicono che puoi destabilizzare un tuo familiare o che è solo uno schiaffo... queste cose si sentono ancora. [...] Moreno, <i>giurista e direttore</i> di diverse società tecnologiche prima di approdare al ministero di Emmanuel Macron. [...] “Non è tanto che possa essere segnalato in un panificio, quanto che sempre più esercizi come farmacie, centri commerciali o luoghi privati possono essere identificati come luoghi sicuri e che una donna possa sapere in una situazione di emergenza che in quello spazio <i>sarai informato</i> dei tuoi diritti”, spiega Montero. [...]</p>

A la hora de analizar los discursos políticos, se observan elementos parecidos. En concreto, en el fragmento del discurso de Bachelet, como en los anteriores, los sistemas traducen los términos ‘presidenta’ y ‘ministra’ en masculino. Además, puede verse cómo los sistemas traducen un elemento que suele emplearse a menudo cuando se utiliza un lenguaje inclusivo, es decir, la duplicación. A este respecto, se ha observado que DL –a diferencia de lo que ocurre en otros párrafos– elimina la duplicación siempre, mientras que GT, en el primer párrafo, la mantiene pero equivocándose, ya que repite dos veces el genérico *amici*; lo mismo ocurre con *Presidenti*, mientras que la reconoce en otro párrafo, cuando traduce *i suoi figli e le sue figlie*. Ambos sistemas traducen empleando el masculino genérico y cometiendo fallos por esto mismo. Incluso cuando el femenino está expresado



claramente ('soy hija'), DL lo traduce con la forma *sono figlio*, mientras que GT lo reconoce como femenino. En líneas generales, ambos sistemas tienden a eliminar las duplicaciones, como puede verse incluso en los 'niños, niñas y jóvenes' que se convierte en *i bambini e i giovani* en ambos sistemas, 'todas y todos' en *tutti* (una vez así en GT, otra en *di ciascuno e di tutti*), y 'chilenos y chilenas' en *cileni*.

<p><a href="http://archivospresidenciales.archivonacional.cl/index.php/discursos-de-michelle-bachelet-al-asumir-el-mando-presidencial">http://archivospresidenciales.archivonacional.cl/index.php/discursos-de-michelle-bachelet-al-asumir-el-mando-presidencial</a></p>	DL	GT
<p>Discurso de S.E. la <b>Presidenta</b> de la República, Michelle Bachelet Jeria, al asumir el mando presidencial</p> <p><b>Amigas y amigos:</b> Hace cuatro años atravesé esta puerta por la que hoy <i>he vuelto a entrar</i>. <b>Y ustedes me acompañaron</b> en esa despedida, y hoy vuelven a estar conmigo aquí, en ésta, la casa de <b>los Presidentes y las Presidentas</b> de Chile. [...] Yo <b>soy hija</b> de la educación pública, y mi compromiso es que en Chile todos tengamos esas mismas oportunidades. [...] <b>Soy médico, y fui ministra de Salud.</b> [...] Sé de primera mano cuáles son las preocupaciones de <b>las jefas de hogar</b>. Sé de la necesidad que cada mujer tiene de trabajar mientras sus <b>hijos e hijas</b> reciben estímulo y educación de calidad. [...] Y también de un país que trabaje incansablemente para que la dignidad y el respeto de los derechos sean la regla para <b>todos y todas</b>. [...] ¡Un programa que compromete una reforma educacional que asegure gratuidad, calidad, inclusión y mejores perspectivas para <b>los niños, niñas y jóvenes</b> de Chile! Un programa que compromete una gran reforma tributaria, que consagrará el principio de que quienes tienen más, contribuyan con más al bienestar de <b>todas y todos!</b> [...] <b>Amigas y amigos:</b> Voy a ser la Presidenta de todos los <b>chilenos y chilenas</b> [...].</p>	<p>Discurso di S.E. il <i>Presidente</i> della Repubblica, Michelle Bachelet Jeria, sull'assunzione dell'incarico presidenziale</p> <p>Amici: Quattro anni fa ho attraversato questa porta dalla quale <i>sono entrato</i> di nuovo oggi. E tu <i>mi hai accompagnato</i> in quell'addio, e oggi sei di nuovo con me qui, in questa casa <i>dei presidenti</i> del Cile. [...] <i>Sono un figlio</i> dell'educazione pubblica, e il mio impegno è che in Cile abbiamo tutti le stesse opportunità. [...] <i>Sono un medico e sono stato ministro della salute.</i> [...] So in prima persona quali sono le preoccupazioni <i>delle donne capofamiglia</i>. Conosco la necessità per ogni donna di lavorare mentre i <i>suoi figli</i> ricevono un incoraggiamento e un'educazione di qualità. [...] E anche un paese che lavora instancabilmente affinché la dignità e il rispetto dei diritti siano la regola <i>per tutti</i>. [...] Un programma che si impegna per una riforma dell'educazione che garantisca istruzione gratuita, qualità, inclusione e migliori prospettive per <i>i bambini e i giovani</i> del Cile! Un programma che si impegna a una grande riforma fiscale, che sancirà il principio che chi ha di più, contribuisce di più al benessere di <i>tutti!</i> [...] Amici: sarò il presidente di tutti i <i>cileni</i> [...].</p>	<p>Discurso di S.E. il <i>Presidente</i> della Repubblica, Michelle Bachelet Jeria, nell'assunzione del comando presidenziale</p> <p>Amici e amici: Quattro anni fa ho varcato questa porta da cui <i>sono rientrato</i> oggi. E <i>mi hai accompagnato</i> in quell'addio, e oggi sei di nuovo con me qui, in questa casa <i>dei Presidenti e dei Presidenti</i> del Cile. [...] <i>Sono figlia</i> della pubblica istruzione e il mio impegno è che in Cile abbiamo tutti le stesse opportunità. [...] <i>Sono un medico, ed ero ministro della Salute.</i> [...] So di prima mano quali sono le preoccupazioni <i>delle donne capofamiglia</i>. Conosco la necessità che ogni donna deve lavorare mentre i <i>suoi figli e le sue figlie</i> ricevono un'istruzione e un incoraggiamento di qualità. [...] E anche da un Paese che lavora instancabilmente perché dignità e rispetto dei diritti siano la regola <i>per tutti</i>. [...] Un programma che impegna una riforma educativa che garantisca gratuità, qualità, inclusione e migliori prospettive per <i>bambini e giovani</i> in Cile! Un programma che impegna una grande riforma fiscale, che sancirà il principio che chi ha di più, contribuisce di più al benessere <i>di ciascuno e di tutti!</i> [...] Amici: sarò il presidente di tutti i <i>cileni</i>: [...].</p>

Por lo que se refiere al discurso de Fernández, también en este caso aparece a menudo la duplicación; una vez más, tanto DL como GT omiten el femenino ('todos y todas' se convierte en *tutti*, 'argentinos y argentinas' en *argentini*, 'compañeros y compañeras' en *compagni* y *colleghi*). Sin embargo, ambos sistemas traducen *sindaci e sindache*, y en un caso GT traduce también la duplicación de *compagni e compagne*. Por lo demás, proporcionan solo el genérico de otras profesiones, como *lavoratori sociali* (DL) o *i miei partner di genere* (GT), y GT traduce en femenino el término 'compañeros'. Ningún sistema reconoce el pronombre femenino y ambos emplean el pronombre masculino plural *li* (*onorarli*). De la misma manera, emplean el masculino para el adjetivo 'orgullosa' y lo mismo ocurre con la expresión 'estoy segura'. La duplicación no puede darse en italiano con el pronombre de primera persona plural (nosotros, -as) ya que esta lengua dispone únicamente del pronombre *noi*. Sin embargo, el hecho de que ambos sistemas usen el pronombre indefinido *ognuno* refuerza el sentido de la frase, haciéndola más inclusiva.

<a href="https://www.casarosada.gob.ar/informacion/archivo/6228-homenaje-a-las-mujeres-en-avellaneda-palabras-de-la-presidenta-de-la-nacion-cristina-fernandez">https://www.casarosada.gob.ar/informacion/archivo/6228-homenaje-a-las-mujeres-en-avellaneda-palabras-de-la-presidenta-de-la-nacion-cristina-fernandez</a>	DL	GT
<p>Palabras de la Presidenta de la Nación Cristina Fernández, en el Acto de Homenaje al Día Internacional de la Mujer y a las Trabajadoras Sociales que llevan adelante el “Plan Vida”, en el Estadio de Independiente, en Avellaneda, Provincia de Buenos Aires. Gracias, muchas gracias, muy buenas noches <b>a todos y a todas</b>. Señor Gobernador de la provincia de Buenos Aires; señor Intendente de Avellaneda; <b>distintos Intendentes e Intendentas; compañeras y compañeros; amigos y amigas</b>: hoy quiero - haciéndome eco de esta convocatoria de Daniel - homenajear, en la figura de <b>las trabajadoras sociales, a mis compañeras de género: a las mujeres de todo el país</b>. Yo siento que en los momentos de mayor sufrimiento y de mayor dolor que atravesó nuestro país, fueron las mujeres -y esto <b>no significa no reconocer a los compañeros</b> [...]</p> <p>Pero yo quiero <b>honorarlas</b> dando de mí lo mejor que tenga porque yo creo que no hay mejor homenaje que desempeñarse con honor, con convicción, con patriotismo y con entrega absoluta en el lugar que cada uno de <b>nosotros o nosotras</b> tenemos. [...]</p> <p>Y si en algo estamos equivocados o tienen una idea mejor, que la traigan, que <b>estoy segura</b> que la podemos llevar adelante juntos. [...]</p> <p>Y yo me siento <b>muy orgullosa</b> de estar hoy aquí, de estar en este espacio. [...]</p> <p>Y nosotros, en este Bicentenario, en estos 200 años de historia, necesitamos también reeditar esa gesta de incorporar más <b>argentinos y argentinas</b> a nuestro espacio político. [...]</p> <p>Por eso, con mucho amor, quiero decirles a <b>todos y a todas</b> que estamos más unidos que nunca [...]</p> <p>Gracias, Daniel, por invitarme; gracias, <b>compañeros y compañeras</b> [...].</p>	<p>Parole del Presidente della Nazione Cristina Fernández, alla cerimonia per rendere omaggio alla Giornata Internazionale della Donna e agli operatori sociali che realizzano il “Piano Vida”, allo Stadio di Independiente, ad Avellaneda, provincia di Buenos Aires.</p> <p>Grazie, grazie mille, una buona serata a tutti voi.</p> <p>Signor governatore della provincia di Buenos Aires; signor sindaco di Avellaneda; <i>vari sindaci e sindache; colleghi; amici</i>: oggi vorrei - facendo eco all'appello di Daniel - rendere omaggio, nella figura <i>dei lavoratori sociali, alle mie colleghe di genere: alle donne di tutto il paese</i>.</p> <p>Sento che nei momenti di maggior sofferenza e dolore che il nostro paese ha attraversato, sono state le donne - e questo non significa che non riconosco i compagni [...]</p> <p>Ma voglio <i>onorarli</i> dando loro il meglio che ho perché credo che non ci sia miglior tributo che esibirsi con onore, con convinzione, con patriottismo e con dedizione assoluta nel posto che <i>ognuno di noi</i> ha. [...]</p> <p>E se ci sbagliamo su qualcosa o se hanno un'idea migliore, possono portarcela, e sono sicuro che possiamo portarla avanti insieme. [...]</p> <p>E mi sento <i>molto orgoglioso</i> di essere qui oggi, di essere in questo spazio. [...]</p> <p>Per questo, con molto amore, voglio dire a <i>tutti voi</i> che siamo più uniti che mai [...]</p> <p>Grazie, Daniel, per avermi invitato; grazie, <i>compagni</i> [...].</p>	<p>Osservazioni del Presidente della Nazione Cristina Fernández, nell'Atto di Omaggio alla Giornata Internazionale delle Donne e degli Assistenti Sociali che realizzano il “Piano di Vita”, allo Stadio Independiente, ad Avellaneda, Provincia di Buenos Aires.</p> <p>Grazie, grazie mille, buona serata a tutti voi.</p> <p>Lord Governatore della provincia di Buenos Aires; Sindaco di Avellaneda; <i>diversi sindaci e sindaci; compagni e compagne; amici</i>: oggi voglio - facendo eco a questo appello di Daniel - onorare, nella figura <i>delle assistenti sociali, i miei partner di genere: donne di tutto il Paese</i>.</p> <p>Sento che nei momenti di maggior sofferenza e dolore che ha attraversato il nostro Paese, sono state le donne - e questo non significa non riconoscere le compagne [...]</p> <p>Ma voglio <i>onorarli</i> dando il meglio che ho perché credo che non ci sia tributo migliore che esibirsi con onore, con convinzione, con patriottismo e con assoluta dedizione nel posto che <i>ognuno di noi</i> ha. [...]</p> <p>E se ci sbagliamo su qualcosa o abbiamo un'idea migliore, sollevale, <i>sono sicuro</i> che possiamo portarla avanti insieme. [...]</p> <p>E mi sento <i>molto orgoglioso</i> di essere qui oggi, di essere in questo spazio. [...]</p> <p>Pertanto, con tanto amore, voglio dire a <i>tutti</i> che siamo più uniti che mai [...]</p> <p>Grazie, Daniel, per avermi invitato; Grazie, <i>collegli</i> [...].</p>

También en el último fragmento, los términos que indican profesión y oficio no se traducen en femenino. La primera vez que aparece el término ‘ministra’, ambos sistemas lo reconocen como femenino, pero en el resto del texto aparece con su forma genérica, como también los demás títulos (delegadas, secretaria de Estado, directora general). Sólo el término ‘científicas’ se traduce en femenino en GT mientras que DL mantiene la forma genérica. Tampoco reconocen el adjetivo ‘todas’, que se traduce como *tutti* (*tutti i laboratori*), mientras que la expresión ‘la mujer trabajadora’ se reconoce como femenina (*donna lavoratrice* en DL y *lavoratrice* en GT). A pesar de la existencia del artículo femenino plural que antecede al epiceno ‘periodistas’, ni DL ni GT lo reconocen como tal (*ai giornalisti*), mientras que ‘las mujeres trabajadoras’ se traduce erróneamente como *del lavoro donne* (GT). Por lo que a la duplicación se refiere, en este caso, ambos sistemas la mantienen la mayoría de las veces (aunque ‘todos y todas’ es *tutti*, ‘amigos y amigas’ es *amici*), pero DL la traduce correctamente (*care amiche, cari amici*) mientras que GT repite dos veces el masculino (*cari amici, cari amici*). Sin embargo, ninguno de los dos reconoce el femenino en la expresión ‘Todos son bienvenidas y todos son bienvenidos’ (*Tutti sono benvenuti e tutti sono benvenuti*). De la misma forma, el ‘Consejo de Ministros y Ministras’ se traduce con *Consiglio dei Ministri* (DL) y *Consiglio dei ministri e dei ministri* (GT), mientras ambos sistemas reconocen ‘niños y niñas’ (*ragazzi e ragazze*). Por lo que al pronombre ‘ella’ se refiere, ninguno de los dos sistemas lo lee (‘porque ellas ya están implicadas’ / *perché sono già coinvolti*) y GT tampoco consigue ver el femenino del gentilicio ‘hondureña’.

https://www.lamoncloa.gob.es/presidente/intervenciones/Paginas/2021/prps08032021.aspx	DL	GT
<p>Comparecencia del presidente del Gobierno en el acto del Día Internacional de la Mujer Querida Irene, <b>querida ministra, queridas delegadas del Gobierno, secretaria de Estado y directora general</b>, querría empezar <b>felicitándoos a todas</b> por el 8 de marzo, hacerlo también, como ha hecho la <b>ministra de Igualdad, a todas las trabajadoras</b> del Ministerio de Igualdad. También a <b>las periodistas</b> que nos van a permitir el poder trasladar un mensaje tan importante [...] en favor de <b>las mujeres trabajadoras</b> [...].</p> <p>Lo que es evidente, querida ministra, <b>queridas amigas, queridos amigos</b>, es que la lucha feminista ha conseguido desde entonces innumerables logros. [...] Hoy es 8 de marzo, es el día de la <b>mujer trabajadora</b>. [...] El feminismo [...] debe ser la <b>lucha de todas y especialmente de todos</b>. No puede existir duda al respecto. [...] <b>Todos son bienvenidos y todos son bienvenidos</b>. [...]</p> <p>Y <b>todos y todas</b> debemos ser conscientes de que en esa gran transformación [...].</p> <p>Lo hemos hablado mucho la ministra y yo, lo hemos hablado mucho en el <b>Consejo de Ministros y Ministras</b> [...]. Hemos aprobado el Ingreso Mínimo Vital [...] contra una de las principales pobreza que más nos avergüenzan como sociedad, que es la pobreza infantil de <b>niñas y de niños</b>. [...] Es <b>para todas y para todos</b>. [...] Hagámoslo por Berta Cáceres, <b>la activista medioambiental hondureña</b> que fue asesinada de un tiro a manos de un sicario. [...]</p> <p>Por <b>todas ellas</b>, por los millones de mujeres que sufren cualquier tipo de discriminación, de opresión, de humillación. [...]</p> <p>Las <b>científicas</b> que fotografían agujeros negros en el espacio exterior. [...]</p> <p>Y en ese propósito justo debemos implicarnos <b>todas y todos</b> también, especialmente en los hombres, especialmente los hombres, <b>porque ellas ya están implicadas</b>.</p>	<p>Apparizione del Presidente del Governo alla cerimonia della Giornata Internazionale della Donna</p> <p>Cara Irene, <i>cara ministra, cari delegati del governo, segretario di Stato e direttore generale</i>, vorrei iniziare <i>congratulandomi con tutti voi</i> per l'8 marzo, e anche, come ha fatto il <i>ministro dell'uguaglianza</i>, con tutti i <i>lavoratori del ministero dell'uguaglianza</i>. Anche ai <i>giornalisti</i> che ci permetteranno di trasmettere un messaggio così importante [...] a favore <i>delle donne lavoratrici</i> [...].</p> <p>Ciò che è chiaro, cara ministra, <i>care amiche, cari amici</i>, è che la lotta femminista ha ottenuto innumerevoli successi da allora. [...] Oggi è l'8 marzo, il giorno della <i>donna lavoratrice</i>. [...]</p> <p>Il femminismo [...] è una lotta di e per le donne, ma deve essere la lotta di tutti e soprattutto di tutti. [...]</p> <p>Tutti sono <i>benvenuti e tutti sono benvenuti</i>. [...]</p> <p>E dobbiamo essere <i>tutti</i> consapevoli che in questa grande trasformazione [...].</p> <p>Il ministro ed io ne abbiamo parlato molto, ne abbiamo parlato molto nel <i>Consiglio dei Ministri</i> [...].</p> <p>Abbiamo approvato il Reddito Mínimo di Vita [...] contro una delle principali forme di povertà che più ci fa vergognare come società, che è la povertà infantile tra le <i>ragazze e i ragazzi</i>. [...] È <i>per tutti noi e per tutti</i>. [...]</p> <p>Facciamolo per Berta Cáceres, <i>l'attivista ambientale honduregna</i> che è stata uccisa da un sicario. [...]</p> <p>Per <i>tutte loro</i>, per i milioni di donne che soffrono qualsiasi tipo di discriminazione, oppressione e umiliazione. [...]</p> <p>Gli <i>scienziati</i> fotografano i buchi neri nello spazio. [...]</p> <p>E dobbiamo essere <i>tutti</i> coinvolti in questo giusto obiettivo, soprattutto gli uomini, soprattutto gli uomini, <i>perché sono già coinvolti</i>.</p>	<p>Apparizione del Primo Ministro all'atto della Giornata internazionale della donna</p> <p>Cara Irene, <i>cara Ministro, cari delegati di Governo, Segretario di Stato e Direttore Generale</i>, vorrei iniziare <i>congratulandomi con tutti voi</i> l'8 marzo, facendolo anche, come ha fatto il <i>Ministro della Parità</i>, a tutti i <i>lavoratori del Ministero di Uguaglianza</i>. Anche ai <i>giornalisti</i> che ci permetteranno di trasmettere un messaggio così importante [...] a favore <i>del lavoro donne</i> [...].</p> <p>Ciò che è evidente, caro Ministro, <i>cari amici, cari amici</i>, è che la lotta femminista da allora ha ottenuto innumerevoli risultati. [...]</p> <p>Oggi è l'8 marzo, è il giorno della <i>lavoratrice</i>. [...]</p> <p>Il femminismo [...] è una lotta collettiva [...] è una lotta per e per le donne, ma deve essere la lotta di tutti e soprattutto di tutti. Non ci possono essere dubbi al riguardo. [...]</p> <p>Tutti sono <i>i benvenuti e tutti sono i benvenuti</i>. [...]</p> <p>E dobbiamo essere <i>tutti</i> consapevoli che in questa grande trasformazione [...].</p> <p>Io e il ministro ne abbiamo parlato molto, ne abbiamo parlato molto in <i>Consiglio dei ministri e dei ministri</i> [...].</p> <p>Abbiamo approvato il Reddito minimo di vita [...] contro una delle principali povertà che più ci fa vergognare come società, ovvero la povertà infantile delle <i>ragazze e dei ragazzi</i>. [...] È <i>per tutti e per tutti</i>. [...]</p> <p>Facciamolo per Berta Cáceres, <i>l'attivista ambientalista honduregno</i> che è stata colpita e uccisa da un sicario. [...]</p> <p>Per <i>tutti loro</i>, per i milioni di donne che subiscono ogni tipo di discriminazione, oppressione, umiliazione. [...]</p> <p><i>Scienziati</i> che fotografano i buchi neri nello spazio. [...]</p> <p>E in quel giusto proposito dobbiamo coinvolgerci anche tutti, specialmente gli uomini, soprattutto gli uomini, <i>perché già sono coinvolti</i>.</p>

## 6. CONCLUSIONES

En este trabajo, hemos intentado ofrecer una idea de cómo dos de los principales motores de TA enfrentan problemas relacionados con la TA del género. Como se ha dicho, estos sistemas siguen mejorando día tras día, pero de nuestro análisis se desprende que no siempre consiguen reflejar el estado del mundo real ni del lenguaje actual, como tampoco consiguen entender siempre el contexto de forma correcta.

Pensamos que los ejemplos proporcionados para demostrar cómo solventan la cuestión del género son bastante representativos como primera aproximación al estudio de un asunto tan polémico y actual. De hecho, se han analizado fenómenos como la traducción de las profesiones laborales femeninas o los cargos, los pronombres, la duplicación, etc., y a menudo se ha observado que los sistemas no siempre reflejan la evolución –social y lingüística– que se ha producido en los últimos años.

A través de los ejemplos y su análisis, se ha visto que siguen existiendo unos valores masculinos predeterminados en ambos sistemas, a pesar de que sea un poco más evidente en DL que en GT, que prefiere emplear epicenos. Sin embargo, y a pesar de aplicar una política dirigida a evitar estereotipos de género, también GT sigue sin atribuirles el femenino a muchas profesiones, en especial cuando estas tienen que ver con ámbitos culturalmente atribuidos a los hombres. Este estado de cosas sugiere que los sistemas automáticos de traducción necesitan ser mejorados aún, tomando en consideración la actualidad del lenguaje natural y, por lo tanto, las cuestiones relacionadas con el género y el lenguaje inclusivo, y teniendo en cuenta, además, la nueva distribución laboral de las trabajadoras, que ocupan cargos cada vez más relevantes en la sociedad, hasta hace poco reservados casi exclusivamente a los hombres. Habrá que seguir entrenando los sistemas para que reflejen cada vez con más precisión e inclusión el mundo real.

- Álvarez de Miranda P. (2018) *El género y la lengua*, Madrid, Turner.
- Castilho S., Moorkens J., Gaspari F., Calixto I., Tinsley J. y Way A. (2017) "Is neural machine translation the new state of the art?", *The Prague Bulletin of Mathematical Linguistics*, 108, pp. 109-120, <https://ufal.mff.cuni.cz/pbml/108/art-castilho-moorkens-gaspari-tinsley-calixto-way.pdf> (consultado el 12/07/2021).
- Conde Noguero M. E. (2018) "Los traductores automáticos en línea como recurso metodológico en el aula de Español como lengua Extranjera". En López García, C. y Manso, J. (eds.), *Transforming education for a changing world.*, Eindhoven, NL, Adaya Press, pp. 304-312.
- Díaz Prieto P. (2012) "Luces y sombras en los 75 años de traducción automática". En Lanero Fernández J. J. y Chamosa J. L. (coords.), *Lengua, traducción, recepción: en honor de Julio César Santoyo*, Vol. 2, León, Publicaciones de la Universidad de León, pp. 139-175, <https://buleria.unileon.es/bitstream/handle/10612/4712/D%c3%adaz%20Prieto%20139-175.pdf?sequence=1&isAllowed=y> (consultado el 10/07/2021).
- Dinu G., Mathur P., Federico M. y Al-Onaizan Y. (2019) "Training neural machine translation to apply terminology constraints". En *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, Florence, Association for Computational Linguistics, pp. 3063-3068, <https://aclanthology.org/P19-1294.pdf> (consultado el 21/06/2021).
- Escudé Font J. y Costa-jussà M. R. (2019) "Equalizing Gender Bias in Neural Machine Translation with Word Embeddings Technique". En Costa-jussà, M. R., Hardmeier, C., Radford, W. y Webster K. (eds.), *Proceedings of the First Workshop on Gender Bias in Natural Language Processing*, Florence, Association for Computational Linguistics, pp. 147-154, <https://arxiv.org/pdf/1901.03116.pdf> (consultado el 05/07/2021).
- Farajian M. A., Bertoldi N., Negri M., Turchi M. y Federico M. (2018) "Evaluation of terminology translation in instance-based neural mt adaptation". En Pérez Ortiz, J. A. et al. (eds.), *Proceedings of the 21st Annual Conference of the European Association for Machine Translation: 28-30 May 2018*, Alacant, Universitat d'Alacant, pp. 149-158, <http://eamt2018.dlsi.ua.es/proceedings-eamt2018.pdf> (consultado el 21/06/2021).
- Frank A., Hoffmann C. y Strobel M. (2004) *Gender Issues in Machine Translation*, Bremen, University of Bremen, <http://citeseer.ist.psu.edu/viewdoc/download?doi=10.1.1.126.1237&rep=rep1&type=pdf> (consultado el 20/07/2021).
- Grijelmo Á. (2019) *Propuesta de acuerdo sobre el lenguaje inclusivo*, Barcelona, Taurus.
- Haque R., Hasanuzzaman M. y Way A. (2019) "Investigating terminology translation in statistical and neural machine translation: A case study on english-to-hindi and hindi-to-english". En *Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP 2019)*, Varna, INCOMA Ltd., pp. 437-446, <https://aclanthology.org/R19-1052.pdf> (consultado el 20/07/2021).
- Hasler E., De Gispert A., Iglesias G., y Byrne B. (2018) "Neural machine translation decoding with terminology constraints". En *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, Volume 2, New Orleans, Association for Computational Linguistics, pp. 506-512, <https://aclanthology.org/N18-2081.pdf> (consultado el 20/07/2021).



- Hutchins J. (1986) *Machine Translation: Past, Present, Future*, Chichester, Ellis Horwood Ltd, Publisher.
- Hutchins J. (1997) "First Steps in Mechanical Translation". En Teller, V. y Sundheim, B. (eds.), *Proceedings of MT Summit VI: Past, Present, Future*, San Diego, CA, pp. 14-23, <https://pdfs.semanticscholar.org/8241/607f07ca47751bf89e5f173158999d07f8a6.pdf> (consultado el 02/09/2021).
- Macketanz V., Avramidis E., Burchardt A., Helcl J. y Srivastava A. (2017) "Machine translation: Phrase-based, rule-based and neural approaches with linguistic evaluation", *Cybernetics and Information Technologies*, 17(2), pp. 28-43.
- Michon E., Crego J. M. y Senellart J. (2020) "Integrating Domain Terminology into Neural Machine Translation". En *Proceedings of the 28th International Conference on Computational Linguistics*, Barcelona, International Committee on Computational Linguistics, pp. 3925-3937, <https://aclanthology.org/2020.coling-main.348.pdf> (consultado el 18/07/2021).
- Monti J. (2017) "Questioni di genere e traduzione automatica". En De Meo, A., Di Pace, L., Manco, A., Monti, J. y Pannain, R. (eds.), *Al femminile. Scritti linguistici in onore di Cristina Vallini*, Firenze, Franco Cesati Editore, pp. 411-431.
- Monti J. (2019) *Dalla Zairja alla traduzione automatica. Riflessioni sulla traduzione nell'era digitale*, Napoli, Paolo Loffredo Editore Srl.
- Niño A. (2009) "Machine translation in foreign language learning: Language learners' and tutors' perceptions of its advantages and disadvantages", *ReCALL*, 21(2), pp. 241-258.
- Núñez Cabezas E. A. y Guerrero Salazar S. (2002) *El lenguaje político español*, Madrid, Cátedra.
- Prates M., Avelar P. y Lamb L. (2019) "Assessing Gender Bias in Machine Translation - A Case Study with Google Translate". En *Neural Computing and Application*, Londres, Springer London, <https://arxiv.org/pdf/1809.02208.pdf> (consultado el 18/07/2021).
- Rescigno A. A., Vanmassenhove E., Monti J. y Way A. (2020) "A Case Study of Natural Gender Phenomena in Translation. A Comparison of Google Translate, Bing Microsoft Translator and DeepL for English to Italian, French and Spanish", *CEUR*, vol. 2769, pp. 62-90, [http://ceur-ws.org/Vol-2769/paper\\_\\_26.pdf](http://ceur-ws.org/Vol-2769/paper__26.pdf) (consultado el 05/07/2021).
- Sánchez Ramos M. M. y Rico Pérez C. (2020) *Traducción Automática. Conceptos clave, procesos de evaluación y técnicas de posesición*, Granada, Editorial Comares.
- Sanz Romero M. (2021) "El sesgo de género del Traductor de Google persiste: Ella cose, él conduce" <https://computerhoy.com/noticias/tecnologia/sesgo-genero-traductor-google-persiste-ella-cose-conduce-834637> (consultado el 05/07/2021).
- Tertoolen R. C. (2012) "Desarrollos en la traducción automática: esperando aún una traducción de alta calidad", *Caracteres*, nov. 2012, Vol. 1, n. 2 [14], pp. 141-147 <https://gredos.usal.es/bitstream/handle/10366/124880/Caracteresvol1n2noviembre2012-traducion.pdf?sequence=1&isAllowed=y> (consultado el 05/09/2021).
- Vanmassenhove E., Hardmeier C. y Way A. (2019) "Getting Gender Right in Neural Machine Translation". En *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, October-November, 2018, Brussels, Association for Computational Linguistics, pp. 3003-3008, <https://arxiv.org/pdf/1909.05088.pdf> (consultado el 05/09/2021).
- Zhang H., Torres-Hostench O. (2019) "Cómo enseñar posesición de traducción automática a una segunda lengua: una propuesta didáctica para el aprendizaje de lenguas", *Revista Tradumàtica. Tecnologies de la Traducció*, 17, pp. 153-161, <https://doi.org/10.5565/rev/tradumatica.237> (consultado el 02/07/2021).





# The “vaccine race”: Metaphorical conceptualizations of the search of an immunization against Covid-19

GIULIA PEDRINI

Università di Trieste, Università di Udine  
[pedrini.giulia@spes.uniud.it](mailto:pedrini.giulia@spes.uniud.it)

## ABSTRACT

Metaphors have been shown to be paramount in health communication in the light of their capacity to make concepts more concrete and, consequently, simple. In the wake of the #ReframeCovid initiative (2020) and similar researches, which highlighted the ubiquity and relevance of figurative language during the coronavirus pandemic, this paper aims at providing a qualitative analysis of metaphors in mass media content with respect to the search of a coronavirus vaccine. The analysis was conducted from a contrastive perspective in Italian, English, German, and Russian. It pinpoints the most frequent source domains of such metaphors and provides an overview of the contexts in which they are used, i.e. the aspects of this phenomenon they frame, through practical examples. Finally, some considerations are made with respect to issues connected to the use of some of these metaphors, also from the translational viewpoint.

## KEYWORDS

Conceptual metaphors, health communication, coronavirus vaccine, mass media content

## 1. INTRODUCTION

The Covid-19 health crisis, an unprecedented event in contemporary history, is having an immeasurable impact on our lives and has highlighted the importance of effective communication. Health communication, however, raises several critical points in normal settings, let alone during a pandemic. Many definitions have been provided for the concept of health communication ever since it gained importance, starting from the 70's – when it reached the status of an independent academic field in the English-speaking areas. One of the most fitting ones is the following: health communication refers to the “conveyance and exchange of knowledge, experience, opinions and emotions which concern health or illness, prevention or treatment, health economics or politics” (Baumann & Hurrelmann 2014: 13, my translation). According to Signitzer (2001), it may be divided into four levels: (1) intrapersonal communication, (2) interpersonal communication (doctor-patient communication), (3) organizational communication (e.g. hospital press releases), (4) mass media or social communication (e.g. medicine journalism, communication campaigns, and TV series regarding medical topics). The present study focuses on this last level.

Information about Covid-19 has been ubiquitous in mass media ever since the beginning of the pandemic, so much that the term “infodemic”, coined during the SARS outbreak in 2003, has gained popularity. This neologism, a portmanteau of the words “information” and “epidemic”, is defined as “a rapid and far-reaching spread of both accurate and inaccurate information about something, such as a disease” (Merriam Webster 2020). Indeed, it is not at all easy to convey proper information about such topics. In this respect, an important role is played, among others, by figurative language.

### 1.1. METAPHORS IN HEALTH COMMUNICATION

The claim that figurative language may help the general public gain an insight into complex subjects like the current pandemic becomes clear if one considers that

[m]etaphors involve talking and thinking about one thing in terms of another, on the basis of perceived similarities or correspondences between them. As such, we use them to make sense of and communicate about new, complex, abstract and sensitive experiences in terms of more familiar, simpler and intersubjectively accessible ones (Semino 2020).

Ortony (1979), in his pre-cognitive view, had already rejected the non-constructivist perspective, which regarded metaphors as mere decorative means, and started considering them as a way to explain abstract and at times also obscure concepts (Colacci 2018: 40). In their well-known Conceptual Metaphor Theory, La-

koff and Johnson (1980: 5) went as far as to state that “the essence of metaphor is understanding and *experiencing* one kind of thing in terms of another” (my italics). Their whole hypothesis, which marked the beginning of modern considerations about metaphors and figurative language from a cognitive point of view,<sup>1</sup> is based on the assumption that “metaphor is pervasive in everyday life, not just in language but in thought and action” (ibid. 3) – as suggested already by the title of their book *Metaphors We Live By*. To provide an even clearer explanation of this theory, “[c]onceptual metaphors are defined as systematic sets of correspondences, or ‘mappings’, across conceptual domains, whereby a ‘target’ domain (e.g. our knowledge about arguments) is partly structured in terms of a different ‘source’ domain (e.g. our knowledge about war)”<sup>2</sup> (Semino 2008: 5). Other similar terms frequently occurring in metaphor studies and first introduced by Richards in *The Philosophy of Rhetoric* (1936) are “tenor” and “vehicle”:<sup>3</sup> the former refers to the concept, object, or person meant, while the latter is the image that actually “carries” the metaphor.

The use of metaphors and, in general, of figurative language is no occasional phenomenon in health communication, as may be inferred from the following excerpt from a paper written by medical experts:

In health communication, non literal meanings play an important role as they frame patient’s [sic] experience of illness, (positively or negatively) influencing their perception of therapy or even of themselves as individuals. [...] Metaphor is indeed a useful device to grasp an unknown (target) concept by using a known (source) concept, thus naming and explaining a phenomenon, i.e. illness, which otherwise would remain unintelligible, obscure (Ervás et al. 2016: 92).

This field has been widely explored by several linguists and scholars, such as Magris (2009; 2021) and Semino. The latter focused many of her studies on metaphors referring to cancer and end-of-life care (see for instance Semino, Heywood & Short [2004], Potts & Semino [2017], Demjen et al. [2016], and Semino et al. [2018]). Magris, instead, studied metaphors in the popularization of science, as they are a means to make complex technical terms and concepts comprehensible to a non-expert audience (cf. also Niederhauser 1999: 217-219), as well as in vaccination campaigns, where beside their explicative role they also have a persuasive function.

- 1 *Metaphors We Live By* may be regarded as “the foundation text for a new, ‘cognitively’ oriented theory of meaning” (Musolf 2012: 302).
- 2 The example refers to the ARGUMENT IS WAR metaphor, highlighted by Lakoff and Johnson in their book in order to clarify the influence that conceptual metaphors have on our everyday life. According to them, not only do we often use expressions which belong to this field (such as “indefensible claims”, “attack weak points”, “demolish arguments”, “win an argument”, “shoot down someone’s arguments”), but “many of the things we do in arguing are partially structured by the concept of war” (Lakoff & Johnson 1980: 4).
- 3 Some decades later, Black (1962) also talks about “topic” and “vehicle”.

## 1.2. METAPHORS AND COVID-19

Metaphors are a prominent feature in the pandemic discourse. This did not go unnoticed, and many researches have been recently devoted to this topic. Some examples are the studies carried out by Spina (2020), Semino (2020), Olšvang (2021), Komatsubara (2021), Signorini & Neves de Brito (2021), Musolff (2021), and Liu & Lu (2021). Furthermore, an interesting initiative concerning metaphors which deal with the pandemic is #ReframeCovid (2020), a collection of metaphors in several languages which attempts at providing the public with source domains to conceptualize Covid-19 that differ from the ones referring to *WAR*. As a matter of fact, it has been shown that a considerable number of metaphors used in discourse regarding the pandemic stem from the military field. Such conceptualizations occur in several contexts: according to Antonelli (cf. Forderi 2020), military metaphors have been used mainly by politicians in order to draw people's attention to the ongoing emergency and thus justify the measures to be taken and sacrifices to be made. However, this raises many critical points, as highlighted by Semino:

[w]ar metaphors, which were widely used at the start of the pandemic, have been criticised for inappropriately personifying the virus as a malevolent opponent, creating unnecessary anxiety, dangerously legitimising authoritarian governmental measures, and implying that those to die did not fight hard enough.<sup>4</sup> Research has shown that war metaphors can actually be useful in some contexts (for example, to convey the need for urgent collective effort), but they can also discourage self-limiting behaviours, such as refraining from our usual activities and just staying home (Semino 2020).

## 1.3. ANALYZED MATERIAL

In the light of the major role played by metaphors in the pandemic discourse, I decided to focus on a specific subfield which has not been widely explored yet, i.e. mass media content concerning the attempts to produce a coronavirus vaccine. For this purpose, mostly articles – but, in limited cases, also online videos of TV channels and podcasts – dating from the beginning of the pandemic until January 2021 were analyzed. The languages examined are Italian, English, German, and Russian.<sup>5</sup> Since metaphorical conceptualizations often pervade illustrations, as well, the analysis will be multimodal and include them, too – as do many other metaphor studies, like the aforementioned initiative #ReframeCovid.

4 The same was observed with respect to cancer and end-of-life care, which, according to some scholars, should not be discussed in military terms (cf. Miller 2010: 20; Semino, Demjen & Demmen 2018).

5 Short phrases were transliterated according to the scientific transliteration of Cyrillic, while longer excerpts were translated altogether, so as not to limit their interest to the field of Slavonic studies.

## 2. SOURCE DOMAINS

Some of the most significant metaphors identified will be listed below and categorized according to their source domain: *inter alia* races in general, horse and car races, hunting, strategy games, gambling, journey, war, mythology and folklore, as well as historical events.

### 2.1. RACES



“Welches Forscherteam macht das Rennen?” in *ZDFheute* (2020)

The most frequent conceptualization undoubtedly sees the SEARCH OF A VACCINE in terms of a RACE: this is shown clearly by noun phrases such as “corsa al vaccino” or “corsa alla sperimentazione” in Italian, by compound nouns such as “Impfstoff-Rennen” and “Impfstoff-Wettlauf” in German, as well as “vaccine race” or “race to the vaccine” in English, and “gonka vakcin” or “gonka za privivkoj” in Russian. Furthermore, it may also be inferred from the frequent verb phrases in metaphorical scenarios framing the competitive nature of races, such as “tenere testa ai leader”, “recuperare”, “seinen Vorsprung verlieren”, “be close behind”, or “dyšat’ v spinu lideru”, just to mention a few. Beside these verb phrases, other noun phrases often highlight the competitiveness of the search: “Despite the advantage of being able to ship and store Moderna’s vaccine using ordinary refrigeration, it’s too early to declare one version a **winner**. That’s because these **frontrunners** might perform differently depending on the population’s demographics” (Richards 2020).<sup>6</sup> Similar is also the instance: “Covid-19 – Risultati in-

<sup>6</sup> Bold letters in the instances added by the author of the article to better highlight metaphors in long excerpts.

coraggianti, ma ancora nessuno **sul podio nella ricerca** per immunizzarsi contro il virus sempre più contagioso” (Buletti 2020).

In race metaphors, references to specific phases of the competition are quite frequent, as is the case in “Corsa al vaccino per Covid-19, quanto manca al **traguardo?**” (Bidetti 2020) or “Covid, piano vaccini **ai blocchi di partenza**” (Tgcom24 2020). In German, examples could be found only with reference to the finishing line,<sup>7</sup> such as “Corona-Impfstoffe: Die **Ziellinie** im Blick?” (Müller 2020).

An interesting feature about such metaphors is that they frame different aspects of the same issue depending on their use: at times the opponent is just the virus, see “[i]m **Wettkampf gegen den tödlichen Erreger** sollen nun Menschen experimentelle Wirkstoffe verabreicht werden” (Schlak 2020). In other frequent phrases, such as “corsa contro il tempo” or “Wettkampf gegen die Zeit”,<sup>8</sup> the need to act hastily is highlighted, although the real opponent is clearly still the virus itself.

The search of a vaccine not only has medical implications, but it is also a chance for countries to affirm themselves politically and economically, hence the frequent expressions such as “corsa geopolitica” or “gara geopolitica”. Similarly, *The Washington Post* stated as follows:

“[t]his is **not a gentlemanly stroll in the park** by a bunch of people who all agree that there’s some common public good we all need to strive for,” said J. Stephen Morrison, director of the Center on Global Health Policy at the Center for Strategic and International Studies in Washington. “This has become a **geopolitical race**, and it’s one that’s seen as tied to domestic stability and support amid lots of adversity.” (Khurshudyan 2020).

Beside the conceptualizations mentioned so far, where the source domain consists in races in general, it was possible to identify a large number of metaphors which refer to more specific types of races and thus frame different aspects of reality: HURDLE RACES, RELAY RACES, SPRINT RACES, and MARATHONS. In an article titled “All eyes on a hurdle race for a SARS-CoV-2 vaccine”, this theme is addressed several times to discuss the difficulties faced by scientists and researchers:

The good news is that the **final hurdle on the way to the finishing line** – the completion of a properly controlled phase III clinical trial – is in sight. Ideally, this process will not be jeopardized by a premature rush, through an Emergency Use Authorization by the US Food and Drug Administration or other international regulators, to get a vaccine into use in the clinic before the trial has generated sufficient safety and efficacy information. As in any **hurdle race**, skill, speed and judgement are all needed to successfully and safely cross the **finishing line** (Gaebler & Nussenzweig 2020).

7 See also the excerpt from the article by Gaebler & Nussenzweig (2020) below for further examples in English.

8 Here, the TIME IS A MOVING OBJECT metaphor identified by Lakoff and Johnson (1980: 41-45) may be found as well.

The attempt to produce a vaccine may be conceptualized as a *RELAY RACE*, too. In this case, the metaphor frames the cooperation among scientists which is needed for the search to be successful:

Scientists worldwide are working against the clock to find a viable coronavirus vaccine – but **are corners being cut** for the sake of political gain and profit?

To begin with, it felt like **a sleek performance from a well-honed relay team**. On 11 January, only 10 days after reporting a new respiratory disease, the Chinese published the genome sequence of the virus that causes it. Researchers around the world set to work building vaccines against Covid-19, as the disease became known, and the first candidate entered human trials on 16 March; it was joined, as the months passed, by dozens of others. Scientists were jubilant, and they had every right to be. They'd **broken all vaccinology records** to get to that point (Spinney 2020).

Similar is also the following instance: “The **race for a vaccine** will continue at ‘warp speed’ – but it’s **not an individual sprint**, it’s a **relay race**. **The baton will shortly be passed** to distribution logistics and patient compliance by citizen-patients” (Pitts 2020). Furthermore, both articles are multimodal texts, highlighting the conceptualization through illustrations, as well:



Illustration on the left by James Melaugh, published in *The Guardian* (Spinney 2020); illustration on the right by Linas Garsys (*The Washington Times*) published in the article by Pitts (2020).

The excerpt by Pitts also introduces the expression *sprint*, which is obviously used to underline the speed at which researchers were working; this characteristic is inherent in the term “race”, too, but is even more evident here. See, for instance, “Der **Impfstoff-Sprint**. Es dauert meist viele Jahre, einen Impfstoff zu entwickeln. Die Corona-Pandemie zwingt Unternehmen, Institute und Behörden zu neuen Wegen” (SRF 2020) or “Nello **sprint per il vaccino**, Moderna **tenta il sorpasso**” (Dusi 2020).

However, many view the search of a vaccine in the opposite way, since – as pointed out by SRF 2020 – it takes a long time to achieve such results: hence the frequent metaphors referring to *MARATHONS*. See, for instance: “Warum die Impf-



stoffentwicklung gegen das Coronavirus so lange dauert. Die Entwicklung neuer Impfstoffe ist **ein Marathon, kein Sprint**. Das ist auch gut so. Fatal wäre, wenn am Ende etwas gespritzt wird, das nicht hilft” (Diehm 2020). Similar considerations were made in the following Swiss Italian article with the title “La maratona del vaccino”:

Sulle misure per contrastare il Coronavirus subentra il discorso della ricerca di un vaccino dalla provata efficacia. Le notizie si susseguono e, **più che uno sprint, la corsa** intrapresa da decine di laboratori in tutto il mondo nella ricerca di un vaccino contro il Covid-19 appare come una **maratona**. È oramai palese che il processo di ricerca di un vaccino efficace e sicuro necessita tempi lunghi, così come, a giusta ragione, è lungo il processo di concessione delle licenze, anche in Svizzera (Buletti 2020).

In another article, the contrast between sprint races and marathons is employed in order to frame the logistical issues posed by the vaccination campaign:

Das größte Problem könnte jedoch aus seiner Sicht beim Personal drohen. Die Impfungen können von medizinischen Fachkräften vorgenommen werden, aber es müssen auch immer Ärzte dabei sein, etwa für den Fall, dass Nebenwirkungen auftreten. Und in der aktuellen Corona-Situation würde das medizinische Personal in Kliniken und Praxen bereits “an der Kapazitätsgrenze arbeiten”, sagt Klumpp. “Da ist aus unserer Sicht eigentlich der größte Engpass”, sagt Klumpp. Es müsse jedem klar sein, das sei “**kein 100-Meter-Sprint, das ist ein Marathon**” (Baars & Lambrecht 2020).

In contrast to these examples, which highlight the opinion that the process of finding and distributing a coronavirus vaccine would take a long time, marathon metaphors have the opposite function in the following instance: “Ugur Sahin, CEO der Mainzer Firma BioNTech, erklärt, wie **der jahrzehntelange Marathon der Impfstoffentwicklung** auf ein **4 x 100 m Staffelrennen** reduziert wird” (Müller 2020). An analogous example – albeit not concerning marathons, but 10,000-M RUNS – is the following:

Das wäre wie **ein 10000-Meter-Lauf in 8 Minuten**

Was es hieße, wenn tatsächlich noch im nächsten Jahr ein Impfstoff gegen das neuartige Coronavirus zugelassen würde, illustriert eine Metapher: Ein trainierter 30-jähriger Freizeitsportler läuft 10000 Meter in rund 55 Minuten – zumindest ist das die Anforderung des Deutschen Sportabzeichens. Der Äthiopier Kenenisa Bekele hat bei seinem Weltrekord für diese Strecke noch nicht einmal halb so lang gebraucht, nämlich 26 Minuten und 18 Sekunden. Und nun soll es auf einmal einen Menschen geben, der diese Strecke in 8 Minuten läuft? So unmöglich das klingt – bei der Impfstoffentwicklung ist solche eine immense Steigerung der Schnelligkeit zumindest denkbar (Frick 2020).

Finally, the following marathon metaphor does not present the production of a vaccine as the finishing line, but as a mere intermediate step towards normality: “Ein wirksamer Impfstoff wäre ein **Meilenstein**, aber noch lange nicht die **Zielgerade** auf dem **Marathon zur Normalität**” (Addendum 2020).

## 2.2. CAR AND HORSE RACES

Other conceptualizations are the SEARCH OF A VACCINE IS A CAR RACE and the SEARCH OF A VACCINE IS A HORSE RACE. In the present categorization of metaphors, these fields were separated from other races in the light of their marked differences – while all types of races in 3.1 undoubtedly retain more similarities. An instance is the following excerpt: “[i]n molti si sono domandati perché i dirigenti di una azienda **in pole position nella ricerca di un vaccino** contro il Covid inondassero il mercato con le loro azioni” (Paudice 2020). This metaphor, originally pertaining to the field of horse races, is nowadays used mainly to refer to car races. Another metaphor ascribable to this category, having horse races as a source domain, is the following. *To nobble* is a British English verb which means “to incapacitate (a racehorse) especially by drugging” (Merriam Webster):

[T]ensions began to surface among the team members [of scientists], and lately even the most distracted spectator will have noticed that they appear to be trying to **nobble each other** openly **on the track**. With accusations that the Russians and Chinese hacked research groups in other countries, biotech executives criticised for cashing in on their own, as yet unapproved vaccines, and Russia approving a vaccine that is still in clinical trials, the quest for a vaccine seems to have turned sour (Spinney 2020).

Finally, another example with horse races as a source domain is the following: “The UK has bought six of the hundreds of vaccines under development. It has two of the three companies **heading down the final furlong** – AstraZeneca’s and Pfizer’s. Bingham says she thinks there is a chance of a vaccine before Christmas” (Boseley 2020). *Furlong* is a measurement unit used mainly in horse races, so in this excerpt *the final furlong* refers to the last phases of a race and may be compared to phrases like “die Ziellinie im Blick” from 2.1.

## 2.3. HUNTING

Another quite effective metaphor conceptualizes the SEARCH OF A VACCINE AS a HUNT. In this case, the German language shows a remarkable predilection for this domain, with a variety of noun phrases such as “Jagd nach dem Impfstoff”, “Jagd nach Corona-Impfstoff”, or “Jagd nach dem Corona-Gegenmittel”. “Die Jagd nach dem Impfstoff” was even the title of a podcast by the NDR television channel which had the aim to keep people updated about the latest research developments. This expression can also be found on an illustration<sup>9</sup> presenting the NDR web page dedicated to this topic:

9 This picture also contains other conceptualizations, such as the widely discussed SEARCH OF A VACCINE seen in terms of a RACE OR even as a CLIMB, highlighting the difficulties and obstacles faced by scientists.



The same conceptualization may be inferred from a video by the German channel ZDF (2020): “Auch andere **sind dem Virus auf der Spur**, wie Professorin Marylyn Addo vom Deutschen Zentrum für Infektionsforschung in Hamburg”.

With respect to the English language, an example is “Novartis jumps into **Covid-19 vaccine hunt**, as Big Pharma and big biotech commit to billions of doses” (Mast 2020), whereas for Italian “caccia al vaccino anti-covid” was found in an article published by *Il Giorno*. However, it may be interesting to note that most phrases like “caccia al vaccino” occurring in Italian newspapers in 2020 actually refer to the distribution of flu vaccines before the winter season and their scarcity in some regions.

Finally, a Russian noun phrase, “ochota na Covid-19” (Strana.ua 2020), is deemed to be of interest, since it implies a change of perspective from most aforementioned metaphors (except for the German example with the expression “dem Virus auf der Spur sein”). As a matter of fact, here the prey to be hunted by scientists is no longer the vaccine, but the virus itself.<sup>10</sup>

#### 2.4. GAMES

A fourth interesting metaphor depicts the SEARCH OF A VACCINE AS A GAME – usually, a strategy one. An example is the reference to chess in the following excerpt from *La Repubblica*:

<sup>10</sup> As shown by the article title, which may be translated as “When will a vaccine be created and by whom”, the topic is still the vaccine.

Il titolo della biotech americana impegnata contro il coronavirus [Moderna] vola in Borsa. La settimana scorsa Pfizer aveva annunciato il successo nel 90% dei casi. Pronti tra poche settimane i dati di AstraZeneca, a gennaio quelli di Sanofi e Johnson&Johnson. La **scacchiera** comincia a riempirsi. Presto **muovere tutti i pezzi** richiederà **strategia**. [...] Due giorni dopo [l'11 novembre] era arrivata la Russia, con il 92% per il suo Sputnik 5, che al momento però non è destinato a **giocare sul terreno europeo** (Dusi 2020).

In another instance, found in the newspaper *Il Resto del Carlino*, the vaccine distribution is conceived of in terms of a REBUS: “Vaccino Coronavirus, nelle Marche resta un rebus. La Protezione civile al lavoro per predisporre i luoghi dove conservare le circa 40mila dosi. Ancora nulla su dove e come saranno somministrate” (Curzi 2020). Finally, an illustration in the newspaper *Der Spiegel* represents a dartboard in the shape of a coronavirus and the dart as a vaccine syringe:



Illustration in *Der Spiegel* (Weber 2020), by KrizzDaPaul / Getty Images

## 2.5. GAMBLING

Another – in this case not really frequent, but still expressive – conceptualization is the one that refers to THE SEARCH OF A VACCINE AS GAMBLING, thus underlining its unpredictability. An example is to be found in the following excerpt, from the Russian news agency TASS, where the global efforts to invent a vaccine are depicted as a lottery:

[...] observing the simultaneous efforts of several countries to find a vaccine, it is more correct to talk about **a lottery rather than a race**. Each team has its own **ticket**, and its winning does not depend on its speed or abilities as much as on fortunate circumstances. Here it is worth noting that the more teams are playing, the more chances

there are to be successful soon. From the point of view of business, elaborating a vaccine is a risky undertaking. In fact, only a few candidates will be able to make it to clinical trials (Beljaev 2020).

Furthermore, a German article from ZDF (Hickmann 2020), discussing the distribution of the vaccine doses, mentions “der unfaire Poker um Impfstoffdosen” in the attempt to highlight that poorer countries are doomed to have more difficulties buying the vaccine, which will exacerbate social differences. Another instance is offered by an Italian article (Bidetti 2020), which states that “[l]a somministrazione di massa di un vaccino – quale che sia – senza aver portato a termine tutte le fasi della sperimentazione è un vero e proprio **azzardo** sulla pelle delle persone”. Here, however, the target domain is slightly different, since it is the VACCINE itself rather than its SEARCH, thus implying skepticism with regard to vaccination. Finally, the English play on words “a shot in the arm or a shot in the dark” (Sheen 2020) may be ascribed to the latter mapping, too, based on the meaning of *shot* as a vaccination and the idiom *shot in the dark* that – although not pertaining to gambling – denotes something arbitrary, subject to chance, exactly like gambling.

## 2.6. WAR

Some of the most frequent metaphors concerning the pandemic are – as mentioned in 2.1 – to be ascribed to the military field. Some conceptualizations are THE PANDEMIC IS WAR, THE VIRUS IS AN ENEMY, THE VACCINE IS A WEAPON and THE SEARCH OF A VACCINE IS WAR. So, in the instances “Der Impfstoff, den ‚Biontech‘ aus Deutschland und ‚Pfizer‘ aus den USA gemeinsam entwickeln, hat am Montag Optimismus verbreitet, dass die Pandemie bald **besiegt sein** könnte” (Dittrich 2020) or “Non ci sono ancora terapie capaci di **vincere il Covid**” (Verdelli 2020), the verbs *besiegen* and *vincere* imply a battle with winners and losers. Another example is the following title, referring to pharmaceutical companies attempting to find a vaccine: “Welche Pharmafirmen sich am Corona-Abwehrkampf beteiligen” (Ehrhardt 2020). Similar is also another Italian excerpt:

“Per dare una svolta è necessario che i vaccini non proteggano solo dai sintomi della malattia o dalle sue forme più gravi ma che siano anche in grado di bloccare la trasmissione del virus”. Così la giornalista e divulgatrice scientifica bergamasca Roberta Villa, laureata in medicina e chirurgia, evidenzia quello che potrà rappresentare uno snodo cruciale nella **lotta all’epidemia da Covid-19** (Ghisleni 2020).

Another instance highlighting the frequently military nature of pandemic mappings is to be found in an Italian article reporting an interview with Maria Rosaria Capobianchi, the virologist who isolated the coronavirus, which describes her undertaking as “prendere prigioniero ‘il nemico’” (Verdelli 2020).

Furthermore, a whole subclass of metaphors consists in spatial ones, i.e. those which describe where war is being waged, such as: “azienda **in prima linea** nella produzione di un vaccino contro il Covid” (Paudice 2020) or “L’Emilia Romagna è in prima linea contro il Covid non solo **sul fronte della cura**, ma anche sotto il profilo della ricerca. E ricerca oggi vuol dire soprattutto vaccino” (Il Resto del Carlino 2020a). The *front* is also mentioned in an article about Moderna’s vaccine, which states that “[i]t’s still early days, but the news from the COVID-19 vaccine front continues to bring much-needed hope to a virus-ravaged world” (Richards 2020). Besides, the aforementioned article reporting the interview with Maria Rosaria Capobianchi refers to her as

[u]na **veterana** molto autorevole, che non ha partecipato alla recente e permanente giostra degli esperti in tv, ma che ha **combattuto nella trincea della ricerca molte battaglie insidiose**, dalla prima Sars del 2002 all’influenza suina del 2009, fino all’Ebola” (Verdelli 2020).

Another interesting subclass of war metaphors is the one referring to WEAPONS. The most general instances are: “Weltweit suchen Experten nach einem Impfstoff, etwas, was die Pandemie aufhalten kann. [...] Die Forscher sind optimistisch, **eine Waffe gegen das neuartige Coronavirus** bauen zu können” (ZDF 2020) or “Der Impfstoff CoronaVac des chinesischen Herstellers Sinovac Biotech gilt als vielversprechendes **Mittel im Kampf gegen die Corona-Pandemie**” (FAZ 2020). An analogous instance is the Russian “sredstvo dlja bor’by s COVID-19”, referring to the attempt to find a vaccine made by researchers in the EU, USA, Australia, Israel, China, Japan, and Russia (TASS 2020). Another general example is provided by the Italian *Il Sole 24 Ore* with the following excerpt:

Tra gli **avamposti** mondiali **nella lotta al Covid** c’è un pezzo di Centro Italia. Nel Lazio e in Toscana si lavora infatti alle due **armi** più attese in questa **battaglia contro il virus** che ha messo in ginocchio tutto il mondo: innanzitutto il vaccino che vede collaborare insieme, a due passi da Roma, piccole biotech company e colossi farmaceutici. [...] (Bartoloni & Pieraccini 2020).

A more specific instance is “La Regione [Emilia-Romagna] **schiera l’artiglieria pesante**: il Policlinico Sant’Orsola di Bologna e l’azienda ospedaliera-universitaria di Modena” (Il Resto del Carlino 2020a).

As far as vaccination itself is concerned, in a German article in the *Augsburger Allgemeine* it is defined as an “Impf-Offensive” (Lange 2020). Also an Italian article from *La Repubblica*, while discussing vaccine administration, employs army metaphors, such as “Giovani laureati, infermieri, assistenti sanitari: la **carica** dei 20mila vaccinatori”.

A last interesting subclass of war metaphors is the one which frames the competition between countries, as in 2.1. A podcast by *Die Zeit* is titled “Der **globale Impfkrieg**: Wie China mit Spritzen seine Macht ausbaut” (Brost & Grabitz 2020).

Similarly, an article by the Russian agency MBK mentions a “vojna vakcin” between Sputnik, Moderna, and other vaccines elaborated by several pharmaceutical companies (Tajzhetinova 2020). Finally, an article title by *Il Corriere della Sera* reads “Vaccino Covid, la **battaglia** tra Pfizer e Moderna per la ‘rivoluzione mRNA” (Modeo 2020).

Finally, in German an interesting metaphor which could be added to this category is the following: “Wann kommt **der Corona-Killer**?”. This instance may also remind of terms such as “natürliche Killerzellen”, which entail a metaphorical connotation – that is also quite frequent in immunology (cf. Magris 2021).

## 2.7. JOURNEY

One of the conceptual metaphors *par excellence*, discussed in Lakoff and Turner (1989), is LIFE IS A JOURNEY.<sup>11</sup> Similarly, a frequent conceptual mapping in mass media narration in this field is THE SEARCH OF A VACCINE IS A JOURNEY. Some of the most evident instances are the following: “I risultati dello studio segnano un passo importante in questo storico **viaggio di 8 mesi per un vaccino**” (Il Messaggero 2020), “Pfizer e BioNTech hanno fatto sapere di essere **‘sulla buona strada**’ per sottoporre il loro vaccino alla *regulatory review* già in ottobre” (Paudice 2020), “Le prime dosi del vaccino targato Pzier-BioNTech arriveranno in Europa all’inizio del 2021, ma ‘bisogna essere chiari con i cittadini: **la strada da percorrere è ancora lunga**” (Bresolin 2020), or the German title “**Der weite Weg** zur Corona-Impfung” (Lambrecht & Baars 2020).

Other metaphors ascribable to this source domain are: “È vero, quindi, che questi vaccini sono soggetti allo stesso iter degli altri, **non sono state prese scorcioie** e avranno le stesse garanzie di un vaccino nuovo, appena uscito sul mercato di un altro tipo” (Ghisleni 2020) and the English equivalents “**are corners being cut** for the sake of political gain and profit?” and “**Short cuts** in testing for vaccine safety and efficacy endanger millions of lives in the short term” (both Spinney 2020).

Some interesting conceptualizations concern the obstacles that may be found on a journey: the title of an article published in *The Economist*, for instance, reads “The **path** to a Covid-19 vaccine may be quick, but it will be **bumpy**” (Carr 2020). An article in *Die Zeit*, instead, states that “[e]rst vor zehn Tagen berichtete die Nachrichtenagentur Reuters, dass mehrere Moskauer Kliniken wegen **Engpässen** beim Impfstoff kurzzeitig keine neuen Freiwilligen mehr für die Tests aufnehmen konnten” (Kireev 2020).

Another interesting example in Italian frames competition between countries. After pointing out that research in China had slowed down as cases of coronavirus infection were decreasing, but that it gained momentum again after the

11 See also, *inter alia*, the AN ARGUMENT IS A JOURNEY metaphor in Lakoff and Johnson (1980).



progress of Pfizer, Moderna, and AstraZeneca, the article states as follows: “I cinesi insomma non vogliono **perdere il treno** dopo gli annunci di Pfizer e Moderna che hanno suscitato molto interesse degli investitori” (Il Sole 24 Ore 2020).

## 2.8. MYTHOLOGY AND FOLKLORE

An interesting source domain consists in MYTHOLOGY and FOLKLORE. A frequent metaphorical expression is “panacea”, mainly used in negative contexts, with the aim to warn readers that vaccines may not mark the end of the pandemic. See: “Kyriakides: “Un solo vaccino non sarà la panacea. La strada è lunga” (Bresolin 2020), or “Die Impfungen gegen Covid-19 sind ein großer Hoffnungsträger im Kampf gegen die Pandemie, jedoch kein Allheilmittel” (Rohwedder 2021). In an article title, a Russian Ukrainian newspaper goes as far as to wonder whether the Russian vaccine will be a panacea or a poison (Iščenko 2020).

Mythological metaphors are also quite effective in framing the complexities with which a vaccine production and distribution is riddled. A frequent compound noun in German is “Herkulesaufgabe”, as in “Das Verteilen des Impfstoffs wird eine Herkulesaufgabe” (Walker 2020) or – further strengthened by the adjective *wahr* – “Vor allem deshalb, weil das Land besonders an diesem Wochenende die Durchimpfungsrate stark steigern möchte, wäre die Impfung der Hochrisikopatienten eine wahre Herkulesaufgabe” (Rufin 2021). In English, the noun phrase “gargantuan task” was recorded – its origins date back to the 16th century, with the satiric novel *Gargantua* by François Rabelais, which depicts a giant king (cf. Merriam Webster):

After you identify the neutralising antibodies for different coronaviruses, then you've got to work out how to induce them through a vaccine, which could take years, he [Prof Dennis Burton, chair of the department of immunology and microbiology at Scripps Research] said. Altogether, at least five companies are taking on the gargantuan task of developing vaccines that can broadly protect against beta-coronaviruses (Grover 2020).

Another interesting metaphor, “silver bullet”, has its roots in mythology and in the fantasy genre, since this weapon was thought to be capable of killing supernatural creatures, typically werewolves. Many newspapers reported a statement by Tedros Adhanom Ghebreyesus, head of the World Health Organization (WHO), claiming that there was “no silver bullet at the moment - and there might never be” (BBC News 2020). In Italian, an analogous instance referring to a “proiettile d'argento” was found in *Il Sole 24 Ore* (Finotto 2020). However, according to terminologist Corbolante, who in August 2020 discussed the translation of “silver bullet” in her blog *Terminologia.etc*, Italian newspapers mentioning a “pallottola d'argento” or “proiettile d'argento” translated the figurative expression used by Tedros Adhanom Ghebreyesus too literally, since these Italian equivalents are

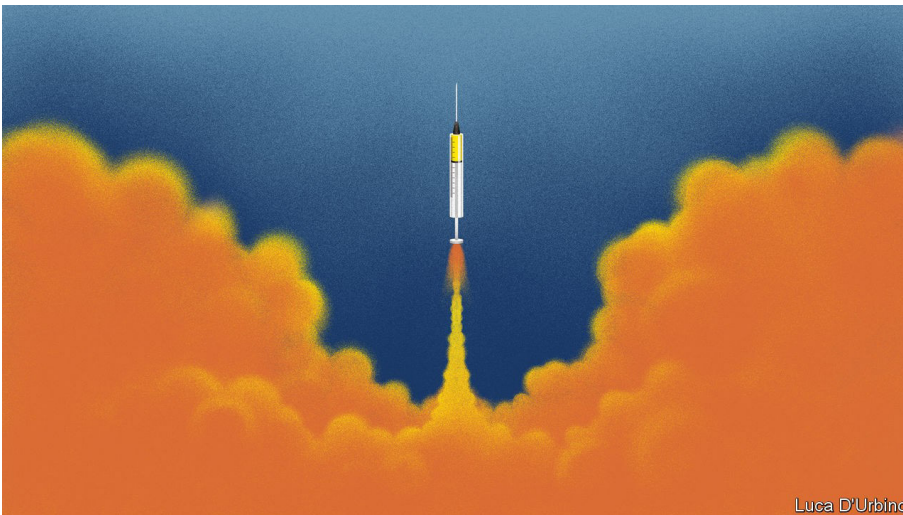


not as idiomatic as the English expression. In her opinion, in fact, “silver bullet” should have been translated as, for instance, “formula magica”.

Beside “silver bullet”, another metaphor mentioned by Corbolante in her blog is “magic bullet”. This figurative expression was indeed found in a German newspaper during the present research: *FAZ* mentioned a “magische Kugel”, “Zauber-kugel”, or “‘magic bullet’ frei nach Paul Ehrlich”<sup>12</sup> referring to the attempts to produce a vaccine made by chemists Rittner and Bourdeaux (Zoske 2020).

## 2.9. HISTORICAL EVENTS

Another source domain consists in *HISTORICAL EVENTS*. An extremely frequent parallelism refers to the space race, which is similar to the source domain of *RACES* for what it frames (competition and speed) but with the addition of the historical reference. An instance is the following: “The spectre arises of nations already engaged in a vaccine ‘space race’ trying to corner supplies for their own populations” (Spinney 2020) or “With governments all over the world pouring money into it, the effort to get a vaccine has come to resemble a more urgent version of the space race” (The Economist 2020). The latter newspaper article highlights the reference to the source domain of space exploration through the following picture, too:



As far as Russia is concerned, the reference to the space race was made explicit by the name of its vaccine Sputnik V, showing its pride in being the first country to register one. The parallel is further enhanced by the promotional

12 The Nobel Prize scientist who coined the term “magic bullet”, referring to substances that could target specific diseases. Therefore, this reference now pertains to the field of medicine rather than to the military one from which it originally stemmed.

video of this vaccine by the Russian Direct Investment Fund,<sup>13</sup> which shows a view of the Earth from space, menacingly surrounded by a brownish layer with spikes, depicting the coronavirus. This layer is gradually destroyed by the satellite Sputnik, which starts revolving around the Earth until it becomes covid-free.<sup>14</sup>

Getting back to newspapers, an article by *Lenta.ru* discussing vaccines was titled “Eto kak cholodnaja vojna” (Švarcman 2020), thus entailing a simile which compares the vaccine search to the Cold War due to its political implications.

Furthermore, another frequent parallelism refers to the arms race, as is the case in the simile “Covid, la corsa ai vaccini come quella agli armamenti: ecco il ‘nazionalismo antiviruses’” (Sabadin 2020).

Finally, another interesting historical reference is that to the D-Day, i.e. the Normandy Landings, which may be inferred from the front page of *La Repubblica* on the 9th December 2020: “V-day, il vaccino sbarca in Europa”.

## 2.10. MISCELLANEOUS SOURCE DOMAINS

Here we will discuss metaphors ascribable to several source domains, which, due to the overall scarcity of examples to be provided, were not reported in separate sections.

A relatively frequent metaphor in Italian refers to vaccines as ANTIDOTES: “Dall’Oms arriva la notizia che si sta lavorando agli standard per la privacy dei certificati vaccinali, pensati in modo tale da consentire agli individui di dimostrare di essere aver [sic] ricevuto l’antidoto contro il Covid-19” (Capponi 2021). It is not rare in Russian newspapers, either: “protivojadie koronavirusu možet ob”javitsja [sic] uže k sentjabrju”<sup>15</sup> (Dmitrieva 2020).

A quite transparent metaphor, referring to VACCINES as a LIFEBOUY, was found in the Russian newspaper *Lenta.ru*: “Spasatel’nym krugom v nynešnej situacii dolžny stat’ vakciny, kotorye ostonovjat rasprostranenie zaraženija”<sup>16</sup> (Švarcman 2020).

Other metaphors are those with the source domain FIRE – highlighted by Semino (2020) for their efficacy, as mentioned in 2.1. However, while they are quite frequent for the target domain PANDEMIC, they are not as common for the more circumscribed domain of VACCINES. Two analogous instances were identified in German and Italian:

13 <https://sputnikvaccine.com/rus/newsroom/press-kit/video/>

14 This instance does not stem from articles (or channels and podcasts) like all the other examples provided in the present paper, and neither is it linguistic content *stricto sensu*, but it is a clear instantiation of the mapping which is dealt with in this paragraph.

15 “An antidote for the coronavirus could appear already in September.”

16 “A lifebuoy in the current situation should be provided by vaccines, which will stop the coronavirus from spreading.”

Medikamente gegen Corona sind also **wie ein Löschwasserschlauch im Kampf gegen einen Waldbrand** – mal hier, mal da wird vielleicht ein Haus vor den Flammen gerettet. Aber nur **eine Schneise, die in den Wald geschlagen wird**, kann die Ausbreitung stoppen – und das ist in dieser Metapher ein Impfstoff. Denn nur wenn genügend Menschen immunisiert werden, kann die Ausbreitung effektiv gestoppt und Menschenleben gerettet werden. (Beck 2020).

In this excerpt, the concept of fire hoses is first introduced through a simile describing drugs used against the virus, and then followed by a metaphor about firebreaks to refer to vaccines. In the Italian version, instead, only firebreaks are hinted at:

Il professor Burioni ha [...] illustrato come il vaccino anti-Covid – “un’impresa incredibile avere un farmaco efficace così presto” - possa “impedire al virus di propagarsi come si farebbe con un incendio in una foresta: anziché tagliare gli alberi, si vaccinano le persone per isolare il virus” (Il Resto del Carlino 2020b).

Interestingly, in the anglophone area this metaphor is used only with regards to Wales and its “fire-break lockdown”, introduced in 2020 so as to reduce cases of coronavirus infection.

A last source domain is RELIGION. It is clear in examples such as *vaccine miracle* (Allen-Mills 2020) or *Impfwunder* (Von Altenbockum 2020), in the following contexts, respectively: “One newspaper described the results from a trial conducted by Pfizer, the American pharmaceutical giant, as a ‘vaccine miracle’”, and “Das Impfwunder wird nicht nur mit Begeisterung aufgenommen werden, sondern, eben weil es alle Rekorde bricht, auch auf Skepsis und Ablehnung stoßen – von der grundsätzlichen Impfgegnerschaft einmal ganz abgesehen”. Also in Italian miracle references are not rare, see: “Miozzo: ‘Il vaccino anti-Covid è un miracolo scientifico. In sei mesi risultati tangibili’” (Carratelli 2020). Other religious references in this language are “E le cure? I vaccini tanto attesi? Non saranno loro a **liberarci dal Male?**” (Verdelli 2020, with a clear allusion to the Lord’s Prayer, further underlined by the capitalization of *Male*) and “L’attesa messianica del vaccino”, title followed by this paragraph:

Una costante dalla scorsa primavera sono stati i periodici annunci dei leader per il prossimo vaccino risolutore – Putin l’ha annunciato per inizio autunno, Trump lo assicurava in perfetto tempismo con le elezioni Usa, Conte lo prevede per la primavera 2021 – generando una sorta di attesa messianica (Finotto 2020).

### 3. CONCLUSIONS AND OUTLOOK

In sum, the most frequently framed concepts are the speed at which research had to be conducted (e.g. through races, horse and car races metaphors), competition between countries as well as against the virus (e.g. through races, games,

gambling, war metaphors, historical references), and the general state of danger (particularly well-conveyed by war metaphors).

With respect to the source domains, the most frequent ones were definitely RACES and WAR. As far as the former is concerned, it must be noted that such metaphors are quite productive: see, for instance, well-attested terms like “arms race” or “space race”<sup>17</sup> (both also used as source domains in some examples provided in 2.9). This may be one of the factors that have led to the high frequency of this source domain.

Conceptualizations are sometimes combined or put in contrast in order to highlight some particular aspects of the target domain (see the instances referring to SPRINTS and MARATHONS in 2.1, or the podcast illustration with HUNTING as well as RACE and CLIMB as source domains in 2.3). However, it should be noted that the phenomenon of mixing metaphors ascribable to different source domains, known as “conceptual blending”, is not always ideal and is sometimes due to insufficient mastering of metaphor use (cf. Magris 2009: 99).

Another issue connected to the use of metaphors is the confusion they may cause in certain cases. In Italy, for instance, there have been debates about the use of “siero” and “antidoto” instead of “vaccino” in the media (cf. Accademia della Crusca, Di Carlo 2021; Corbolante 2021). Here, the question arises as to what extent “antidote” is a clear-cut metaphor and when, instead, it becomes an improper use of a technical term. According to me, such uses are appropriate if the metaphorical connotation is easily inferable from the context, otherwise this type of metaphorization ought to be avoided. In the example provided in 2.10, the Italian use of “antidoto” seems to be particularly ambiguous, since it refers to a concrete setting – where a vaccine has already been injected and a certificate needs to be issued – and not to vague assumptions and hopes for the future, where, in my opinion, “antidoto” would be more acceptable as a metaphor.

Another potential issue is connected to the prominent use of war metaphors, already pointed out in 1.2. While metaphors are influenced by what is occurring around us, our perception of reality, and the most widespread ideologies, the opposite is also true, as their use impacts on our perceptions, ideas, and opinions. In this respect, Komatsubara (2021) states that “metaphors build, bias, and balance our views on the situation we are faced with”. In the same vein, Colaci (2018: 38) – discussing the language of politics, its conative function and its realizations – maintains that metaphors activate the hidden associations on which our value system is based, and are therefore able to influence the audience’s emotional response. It goes without saying that this potential to manipulate readers’ perception may be dangerous. At the same time, war metaphors should not be demonized altogether.<sup>18</sup> A possible solution may be that provided by Komatsub-

17 Or, in Italian, also other similar terms like “corsa all’oro”, “corsa al ribasso”, etc.

18 Cf. also Musolff (2021) and his reflections about war metaphors in the Covid-19 discourse: “blanket criticism’ of all war/fighting terminology seems unjustified”.

ara (2021), who suggests that variation in metaphorical sources<sup>19</sup> could reduce conceptual bias and the risks associated with it.

Finally, another issue highlighted by the present analysis – specifically, by the case of “proiettile d’argento” in Italian newspapers, see 2.8 – consists in the translation of metaphors. This aspect, incidentally, has attracted the attention of many linguists and translation scholars for quite some time. One of the most celebrated and known models was put forward by Newmark (1981: 87-91), who proposed seven translation strategies to tackle the rendition of metaphors in another language: 1) reproducing the same image in the TL; 2) replacing the image in the SL with a standard TL image; 3) translating the metaphor by simile; 4) translating the metaphor (or simile) by simile plus sense; 5) converting the metaphor to sense; 6) deleting the metaphor; 7) combining the same metaphor with its sense. Similarly, Magris (2009: 100-102) highlights the need to select well-attested, natural-sounding metaphors in the TL, as metaphorical conceptualizations and phraseology are tightly intertwined. Consequently, the most suitable translation of the “silver bullet” metaphor would have been provided through the second strategy put forth by Newmark, i.e. the replacement of the SL image through a standard TL image. This instance, along with the discussion mentioned above with respect to the incorrect use of “antidoto” and “siero” in Italian newspapers, unfortunately confirms what has already been noted by Magris (*ibid.* 100), i.e. the fact that Italian journalists seem to use metaphors more erratically than, for example, German journalists (and, even more so, when translation comes into play).

In conclusion, a potential outlook of the present study, suggested by instances like the play on words “shot in the arm or shot in the dark”, the newspaper title which wonders if the vaccine will be a panacea or poison, or the skeptical excerpt regarding the “Impfwunder” (2.5, 2.8, and 2.10, respectively), is to examine in detail metaphors employed for the vaccination campaign in order to decrease vaccine hesitancy – as well as, conversely, those used by no-vax movements and in similar skeptical environments, as has been done by Signorini and Neves de Brito (2021) with respect to Facebook discussions on coronavirus vaccines in Brazil or Magris (2021) for general vaccination campaigns.

19 In my opinion, though, this should not be done within the same sentence or paragraph, so as to avoid the aforementioned conceptual blending.

## REFERENCES

- Black M. (1954) "Metaphor". In *Proceedings of the Aristotelian Society*, 55, Hoboken, Blackwell Publishing, pp. 273-294.
- Colaci L. A. (2018) *Politologia del linguaggio italiano e tedesco. Metafore concettuali e strategie retorico-narrative al Parlamento Europeo*, Milano, FrancoAngeli.
- Corbolante L. (2020) "Munizioni d'argento contro il coronavirus?", *Terminologia etc.*, <http://blog.terminologiaetc.it/2020/08/07/significato-silver-bullet/> (consultato il 23.01.2021).
- Corbolante L. (2021) "Promemoria per i media: vaccino ≠ siero ≠ antidoto", *Terminologia etc.*, <http://blog.terminologiaetc.it/2021/01/04/differenza-vaccino-antidoto-siero/> (consultato il 17.07.2021).
- Demjen Z., Semino E. & Koller V. (2016) "Metaphors for 'good' and 'bad' deaths. A health professional view", *Metaphor and the Social World*, 6:1, pp. 1-19.
- Di Carlo M. (2021) "Siero, antidoto e vaccino: facciamo chiarezza", *Accademia della Crusca.it*, <https://accademiadellacrusca.it/consulenza/siero-antidoto-e-vaccino-facciamo-chiarzza/2955> (consultato il 17.07.2021).
- Ervas F., Montibeller M. et al. (2016) "Expertise and metaphors in health communication", *Medicina & Storia*, 16:9-10, pp. 91-108.
- Foderi A. (2020) "Le parole del 2020", *Wired.it*, 31.12.2020.
- Hurrelmann K. & Baumann E. (2014) (Hrsg.) *Handbuch Gesundheitskommunikation*, Mannheim, Huber.
- Komatsubara T. (2021) "Variation in metaphor reduces risk in conceptualization: Metaphors on the COVID-19 pandemic in Japan". In *Paper of the 17th International Pragmatics Conference "The Pragmatics of Inclusion"*, Winterthur, 27th June-2nd July, being printed.
- Lakoff G. & Johnson M. (1980) *Metaphors We Live By*, Chicago/London, The University of Chicago Press.
- Lakoff G. & Turner M. (1989) *More than Cool Reason: A Field Guide to Poetic Metaphor*, Chicago/London, The University of Chicago Press.
- Liu N. & Lu S. (2021) "Metaphor use on the People's Daily Newspaper during public health emergencies: A diachronic and quantitative study". In *Paper of the 17th International Pragmatics Conference "The Pragmatics of Inclusion"*, Winterthur, 27th June-2nd July, being printed.
- Magris M. (2009) "Medizinische Fachbegriffe in der populärwissenschaftlichen Vermittlung: Ein Vergleich zwischen Deutsch und Italienisch". In P. Taino, M. Brambilla & T. Briest (Hrsg.), *Eindeutig uneindeutig. Fachsprachen – ihre Übersetzung, ihre Didaktik*, Frankfurt, Lang, pp. 89-104.
- Magris M. (2021) "Metaphern in der Impfkommunikation: zwischen Wissensvermittlung und Persuasion". In E. Agazzi, R. Calzoni, G. Carobbio et. al. (Hrsg.), *Übersetzen. Theorien, Praktiken und Strategien der europäischen Germanistik*, Bern, Peter Lang, pp. 363-375.
- Merriam Webster (2020) "Words We're Watching: 'Infodemic'", <https://www.merriam-webster.com/words-at-play/words-were-watching-infodemic-meaning> (accessed 23-01-2021).
- Miller R. S. (2010) "Speak up: 8 words and phrases to ban in oncology!", *Oncology Times*, 32:12, p. 20.
- Musolff A. (2012) "The study of metaphor as part of critical discourse analysis", *Critical Discourse Studies*, 9:3, pp. 301-310.



- Musolff A. (2021) "War against COVID-19": Is the pandemic as war metaphor helpful or hurtful?. In *Paper of the 17th International Pragmatics Conference "The Pragmatics of Inclusion"*, Winterthur, 27th June-2nd July, being printed.
- Newmark P. (1981) *Approaches to Translation*, Oxford, Pergamon Press.
- Niederhauser J. (1999) *Wissenschaftsprache und populärwissenschaftliche Vermittlung*, Tübingen, Narr.
- Ol'svang O. (2021) "Metaforičeskaja model' 'Pandemija – eto vojna' v russkom i anglojazыčnom medicinskom diskurse (na materiale publicističeskich tekstov)". In *Conference paper of "Methodological and Linguistic Problems of LSP"*, Samara, 18-19th May, being printed.
- Ortony A. (1979) *Metaphor and Thought*, Cambridge, U.P.
- Potts A. & Semino E. (2017) "Healthcare professionals' online use of violence metaphors for care at the end of life in the US: a corpus-based comparison with the UK", *Corpora*, 12:1, pp. 55-84.
- ReframeCovid (2020) "The initiative", <https://sites.google.com/view/reframecovid/initiative> (accessed 23-01-2021).
- Richards I. A. (1936) *The Philosophy of Rhetoric*, London, Oxford University Press.
- Semino E. (2008) *Metaphor in Discourse*, Cambridge, Cambridge University Press.
- Semino E. (2020) "A fire raging: Why fire metaphors work well for Covid-19", <http://cass.lancs.ac.uk/a-fire-raging-why-fire-metaphors-work-well-for-covid-19/> (accessed 23-01-2021).
- Semino E., Demjen Z. et al. (2018) *Metaphor, Cancer and the End of Life: A Corpus-Based Study*, London, Routledge.
- Semino E., Demjen Z. & Demmen J. (2018) "An integrated approach to metaphor and framing in cognition, discourse, and practice, with an application to metaphors for cancer", *Applied Linguistics*, 39: 5, pp. 625-645.
- Semino E., Heywood J. & Short M. (2004) "Methodological problems in the analysis of metaphors in a corpus of conversations about cancer", *Journal of Pragmatics*, 36: 7, pp. 1271-1294.
- Signitzer B. (2001) "Ansätze und Forschungsfelder der Health Communication". In K. Hurrelmann & A. Leppin (Hrsg.), *Moderne Gesundheitskommunikation*, Bern, Huber, S. 22-35.
- Signorini I. & Neves de Brito L. A. (2021) "Mappings and metaphorical scenarios in contemporary discussions on Facebook about the Brazilian vaccine 'crisis'". In *Paper of the 17th International Pragmatics Conference "The Pragmatics of Inclusion"*, Winterthur, 27th June-2nd July, being printed.
- Spina S. (2020) "La peste, il terremoto e altre metafore. Il coronavirus nel discorso della stampa italiana", *Treccani*, 10.04.2020.
- MASS MEDIA REFERENCES
- Allen-Mills T. (2020) "Who will get the Pfizer Covid vaccine first?", *The Times*, 15.11.2020.
- Addendum (2020) "COVID-19: Ist der Impfstoff-Hype berechtigt?", *Addendum*, 02.06.2020.
- Baars C. & Lambrecht O. (2020) "Coronavirus: Der große Impfplan für Deutschland", *ndr*, 06.11.2020.
- Bartoloni M. & Pieraccini S. (2020) "Vaccini e terapie, il Centro Italia avamposto nella lotta al Covid", *Il Sole 24 Ore*, 30.10.2020.
- BBC News (2020) "Coronavirus: WHO warns of 'no silver bullet' amid vaccine search", 03.08.2020.
- Beck D. (2020) "Was bringen Medikamente im Kampf gegen Corona?", *SWR*, 21.10.2020.
- Beljaev D. (2020) "Eto daže ne gonka, a lotereja. Kto pervym najdet vakcinu ot koronavirusa?", *TASS*, 18.03.2020.
- Bidetti M. (2020) "Corsa al vaccino per Covid-19, quanto manca al traguardo?", *Galileo*, 01.10.2020.
- Bocci M. (2020) "Giovani laureati, infermieri, assistenti sanitari: la carica dei 20mila vaccinatori", *la Repubblica*, 02.12.2020.
- Boseley S. (2020) "'It's possible': the race to approve a Covid vaccine by Christmas", *The Guardian*, 30.10.2020.
- Bresolin M. (2020) "Kyriakides: 'Un solo vaccino non sarà la panacea. La strada è lunga'", *La Stampa*, 12.11.2020.
- Brost M. & Grabitz I. (2020) "Der globale Impfkrieg: Wie China mit Spritzen seine Macht ausbaut", *Die Zeit*, 18.12.2020.
- Buletti M. G. (2020) "La maratona del vaccino", *Azione*, 24.08.2020.
- Capponi M. (2021) "Vaccini, Sanofi al lavoro per produrre quello di Pfizer-Biontech. Impennata di casi in Cina", *Milano Finanza*, 15.01.2021.
- Carr E. (2020) "The path to a covid-19 vaccine may be quick, but it will be bumpy", *The Economist*, 16.11.2020.
- Carratelli N. (2020) "Miozzo: 'Il vaccino anti-Covid è un miracolo scientifico. In sei mesi risultati tangibili'", *La Stampa*, 27.12.2020.
- Curzi P. (2020) "Vaccino Coronavirus, nelle Marche resta un rebus", *il Resto del Carlino*, 27.11.2020.
- Diehm C. (2020) "Warum die Impfstoffentwicklung gegen das Coronavirus so lange dauert", *Handelsblatt*, 26.03.2020.

- Dittrich B. (2020) "Trotz Corona-Impfstoff: Auch 2021 müssen wir noch Abstand halten", *vorwärts*, 10.11.2020.
- Dmitrieva O. (2020) "Vakciny protiv koronavirusa COVID-19 obeščali k sentjabrju", *Rossijskaja Gazeta*, 11.04.2020.
- Dusi E. (2020) "Nello sprint per il vaccino, Moderna tenta il sorpasso: 'Il nostro è efficace al 94,5%'", *la Repubblica*, 16.11.2020.
- Ehrhardt M. (2020) "Welche Pharmafirmen sich am Corona-Abwehrkampf beteiligen", *Deutschlandfunk*, 16.03.2020.
- FAZ (2020) "Machtkampf um Corona-Impfstoff", 10.11.2020.
- Finotto C. A. (2020) "Coronavirus, il vaccino potrebbe non essere il proiettile d'argento e la lotta sarà ancora lunga", *Il Sole 24 Ore*, 14.11.2020.
- Gaebler C. & Nussenzweig M. C. (2020) "All eyes on a hurdle race for a SARS-CoV-2 vaccine", *nature*, 19.10.2020.
- Ghisleni P. (2020) "Vaccino Covid, Roberta Villa: 'Bisogna verificare se bloccherà la diffusione del virus'", *BergamoneWS*, 06.12.2020.
- Grover N. (2020) "The long game: the race for a vaccine against all coronaviruses", *The Guardian*, 13.11.2020.
- Haas S. (2020) "Wettlauf um Corona-Schutz. Diese Firmen liegen im Impfstoff-Rennen vorn", *zdf*, 16.12.2020.
- Hickmann M. (2020) "Kampf gegen Coronavirus. Der unfaire Poker um Impfstoffdosen", *zdf*, 25.07.2020.
- Il Giorno* (2020) "Vaccino anti Covid: in Lombardia si parte il 15 gennaio", 10.12.2020.
- Il Messaggero* (2020) "Vaccino Covid, annuncio di Pfizer: 'Efficace al 95%'. Test su 43.500 persone senza problemi. Oms: seconda ondata combattuta senza cure", 18.11.2020.
- Il Resto del Carlino* (2020a) "Vaccino Covid: Emilia Romagna in prima linea. 'Test a dicembre'", 22.11.2020.
- Il Resto del Carlino* (2020b) "Vaccino Covid, Burioni e Gorini: 'Funziona ed è sicuro'", 29.11.2020.
- Il Sole 24 Ore* (2020) "Da Pfizer a AstraZeneca ai 5 vaccini cinesi: annunci e progressi sulla cura al Covid. Possibile ok in Europa a dicembre", 18.11.2020.
- Iščenko A. (2020) "Rossijskaja vakcina ot koronavirusa dlja Ukrainy: jad ili panaceja", 24 *Novosti*, 15.10.2020.
- Khurshudyan I. (2020) "In race for coronavirus vaccine, Russia ramps up rhetoric to defend Sputnik V", *The Washington Post*, 04.10.2020.
- Kireev M. (2020) "Russlands Impfstoff gerät ins Hintertreffen", *Die Zeit*, 16.11.2020.
- Lambrecht O. & Baars C. (2020) "Der weite Weg zur Corona-Impfung", *Tagesschau*, 06.11.2020.
- Lange S. (2020) "Schutz im Eiltempo: Wie komplex wird die Impfstoff-Offensive?", *Augsburger Allgemeine*, 21.12.2020.
- Mast J. (2020) "Novartis jumps into Covid-19 vaccine hunt, as Big Pharma and big biotech commit to billions of doses", *Endpoints*, 28.05.2020.
- Modeo S. (2020) "Vaccino Covid, la battaglia tra Pfizer e Moderna per la 'rivoluzione mRNA'", *Il Corriere della Sera*, 20.11.2020.
- Müller H. (2020) "Corona-Impfstoffe: Die Ziellinie im Blick?", *Labor Journal*, [https://www.laborjournal.de/rubric/hintergrund/hg/hg\\_20\\_09\\_01.php](https://www.laborjournal.de/rubric/hintergrund/hg/hg_20_09_01.php) (consultato il 23.01.2021).
- NDR (2020) "Das Wichtigste zu Corona-Impfstoffen", <https://www.ndr.de/fernsehen/>
- sendungen/panorama\_\_die\_reporter/Coronavirus-Impfstoff-Zulassung,impfstoff128.html (consultato il 23.01.2021).
- Paudice C. (2020) "Gli affari opachi dietro la corsa al vaccino", *HuffPost*, 14.11.2020.
- Pitts P. J. (2020) "The sprint to a COVID-19 vaccine is a relay race", *The Washington Times*, 25.11.2020.
- Richards S. E. (2020) "Why Moderna may have an edge in the vaccine race: refrigeration", *National Geographic*, 20.11.2020.
- Rohwedder W. (2021) "Geimpft - und dann?", *Tagesschau*, 19.01.2021.
- Rufin M. (2021) "Wer jetzt geimpft wird", *Die Neue Südtiroler Tageszeitung*, 16.01.2021.
- Schlak M. (2020) "Das riskante Wettrennen der Pharmakonzerne", *Der Spiegel*, 14.03.2020.
- Sheen M. (2020) "Shot in the arm or a shot in the dark? Complacency fears over Covid-19 vaccine equity market boost", *Investment Week*, 13.11.2020.
- Spinney L. (2020) "How the race for a Covid-19 vaccine is getting dirty", *The Guardian*, 30.08.2020.
- SRF (2020) "So steht es im Rennen um den Impfstoff", 23.11.2020.
- Strana.ua (2020) "Ochota na Covid-19. Kto i kogda sozdast vakcinu ot koronavirusa", 24.03.2020.
- Švarcman V. (2020) "Eto kak cholodnaja vojna", *Lenta.ru*, 15.10.2020.
- Tajzhetinova M. (2020) "Sputnik ili Moderna: kak prochodit vojna vakcin i kto vpered", *MBK*, 24.10.2020.
- TASS (2020) "Eto daže ne gonka, a lotereja. Kto pervym najdet vakcinu ot koronavirusa?", 18.03.2020.
- Tgcom24 (2020) "Covid, piano vaccini ai blocchi di partenza: possibilità di requisire frigoriferi



alle industrie e fiale tracciate su un portale”, 29.11.2020.

Verdelli C. (2020) “Covid, la scienziata che isolò il coronavirus: ‘Il vaccino da solo non basterà a far sparire il virus’”, *Il Corriere della Sera*, 22.11.2020.

Von Altenbockum J. (2020) “Wann kommt der Corona-Killer?: Das Impfwunder”, *FAZ*, 09.11.2020.

Walker B. (2020) “Das Verteilen des Impfstoffs wird eine Herkulesaufgabe”, *Badische Zeitung*, 24.11.2020.

Weber N. (2020) “Was Allergiker jetzt wissen müssen”, *Der Spiegel*, 23.12.2020.

ZDF (2020) “Suche nach Impfstoff”, <https://www.zdf.de/nachrichten/zdfspezial/zdf-spezial---coronavirus-clip-4-100.html> (consultato il 23.01.2021).

Zoske S. (2020) “Eine magische Kugel aus Tuberkulose-Bakterien”, *FAZ*, 10.11.2020.

# Hacia una sistematización de las disimetrías morfosintácticas en la traducción del régimen preposicional entre lenguas tipológicamente emparentadas: español/italiano

GIUSEPPE TROVATO  
Università Ca' Foscari Venezia  
[giuseppe.trovato@unive.it](mailto:giuseppe.trovato@unive.it)

## ABSTRACT

The aim of this paper is to analyse some aspects of the prepositional system from a translational and contrastive point of view between two related languages (Spanish/Italian). Mastering prepositions in a foreign language is an arduous task and undoubtedly contributes to linguistic improvement. The activity of translation can often be facilitated by linguistic proximity, but when dealing with the prepositional system, numerous dissymmetries can appear, the analysis of which is of the utmost importance in order to avoid linguistic calques and interferences, as well as to favour metalinguistic reflection on the different ways of expressing prepositions between two related languages. For our analysis, we will take into consideration the novel *Caperucita en Manhattan* by the Spanish writer Carmen Martín Gaité and its Italian translation carried out by Michela Finassi Parolo. The corpus created will allow us to carry out a cross-linguistic comparison in order to highlight the existing dissymmetries.

## KEYWORDS

Prepositional system, translation process, related languages (Spanish/Italian), morphosyntactic dissymmetries, *Caperucita en Manhattan*

## 1. INTRODUCCIÓN

El dominio de una lengua extranjera pasa por varios niveles lingüísticos que se supone que un aprendiente tiene que ir asimilando y desarrollando para poder configurarse como un usuario competente. En este escenario, se coloca la amplia temática de las preposiciones cuyo uso, valores y funciones no se corresponden por completo entre una lengua y otra. Este fenómeno no remite solo a lenguas tipológicamente alejadas como por ejemplo el ruso, el japonés o el chino, sino que caracteriza también a lenguas más cercanas como las pertenecientes al ámbito de las lenguas romances. En este sentido, el italiano y el español no son ninguna excepción, por mucho que se califiquen tradicionalmente como lenguas afines.

En este artículo vamos a abordar el régimen preposicional desde el terreno de la traducción, pues el fenómeno traductor es el que de forma más evidente permite realizar una comparación interlingüística con el fin de observar los cambios potenciales que determinados ítems lingüísticos tienden a sufrir a lo largo del proceso traslativo.

Para alcanzar nuestros objetivos, usaremos un corpus creado a partir de la novela *Caperucita en Manhattan* (1990) de la célebre escritora salmantina Carmen Martín Gaité y su correspondiente traducción en lengua italiana llevada a cabo por Michela Finassi Parolo (2011). El cotejo entre la versión española y la traducida pondrá de relieve las variaciones y disimetrías morfosintácticas que se producen en el ámbito del régimen preposicional a la hora del proceso de traducción.

## 2. BREVE PANORÁMICA SOBRE EL RÉGIMEN PREPOSICIONAL DESDE UNA ÓPTICA CONTRASTIVA

Cualquiera que se haya embarcado en el estudio de una lengua extranjera, en algún que otro momento de su recorrido de aprendizaje habrá oído que para conocer realmente bien un idioma extranjero hay que dominar el sistema de las preposiciones. En efecto, entre los errores más recurrentes que cometen los estudiantes de español como lengua extranjera, figuran los relacionados con el uso correcto del régimen preposicional. Basta con poner el ejemplo relativo a la diferencia entre *por* y *para*, para percatarse de las dificultades que puede llegar a suponer para un no nativo el uso efectivo de las preposiciones.

La *Nueva Gramática de la Lengua Española* (2010: 557-560)<sup>1</sup> define las preposiciones como palabras invariables y casi siempre átonas que introducen un complemento llamado “término” con el cual forman un grupo preposicional. El grupo preposicional modifica varias clases de palabras, de particular manera los

1 La *Nueva Gramática de la Real Academia* dedica un apartado muy amplio (557-573) al tema de las preposiciones y al grupo preposicional, contemplando también las locuciones preposicionales y los grupos sintácticos complejos formados mediante preposiciones.

verbos, los sustantivos, los adjetivos y los adverbios. Además, puede desempeñar la función de complemento de régimen en calidad de elemento argumental como en *contar con su apoyo* o de un complemento adjunto como en el caso de *famosa por sus vinos*. Asimismo, el conjunto preposicional puede ejercer la función de atributo como en *habló en voz baja* o como complemento predicativo *Llegó cansada del paseo*. El sintagma preposicional, en algunos casos, admite modificadores, adverbios o locuciones adverbiales que hacen más específico y delimitan el contenido, así como el significado de la expresión, como en *llegó casi al final de la cita*. La gramática de la Real Academia Española postula, además, que existen preposiciones que poseen contenido gramatical y otras que solo cuentan con un contenido propiamente léxico.<sup>2</sup>

Son varios los estudios realizados en torno al proceso de enseñanza-aprendizaje de las preposiciones por parte de varios tipos de aprendientes (Cf. Cano Gines, y Flores Ramírez, 1998; Minguell, 2009; Regueiro Rodríguez, 2014).

En la pareja de lenguas español/italiano, no han faltado aportaciones en torno al sistema preposicional desde varias perspectivas de análisis. Uno de los estudios pioneros llevados a cabo en este marco es el de Katerin Katerinov (1980) titulado *Uso delle preposizioni in italiano e in spagnolo*, en el que se ofrece una interesante aproximación al tema considerando los aspectos simétricos y disimétricos a nivel de uso de las preposiciones en las dos lenguas.

En un estudio nuestro (Cf. Trovato, 2013), abordamos el régimen preposicional desde tres enfoques: contrastivo, traductológico y pedagógico; hemos tomado en consideración solo los casos en los que no se da una uniformidad en términos de uso y funciones de las preposiciones entre el italiano y el español.<sup>3</sup> En efecto, la gramática de dependencias establece reglas de estructura sintagmática que reflejan propiedades léxicas y consideran el contexto de las relaciones sintácticas y semánticas que existen entre los lexemas.

Bermejo Calleja (2007) aborda el tema de las preposiciones a partir de los diccionarios bilingües español-italiano/italiano-español, con el fin de observar su identidad y contraste. Caballero Rubio y Corral Hernández (2006) realizan un tratamiento contrastivo español-italiano de las preposiciones locativas, subrayando las disimetrías que se producen entre una lengua y la otra. Lenarduzzi (2007), en su estudio, baraja la hipótesis de una incidencia de las preposiciones en lo que la tradición gramatical califica como modalidad y basa su trabajo en la teoría de la enunciación. Recientemente Luque Toro (2019) ha afrontado la cuestión de las locuciones adverbiales introducidas por la terna preposicional *a, de, en* y ha estudia-

2 La preposición “a”, por ejemplo, desempeña una función puramente sintáctica cuando introduce el complemento objeto directo o indirecto, como en el caso de *escogieron a él entre todos* y *dieron el libro a su abuela*.

3 Algunos ejemplos de disimetrías en el tratamiento traductológico de las preposiciones serían los siguientes: la conozco de oídas = *la conosco per sentito dire*; el abogado actuó de buena fe = *l'avvocato ha agito in buona fede*; mi dinero no es nada con el que tú tienes = *i miei soldi non sono niente in confronto ai tuoi*.

do cómo, a partir del concepto original de espacio inherente a cada una de ellas, se configura una amplia red de combinaciones cuyos significados aparecen definidos en función de las distintas preposiciones que rigen estas construcciones.

Antes de terminar esta breve panorámica acerca del régimen preposicional desde una óptica contrastiva, también es interesante considerar la perspectiva del uso de las preposiciones por parte de los hispanohablantes, pues la afinidad lingüística puede ser causa de dificultades tanto en una dirección como en la otra. En este sentido, Solsona Martínez (2007) en su tesis doctoral aborda el análisis de errores y de la interlengua en el aprendizaje de las preposiciones italianas por parte de aprendientes hispanófonos y, desde varios puntos de vista, la problemática contrastiva es común a ambos grupos.

### 3. SISTEMATIZACIÓN DE LOS RESULTADOS DE LA TRADUCCIÓN DE LAS PREPOSICIONES EN LA DIRECCIÓN ESPAÑOL > ITALIANO A PARTIR DE LA NOVELA *CAPERUCITA EN MANHATTAN*

Tras enmarcar la cuestión del régimen preposicional desde una perspectiva teórica y contrastiva, podemos proceder a presentar y analizar los resultados del rastreo traductológico que hemos llevado a cabo usando como corpus de nuestro material lingüístico la novela *Caperucita en Manhattan* de Carmen Martín Gaité y su traducción oficial en lengua italiana a cargo de Michela Finassi Parolo.

Dada la amplitud del tema tratado, nuestro estudio se ha ceñido al análisis de tres preposiciones, a saber, las más comúnmente utilizadas en los intercambios comunicativos cotidianos. Se trata, concretamente, de las preposiciones: *a*, *en*, *de*. Además, cabe precisar que no solo hemos tomado en cuenta las preposiciones a solas, sino también las locuciones preposicionales formadas a partir de ellas.

#### 3.1. RASTREO TRADUCTOLÓGICO DE LA PREPOSICIÓN *A* Y ANÁLISIS CONTRASTIVO

Vio a su hija que llevaba un largo rato parada en la puerta de la cocina	Vide sua figlia che da tempo era ferma sulla porta della cucina
Aurelio había enterrado a Gloria Star	Aurelio aveva seppellito Gloria Star
Cuando su padre llamaba “lagarta” a la abuela	Quando suo padre chiamava la nonna “lucertolona”
La señora Allen soltaba unos instantes a Sara de la mano	La signora Allen lasciava andare per qualche momento la mano di Sara
Aquellas historias acompañaban luego a Miss Lunatic	Quelle storie accompagnavano miss Lunatic
Para eso tienes a los Maestros Tartufos	Hai già i maestri del tartufo
Había despedido a sucesivos pasteleros	Aveva licenziato un pasticciere dopo l’altro
Miró a la anciana	Guardò l’anziana donna
Estaba insultando a alguien	Stava insultando qualcuno

Sabes tú si han cogido al vampiro del Bronx?	Non sai se hanno preso il vampiro del Bronx?
Reconocido debajo de su disfraz de mendiga a la Libertad	Aveva riconosciuto sotto la maschera della mendicante la Libertà
Le encargaba que cuidara a aquella criatura	Gli raccomandava di aver cura di quella bimba
Las calles que quedan a la derecha	Le strade sulla destra
Había a la entrada un perchero con cabezas doradas de león	Nell'entrata c'era un attaccapanni con teste di leone dorate
Salió al pasillo	Uscì in corridoio
Había que bajar a él	Vi si arrivava
El señor Allen fue a su dormitorio	Il signor Allen andò in camera da letto
Llegando a la fontanería	Mentre andavano in negozio
Edgard Woolf llamó a su despacho a aquellos consejeros	Edgard Woolf chiamò nel suo ufficio quei consiglieri
Oyó aplausos a su alrededor	Sentì degli applausi intorno a lei
A la espalda	Con le mani dietro la schiena
Si caes al pozo estás perdida	Se cadrai nel pozzo sarai perduta
Hasta que salieron a la superficie	Finché non furono uscite in superficie
Al final de una escalinata	In fondo a una scalinata
Al amparo de una campanita	Schermate con una piccola campana
A la puerta	Sulla porta
Esperando a la puerta	Aspettarlo davanti alla porta
Al otro lado se veían edificios	Dall'altra parte si vedevano edifici
No quiero gente inepta a mi lado	Non voglio degli inetti attorno a me
De quince centímetros a cada lado	Di quindici centimetri per parte
Me he olvidado de poner pilas a la linterna	Mi sono dimenticata di mettere le batterie nella pila
A cualquier lado	Dappertutto
¿Ves la estación terminal del ferri a la isla?	Vedi la stazione dei ferryboat per l'isola?
Pudiera ser que hubiera trasladado a este otro barrio	Poteva benissimo aver trasferito il suo campo in quest'altro quartiere
La cola daba la vuelta a la manzana	La coda proseguiva fin dietro l'angolo
Vio brillar a lo lejos	Vide brillare lontano
Dejando a la derecha	Lasciando sulla destra
Le dijo, al verla tirada en la acera	Le aveva detto, vedendola lunga distesa sul marciapiede
Le parecía maravilloso por el sonido que tenía al decirlo	Le sembrava meraviglioso per il suono che aveva
Para que no tuviera frío al salir	Così che non avesse freddo dopo
Se habían desvanecido al dar la luz	Erano svaniti con la luce
Pero no le daba tiempo a verlos bien	Non aveva tempo di guardarli bene
Crujían al partirlos por la mitad	Scricchiolavano quando li si divideva a metà
Espere un momento a ver	Aspetti un momento

Al pasar por el escaparate	In vetrina
Cuando empezaba a caer la noche	Al calare della notte
(Se oyó el golpe) al cerrarse	(Si senti il colpo) che si chiudeva
(Experimentó inquietud) al reconocer el gesto de la estatua	(Provò una leggera inquietudine) riconoscendo il gesto della statua
(una gota cayó) a manchar la servilleta de cuadros	Una goccia cadde macchiando il tovagliolo a quadretti
Le había dicho al salir del café	Le aveva detto, uscendo dal Caffè
Al abrirlos, vio	Quando li aprì, vide
A saber	Chi lo sa?
Le había dado tiempo a comprobar	Aveva avuto il tempo di constatare
Al salir del ascensor	Mentre usciva dall'ascensore
Volver al mundo civilizado	Tornare nel mondo civile
Su abuela Rebeca, que ya no había vuelto a llamarse Gloria Star	Sua nonna Rebecca che ormai non si chiamava più Gloria Star
Se sentía a disgusto con que se rieran tanto	Era seccata che ridessero tanto
Porque sus padres se ponían a discutir	(Per capire come mai) i genitori discutessero
Para dar pie a las fantasías de Sara	Per accendere le fantasie di Sara
Casi siempre salía a relucir	Quasi sempre veniva tirata in ballo
Segura de que la respuesta iba a ser afirmativa	Sicurissima che la risposta sarebbe stata affermativa
Aurelio ya nunca lo iba a conocer	Ormai non avrebbe più conosciuto Aurelio
Qué me va a dar miedo!	Ma che paura e paura!
¿Adónde va a ir con el frío que hace?	Dove dovrebbe andare con il freddo che fa?
Me voy a llevar una sorpresa	Ci sarà una sorpresa per me
Nunca se atrevía a decirlo	Non aveva mai osato dirlo
Le dio a la tecla de la marcha	Premette il tasto di accensione
No le metas líos en la cabeza a la niña	Non fare andare in confusione la bambina
A tu abuela la tratan como a una reina	Non c'è dubbio che tratti tua nonna come una Regina
Unos asuntos tan insulsos vinieran a ocupar la mente de su madre (..) y a interrumpir	Questioni così insulse occupassero la mente di sua madre e interrompessero i suoi pensieri
Le temo a la vejez de mi madre	Ho paura per la vecchiaia di mia madre
Se echó a reír	Scoppiò in una risata
La niña volvía a hacer pucheros	La bambina stava di nuovo per scoppiare in lacrime
No me atrevo a decirlo	Non oso dirlo
No vuelvas a tener miedo jamás	Non aver mai più paura
Atendiendo a mil detalles rutinarios	Mi occupo di mille cosette di routine
(notó) que se adelantaba a hacerlo	(Si accorse) che l'aveva preceduta
Había llegado a la conclusión de que la más impresionante	Era arrivata alla conclusione che quella più vivida
Él se apresuró a acudir	Lui si affrettò

Solo le faltaba romper a hablar	Gli mancava soltanto la parola
A lo mejor, un día, de pronto, lo siente usted solo y entiende	Forse un giorno, tutto d'un tratto, se lo sente lei da solo e lo comprende
Se toma el aire a gusto	Si può prendere una boccata d'aria
Estaba a oscuras	Era buio
Pelo peinado a raya	(Un uomo) pettinato con la riga da una parte
El corazón le latía a toda prisa	Il cuore le batteva furiosamente in petto
Salía un olor desvaído a papel antiguo, a flores secas	Ne usciva un odore di carta antica, di fiori secchi
Pasaba del oro al verde, casi a la vez	Che passava dall'oro al verde, quasi contemporaneamente
Combinaba a su capricho las letras	Combinava secondo il suo capriccio
Los tapaba a propósito	Glieli copriva di proposito
Mezclando vocales y consonantes a la buena de Dios	Mescolando vocali e consonanti a casaccio
Llegaron a echarla a patadas	L'avevano cacciata via a pedate
Las iluminaba todas a la vez	Li illuminava tutti insieme
Se paseaba a solas y pensativo	Passeggiava da solo immerso nei suoi pensieri
Andaba a toda prisa	Camminava in tutta fretta
Cualquier situación se puede volver al revés	Qualunque situazione può capovolgersi
Le abrió una lata de comida a Cloud	Aprì una scatoletta di cibo per Cloud
Tus consejos nos aburren a todos	Con i tuoi consigli ci hai stufato tutti quanti
Parecía tenerles a los dos más simpatía que su mujer	Sembrava dimostrare più simpatia per quei due che non a sua moglie
(A la señora Allen) le entraba una especie de tic nervioso	(La signora Allen) veniva colta da una specie di tic nervoso
A mí me encantaría tener dinero	Mi piacerebbe da matti avere qualche soldo
¡A mí ya me harta!	Io ne ho fin sopra i capelli!
Inyectar fe a los desesperados	Iniettare fiducia nei disperati
A cuyo cargo corría la revisión	Responsabile della revisione
Para insuflarle vida a ella	Per soffiare la vita in lei
¿Me lo vas a contar a mí?	A me lo racconti?
Me pasa a mi contigo	A me succede lo stesso con te
Se había negado a sí mismo	Si era negato
¡Qué bastante rabia me da ya a mí sola!	Mi fa già abbastanza rabbia così!
Le pasaría lo que a ella	Gli era successo come a lei
Al fin y al cabo es el único que tendría derecho	Dopotutto, è l'unico che avrebbe diritto di farlo
Al fin y al cabo	Tra l'altro
Al fin y al cabo, era una diosa	Alla fin fine era una dea
Se apagaran las diez al mismo tiempo	Si spensero contemporaneamente



Precisamente hoy a la hora de comer	Mentre pranzavamo
Al cabo de diversas llamadas telefónicas	Dopo svariate chiamate
Al son de aquel tintineo	A quel tintinnio
Ornamentación puesta al día	Arredi all'ultima moda
Salen del horno a diario	Escono quotidianamente dai forni
Se iba parando a cada momento	Si fermava ogni minuto
A todas estas, había llegado	Nel frattempo, era arrivato
A los pocos instantes	Pochi istanti dopo
A medida que iban alejándose	Mentre si allontanavano
A la media hora	Mezz'ora dopo
Olía a colillas, a cerrado, a sudor y a perfume barato	C'era odore di cicche, di chiuso, di sudore e di profumo a poco prezzo
Realmente al olor a bollos, tartas y pasteles	Il profumo dei pasticcini, delle torte e dei croissant
(mi tarta) que sabe a jarabe	(La mia torta) sa di sciroppo
Percibió un leve perfume a jazmín	Avvertì un lieve profumo di gelsomino
Toque de animal al acecho	L'aria di un animale in agguato
Olía a tarta de fresa	Aveva il profumo della torta di fragole
A media altura	A metà della sua altezza
Pero a falta de pan, buenas son tortas	Ma chi si contenta gode
Venga o no venga a cuento	(Imbarcarsi) in discorsi inopportuni
No le saca partido a nada	Non prova gioia per nulla
A medida de que su jefe	Via via che il suo capo
Atender así a las cuitas y alteraciones	Prendersi cura delle angosce e dei cambiamenti
Los viejos o nos ponemos al día	Noi vecchi o siamo aggiornati
Había dado pie a sus fantasías nocturnas	Avesse acceso le sue fantasie notturne
Aunque sea a rastras	Anche dovessi trascinarle per i Capelli
Plántale cara a ese miedo	Guardala bene in faccia questa paura
Se le cayó el alma a los pies	Vide tutto nero
Como a las niñas de sus ojos	Come se fosse stata sua figlia
A duras penas	Faceva fatica
Como a las niñas de sus ojos	Come se fosse sua figlia
Se había metido a Peter en el bolsillo	Teneva Peter in pugno
Esta entrada a hurtadillas	Ingresso in punta di piedi
A manera de bombilla dibujada	Come la lampadina disegnata
El señor Allen fue a recogerlo (el teléfono)	Il signor Allen andò per rispondere
A su casa a leerle la mano	A casa sua per leggergli la mano
La apuñalaron sin consideración a su edad	La pugnarono senza nessun rispetto per la sua età
Un elixir bastante bueno contra al miedo	Un elisir ottimo per la paura
Miedo a estarse despidiendo de ella	Timore di separarsi da lei
Miedo a descubrir ante sí	Paura nel trovarsi davanti
¿No sería miedo a la libertad?	Non sarà forse paura della libertà?
En honor al apellido de la familia	In onore del cognome di famiglia

Lo que hacía falta para darle contraste al ambiente	Mi mancava per creare un contrasto nell'ambiente
La gente que tiene miedo a lo maravilloso	La gente che ha paura del meraviglioso
Totalmente ajenas a la cámara	Del tutto inconsapevoli della macchina da presa
Ha sido una fuerza superior a mí	È stato più forte di me
Una curiosidad superior a él	Una curiosità più forte di lui

Tabla 1 – Resultados del rastreo traductológico de la preposición *A* en la novela *Caperucita en Manhattan*

Como se puede comprobar, los casos más cuantiosos remiten al uso de la preposición *a* + complemento objeto directo, fenómeno normal en lengua española y no contemplado en italiano. En estos casos la preposición *A* es un elemento gramatical que introduce el complemento como en *vio a su hija/vide sua figlia*. En muchas oraciones la preposición *A* cuenta con un valor locativo que se expresa de varios modos y matices en la lengua italiana como el caso del locativo-estativo: *sulla, al, intorno, dietro, in fondo, schermate con* (al amparo de). Como es evidente, muy frecuentemente la preposición *A* es traducida mediante la preposición *su*. Otra traslación muy frecuente de esta preposición es *in, nel, nella*. La preposición *A* es utilizada también para señalar un movimiento hacia un lugar que en italiano puede corresponder a *in, al, verso il, vi* (si *arrivava*). En este último caso, se puede constatar que la preposición *A* puede verse al italiano con el adverbio de lugar equivalente a *ivi* utilizado con verbos de movimiento hacia un lugar + verbo impersonal. Hay también casos en los que *A* combinada con la palabra *lado*, indica diferentes partes del espacio y de las perspectivas en italiano: *dall'altra parte, intorno, per parte*.

En algunos casos la preposición española *A* y su complemento resultan corresponderse con un adverbio en italiano: *a cualquier lado/dappertutto; a lo lejos/lontano*. La forma compuesta *AL + infinitivo* se traduce en italiano con un sentido temporal: *dopo, quando, che si* (*chiudeva*). A veces *A* o *AL + infinitivo* cambia ligeramente de significado y se vuelve más simple en su traducción: *espere un momento a ver/aspetti un momento; al pasar por el escaparate/in vetrina*. En algunos casos, en italiano corresponde un gerundio que remite a la contemporaneidad de dos acciones o que una acción sigue inmediatamente a la otra o puede tratarse de una oración introducida con *mentre*. La preposición *A + infinitivo* puede también tener el sentido de sorpresa, exclamación o interrogación como en *a saber/chi lo sa?* La preposición *A* se puede utilizar con verbos preposicionales tanto en la forma *A + sustantivo* (*volver al mundo/tornare nel mondo; dar pie a las fantasías/accendere le fantasie; atendiendo a/mi occupo di*) como en la forma de verbos incoativos con infinitivo. Los verbos incoativos expresan el arranque de una acción y son representados por perífrasis.

En los ejemplos ofrecidos se pueden identificar también los verbos incoativos en su dimensión de cambio de estado: *ya no había vuelto a llamarse Glora Star/ormai non si chiamava più Gloria Star; se echó a reír/scoppiò in una risata; llegaron a echarla a pa-*

tadas/l'avevano cacciata via a pedate. La preposición A española es utilizada también para expresar una modalidad: *se toma el aire a gusto/si può prendere una boccata d'aria; el corazón le latía a toda prisa/il cuore le batteva furiosamente in petto; pelo peinado a raya/pettinato con la riga da una parte*. La expresión *a lo mejor* es traducida frecuentemente con *forse*. Además, A introduce el complemento indirecto que se corresponde con la preposición A italiana (*¿me lo vas a contar a mí?/a me lo raccontì?*). En el caso del complemento indirecto, su equivalente en la lengua italiana puede ser: *con, per, in, con, come a*. Por ejemplo, en *le abríò una lata de comida a Cloud/aprí una scatoletta di cibo per Cloud; inyectar fe a los desesperados/iniettare fiducia nei disperati*. En el caso del complemento indirecto, el pronombre átono redundante es obligatorio con individuos específicos, lo cual no tiene correspondencia en la lengua italiana. Al contrario, en italiano esta construcción se configura como un error gramatical. Lo mismo vale para los verbos reflexivos con pronombre duplicado (*se había negado a sí mismo/si era negato*) y con los verbos pseudo-personales con pronombre dativo: *a mí me encantaría tener dinero/mi piacerebbe da matti avere qualche soldo*. La preposición A se utiliza también para expresar tiempo: *mentre, quotidianamente, pochi instanti dopo*.

La locución adverbial *al fin y al cabo* se puede traducir según la combinación de las palabras que la acompañan: *dopotutto, tra l'altro, alla fin fine*. Con los verbos de sentidos, la preposición A expresa una característica que en italiano se expresa con *di*: *olia a colillas, a cerrado/c'era odore di cicche, di chiuso; mi tarta sabe a jarabe/la mia torta sa di scioppo*. La preposición A se encuentra en numerosas unidades fraseológicas en las cuales la diferencia con el italiano no tiene que ver solo con la preposición, sino que casi siempre supone un cambio de todo el conjunto sintagmático para transmitir el mismo significado: *a falta de pan, buenas son tortas/chi si contenta, gode; no le saca partido a nada/non prova gioia per nulla; se le cayó el alma a los pies/vide tutto nero*. Además, en la última parte de la tabla, se puede ver como A tiene sentido de finalidad, expresado en italiano con la preposición *per* (*el señor Allen fue a recoger el teléfono/il signor Allen andò per rispondere*); indica una parte (*sin consideración a su edad/senza considerazione per la sua età*); sigue en las construcciones ciertos sustantivos y adjetivos (*miedo a la libertad/paura della libertà; totalmente ajenas a la cámara/del tutto inconsapevoli della macchina da presa*); se utiliza en las comparaciones con la palabra *superior* y en italiano corresponde a *di* (*ha sido una fuerza superior a mi/è stato più forte di me*).

### 3.2. RASTREO TRADUCTOLÓGICO DE LA PREPOSICIÓN EN Y ANÁLISIS CONTRASTIVO

La ciudad aparece muy confusa en los atlas geográficos	La città è sempre poco chiara sugli atlanti
Sale en las guías turísticas	Compare sulle guide turistiche
No se le ve alegría en la cara	Ma sui volti non si legge l'allegria
Escapándose en coche de oro	Scappando su automobili d'oro
¿Cuando hablara usted en serio, madre!	Quando vi deciderete a parlare sul serio!

Si estaba tan contento en la isla	Stava così bene sull'isola
Con gorro en punta	Con un cappuccio sulla testa
Que escribió Sara en aquel cuaderno	Che Sara scrisse su quel quaderno
Iba en patines	Andava sui pattini
Han vuelto <i>en autobuses</i>	Ritornato <i>con l'autobus</i>
La gente que viaja en el metro de Nueva York	Chi viaggia sulla metropolitana di New York
(circulaban) en patines	(Si muovevano) sui pattini
Había llegado a conocerlas como las rayas marcadas en las palmas de su mano	Le conosceva come le linee sul palmo della sua mano
Se sentaba en una banqueta del cuarto de baño	Si sedeva sopra una panchetta del bagno
(Aun viejo) sentado en unas escaleras	Un vecchio seduto sulle scale
Se iniciaba en un declive	Cominciava su di un pendio
Recaían siempre en victimas femeninas	Ricadevano sempre su vittime femminili
(Me gustaría ir) en coche de caballos	(Mi piacerebbe di più) andare sopra un calesse
Un disco puesto en el pick-up	C'era un disco sul piatto del giradischi
Se detuvo en el umbral	Si trattenne sulla soglia
Le mandó un beso en la punta de los dedos	Le mandò un bacio sulle punta delle dita
Le servía de pista	La mettevano sulla buona strada
Llegó a sentarse en sus rodillas	Arrivò a sederglisi sulle ginocchia
Se había tumbado vestida en la cama	Si era buttata sul letto vestita
Sabía leer el porvenir en la palma de la mano	Sapeva leggere il futuro sul palmo della mano
(sus ojos brillando) en el rostro pálido	(i suoi occhi Neri che brillavano) sul volto pallido
Lo tenía colgado en el respaldo de la silla	Lo aveva appeso sulla spalliera della sedia
Puede salir a la barandilla que tiene en la corona	Puoi affacciarti alla ringhiera che ha sulla corona
Vio un banco y se sentó en él	Vide una panchina e vi si sedette
Hundió la cabeza en su regazo	Abbandonò la testa sul suo grembo
Sentado en el banco	Seduto sulla panchina
(Desplazándose) en escaleras metálicas	(Spostandosi) su scale mobili lucenti
Deslizándose en trineo	Scivolare sulle slitte
Quedó acreditado en foto de prensa	Come era emerso dalle fotografie sui giornali
¿En la boca del estómago?	Alla bocca dello stomaco?
Emblema que estaba marcado en el papel	Emblema che contrassegnava la carta
Le pedía a Peter que le dejara en Central Park	Chiedeva a Peter di lasciarlo a Central Park
Deseo de llorar en brazos de alguien	Desiderio di piangere tra le braccia di qualcuno
Lo decidiría en la calle	Avrebbe deciso per strada
Allí en el suelo junto a los suyos	Lì per terra, vicino alle sue

Vamos a cruzar en esa dirección	Attraversiamo da quella parte
Llevaban lacitos en el pelo	Avevano nastri fra i capelli
Había bajado en mi ayuda	Era scesa da me per aiutarmi
Rod duerme en su casa	Rod dorme a casa
En la confluencia del Hudston con el East River	Alla confluenza dell'Hudson con L'East River
(una parada) en la base de la estatua	Una sosta ai piedi della statua
Cuando estés en cama	Quando sarai a letto
(anillo) en el dedo índice	(anello) infilato all'indice
En el exterior de Manhattan	Fuori, a Manhattan
Pensando en sus negocios	Pensando ai loro affari
¿En que estará pensando?	A che cosa stara pensando?
Solo piensan en aumentar su riqueza	Pensano solo ad aumentare la propria ricchezza
Mejor, no pienses en nada	Meglio, non pensare a nulla
Conocer Manhattan se había convertido en una obsesión	Conoscere Manhattan era diventata una ossessione
La discusión de sus padres se convertía en una música de fondo	La discussione dei suoi genitori diventava una musica di sottofondo
¿No te vieron convertirte en estatua?	Non ti abbiano visto diventare una statua?
Se empeña en no venirse con nosotros	Vi ostinate a non venire con noi
En cuanto se metían en la cama	Quando andavano a dormire
Se había metido un ratito en un bingo de barrio	Era andata un attimo a giocare al bingo
Metida en el escenario	Prendere parte
Se metió en su <i>limousine</i>	Montò sulla sua limousine
Se quedaba prendida en uno de aquellos extravagantes charlatanes	Veniva catturata da uno di quei ciarlatani stravaganti
Tardó mucho en dormirse	Tardò molto ad addormentarsi
La gente la quería porque no caía en ese defecto	La gente l'amava perché non faceva mai l'errore
El comisario se fijó en aquellos dedos	Il commissario osservò quelle dita
Nadie se fijaba en ellas	Nessuno si accorgeva di loro
Pero he quedado en ir	Siamo rimasti intesi che sarei andata
Eso es lo que se empeñaba en creer	Era questo che si ostinava a credere
Si te empeñas en sufrir	Se proprio vuoi soffrire
Se encontraba ocupado en la revisión	Era occupato a revisionare
Se miró en el espejo	Si guardò allo specchio
Para concentrarse en el recuerdo	Per concentrarsi sul ricordo
Con los ojos fijos en la ventanilla	Con gli occhi fissi sul finestrino
Isla en forma de jamón	Isola a forma di prosciutto
En forma de cruz	A forma di croce
No son normales en una niña de tres años	Non sono normali per una bambina di tre anni

Resultar tan indispensable en El Dulce Lobo	Sarebbe diventato così indispensabile per il Dolce Lupo
Se estuvo gastando todos sus ahorros en ir a oír cantar	Aveva speso tutti i suoi risparmi per andare a sentirla cantare
A estos personajes no los sacaban en sus conversaciones	Non tiravano mai fuori questi personaggi durante le loro conversazioni
En los tiempos de Aurelio	I tempi di Aurelio
En 1905	Il 1905
Había sido, en fin	Insomma, era stato
En seguida	Subito
Así que en el fondo acabó dándole igual	Alla fine per lei faceva lo stesso
Eran fosforescentes en la oscuridad	Erano luminose al buio
Dándole muchos tirones en el pelo	Tirandole i Capelli
Aunque desaparezcan en la realidad	Anche se svaniscono dalla realtà
Llevar la ropa en desorden	Indossavano vestiti sgualciti
Fue media hora que se pasó en un vuelo	Quella mezz'ora era volata
Suplicaba la niña en un susurro fervoroso	Supplicava la bambina con un sussurro fervoroso
Nos habremos cruzado en los ascensores	Ci saremo incrociati con gli ascensori
Le susurró en voz baja	Sussurrò a bassa voce
Estoy gastando saliva en balde	Sto sprecando il mio fiato
Se lo toman tan en serio que se olvidan de vivir	Lo prendono così sul serio da dimenticarsi di vivere
Lo que es tener un problema en serio	Avere un problema serio
Has intentado enamorarte en serio	Hai mai cercato di innamorarti sul serio?
(dejó la mano izquierda) en suspensión	Tenne la mano sinistra sospesa
Era en verdad excelente	Era davvero invitante
(llamarla) en tono autoritario	(Chiamarla) con un tono autoritario
Jersey en punto rojo	Un maglione rosso
Tenía pintados por las paredes en rojo unos pájaros	Sulle pareti rosse erano dipinti uccelli
La advertía la señora Allen en voz baja	Lo richiamava la signora Allen sottovoce
Empezó a cantar en voz en cuello	Comincio a cantare a squarciagola
Lo estaba paladeando con los ojos en blanco	Assaporarlo rovesciando gli occhi al cielo
Embutidos en sus calcetines blancos	Imbottite con i calzettoni Bianchi
En previsión de robos e incendios	Per precauzione contro furti e incendi
Uno de los técnicos especializados en sonido	Uno dei tecnici del suono
Acabó echando mano de él para todos	Aveva finito per fare ricorso a lui per ogni cosa
Parece que viene en buen plan	Sembra gentile
En comparación con la aventura	Paragonato con l'avventura
Se detuvo en seco	Si fermò di colpo

Seguir siempre con la vista fija en el camino	Avrebbe tenuto sempre gli occhi fissi sulla strada
No en vano	Non per nulla
En efecto	Infatti
Él tenía interés en llegar antes	Lui aveva interesse ad arrivare prima
Meterla en un lío	Se cercava di fregarla

Tabla 2 – Resultados del rastreo traductológico de la preposición EN en la novela *Caperucita en Manhattan*

Como se puede observar, la preposición española EN cuenta –en elevada medida– con la correspondencia en lengua italiana *su* que es un locativo y puede hacer referencia bien a un lugar físico o figurado, bien a la posición de un objeto, o de una parte perteneciente a un conjunto. Siempre en su dimensión locativa, esta preposición puede tener varias opciones traductológicas: *en la boca/allà bocca*; *estaba marcado en el papel/contrassegnava la carta* (ausencia de la preposición); *en el pelo/fra i capelli*; *en la calle /per strada*; *en el espejo/allo specchio*.

Existen, además, verbos preposicionales que abogan por la construcción con EN para el complemento objeto o para otros complementos del verbo; EN se emplea asimismo en unidades fraseológicas y en este caso, la traducción italiana difiere casi por completo de la versión original en español (*parece que viene en buen plan/sembra gentile*; *en seco/di colpo*).

En algunos casos, EN se traduce con *per*: al hablar de personas y de sus cualidades (*no son normales en una niña de tres años/non sono normali per una bambina di tre anni*) y cuando expresa una finalidad como en *se estuvo gastando todos sus ahorros en ir a oír cantar* (énfasis en la situación-locación) y *aveva speso tutti i suoi risparmi per andare a sentirla cantare* (énfasis en la finalidad). También para la preposición EN se puede notar el elemento ya presente en la preposición anteriormente analizada: muchas veces en español se prefiere la preposición a su correspondiente nombre, adjetivo o adverbio italiano, por ejemplo en el caso de expresiones que indican una característica, un adjetivo o una modalidad (*jersey en punto rojo/un maglione rosso*; *en voz baja/sottovoce*; *con los ojos en blanco/alzando gli occhi al cielo*). EN puede introducir una expresión temporal sin una correspondencia preposicional total en italiano: *en 1905/il 1905*. Algunas veces EN puede ser trasladada mediante *da*, pero no se trata de un uso muy frecuente: *aunque desaparezcan en la realidad/anche se svaniscono dalla realtà*. Cuando se habla de forma, EN se traduce casi siempre con *a*: *en forma de jamón/a forma di prosciutto*.

### 3.3. RASTREO TRADUCTOLÓGICO DE LA PREPOSICIÓN DE Y ANÁLISIS CONTRASTIVO

Déjate de psiquiatras	Lascia perdere gli psichiatri
¡Pues vaya de unas tonterías que habláis las mujeres!	Voi donne ne dite di scemenze!

El lobo tenía una cara tan buena, tan de estar pidiendo cariño	Il lupo aveva una faccia così buona e così bisognosa di affetto
Ella a Aurelio se lo figuraba de otra manera	Aurelio se lo immaginava diverso
Siempre está de broma	Scherza sempre
Le había costado a Sara la separación de su abuela y el librero	Era costata a Sara la separazione tra sua nonna e il librario
Venía de ganar en el bingo	Aveva vinto al bingo
Cogida del brazo de un hombre	Al braccio di un uomo
Y parecía echarla de menos	Sembrava sentirne la mancanza
Valiéndose de un bolígrafo gordo	Usando un grosso pennarello
Ella se encogía de hombros	Lei si stringeva nelle spalle
(la señora Allen) cambiaba de conversación	(La signora Allen) cambiava argomento
Sara dejaba de escuchar	Sara non l'ascoltava più
Toda una página y no podía dejarla de mirar	Occupava una pagina intera e non riusciva a smettere di guardarla
La señora Allen acababa de comprar	La signora Allen aveva appena comprato
Cuando estaba a punto de nacer Sara	Quando Sara stava per nascere
¿Qué haces ahí enterándote de lo que no te importa?	Che fai lì impalata a sentire cose che non ti riguardano?
No hacía más que acordarse de la abuela	Pensava alla nonna
Nos impide disfrutar del camino por donde vamos andando	Ci impedisce di godere del camino che stiamo percorrendo
(no es bonito) como cuando se gozaba de su tacto como del de una joya	(Non è più bello) come quando era piacevole al tatto come un gioiello
Oyéndola hablar de lo que fuera	Preferiva continuare a sentirla parlare, non importa di cosa
Para dotar de realismo a aquellas fingidas velas	Per conferire realismo a quelle finte candeline
Me moriría de un atracón	Mi verrebbe un infarto
(La tarta de fresa) cambiara de sabor	(il fatto che la torta di fragole) cambiasse sapore
Se había dado cuenta de que su madre	Si era accorta che sua madre
La niña sin soltarse de la mano de miss Lunatic	La bambina, pur non lasciando la mano di Miss Lunatic
Como si quisiera abrigarla de peligros	Come se volesse proteggerla da pericoli
Debían de ser cosas buenas	Doveva essere qualcosa di bello
Los niños suelen encapricharse precisamente de lo más peligroso	I bambini si fissano proprio sulle cose più pericolose
No se acababa de entender	Non riusciva a comprendere
Sin dejar de sentir	Continuando a sentire
El sándwich de tu padre	Il panino per tuo padre
Era un mueble de secretos	Era un mobile per i segreti
Ya tiene edad de hacerse cargo de las cosas	Ormai ha l'età per occuparti di ogni cosa
(Recorría Manhattan) a unas velocidades impropias de su edad	(Attraversava Manhattan) a una velocità impropria alla sua età



Lo que pasa es que tengo poco tiempo de leer	Purtroppo ho poco tempo per leggere
Recogiendo del suelo su sombrero	Raccogliendo il cappello da terra
Había sacado de una bolsita (el aroma) que salía del interior por las puertas giratorias	Aveva tirato fuori da un borsellino (L'aroma) che arrivava attraverso le porte girevoli
(Los diferentes oficios) le alejaron de Nueva York	(I diversi impieghi) che lo allontanarono da New York
El puesto de segundo de a bordo	Posto di copilota
Enterarse de por dónde iba	Capire dove stava andando
Estaba su nieta recogiénolo del suelo	La nipote stava raccogliendoli sul pavimento
Lo que ella quería de mayor era ser actriz	Lei che da grande voleva diventare una attrice
(a punto de producirse) accidentes de coches	(In cui stavano per verificarsi) incidenti automobilistici
De principios de diciembre	Ai primi di dicembre
Hay que hacerlas de vísperas	Se le fai la sera prima (le torte restano più buone)
Decían de vez en cuando	Se dicevano ogni tanto
En estado de letargo	In letargo
(la receta) de la tarta de fresa de toda la vida	(La ricetta) della gloriosa torta di fragole
Llevaba cerca de una hora esperando	Era quasi una ora che aspettava
De momento	Per il momento
De pronto	Però a un tratto
De repente	A un tratto
De vez en cuando	Ogni tanto
Interrumpido de vez en cuando por el correteo de alguna ardilla	Interrotte di tanto in tanto dalle scorribande
Un niño de tierna edad	Un bambino in tenera età
Debe de ser de tu edad	Deve avere la tua età
Ha salido lista de verdad	È davvero sveglia
Donde vivía de verdad en Brooklyn	Dove viveva in realtà
Te lo digo de verdad	Dico sul serio
No era seguro que se hubieran visto de verdad	Non poteva giurare di averli visti per davvero
Luego de reojo	Con la coda dell'occhio
Se levantó varias veces de puntillas	Si alzò diverse volte in punta di piedi
Fuertemente de la mano	Stringendola forte
Hablando muy de prisa	Parlando in fretta
Se había puesto triste de pronto	Era diventata improvvisamente
De un salto	Con un salto
Iba ella de la mano de su madre	Con la mano in quella di sua madre
Puede llorar de tres maneras distintas	Può piangere in tre modi diversi

Persiguiéndose de forma vertiginosa	Disegni che si inseguivano in modo vertiginoso
Está resuelto de un modo bastante ingenioso	È una soluzione abbastanza ingegnosa
Un hombre que estaba de pie	Un uomo che se ne stava in piedi
Cayó de rodillas	Cadde in ginocchio
Las explicaciones del chico de la gorra	Le spiegazioni del ragazzo col berretto
Un viaje de muchas millas	Un viaggio lungo molte miglia
¿Viviste allí de soltera?	Prima di sposarti?
¿Por qué pones cara de mártir?	Perché fai quella faccia da martire?
La librería de viejo de Aurelio	Il negozio di libri usati
Como los gusanos de luz	Come le lucciole
(Le parecía) una casa de novela	(Le) sembrava una casa da romanzo
Aunque fuera teniendo que reducirse de tamaño como Alicia	Sebbene per farlo dovesse rimpicciolirsi come Alice
A una rubia de tacones altos	Una bionda coi tacchi alti
Estoy tomándome una copita de licor	Sono (nel bar) a bere un bicchierino
El hombre del tiempo ha dicho que la nubosidad es variable	Le previsioni hanno detto che la nuvolosità è variabile
Dentro de un florero de china	Dentro un vaso da fiori cinese
El mantel era de flores bordadas	La tovaglia aveva fiori ricamati
La abuela fue al cuarto de estar	La nonna andò in soggiorno
Era de papel cuadriculado	Era a quadretti
Que imitaban cacharros de cocina	Che imitavano le stoviglie
Historias de miedo	Storie spaventose
Mentiras de niña chica	Bugie da bambina piccola
Estaba sentada ante un tocador de tres espejos	Era seduta davanti a un mobile di toeletta con tre specchi
Una bolsita de raso azul bordada de lentejuelas	Una borsetta di raso blu ricamata con le paillettes
Iba siempre de manga corta	Era sempre in maniche corte
Está en un grado de descuido	E' trascurato da far pietà
Como de aleteo de pájaros	Come un frullo d'ali
Sara estrenó un conjunto de falda plisada	Sara sfoggiava un completino con la gonna plissettata
Rod estaba igual de zoquete	Rod era sempre lo stesso giuggiolone
De patas muy largas	Dalle zampe lunghissime
El papelito era de color malva	La strisciolina era color malva
En una computadora de bolsillo	Con una calcolatrice tascabile
La primera maqueta solo tenía nueve pies de alto	Il primo plastico che era alto meno di tre metri
Arrastraba un coche de niño vacío	Trascinava una carrozzina vuota
(Era un modelo) de gran tamaño	(Era un modello) piuttosto grande
(solía ir) con aire de balada o de nana	(Era solita) canticchiare ballate o ninnenanne

Los garitos de mala vida	Le bische malfamate
Con un aire de altivez e independencia	Con un'aria altera e indipendente
Diseñadores de mayor renombre	I designer più rinomati
Las máquinas de aire acondicionado	I macchinari per l'aria condizionata
Flojedad, como de convalecencia juvenil	Mollezza, come una convalescenza giovanile
El hombre del pelo gris	L'uomo dai Capelli grigi
Miss Lunatic llevaba mediado su segundo cóctel de champán	Miss Lunatic era a metà del suo secondo cocktail champagne
El traje de punto	Il vestito
Andando sola camino de salida	Mi sono trovata a camminare da sola verso l'uscita
El hombre-muñeco de pelo rizado	L'uomo burattino con i Capelli ricci
Un batido grande de chocolate	Un frappè grande al cioccolato
Saca la bolsita de lentejuelas	Prende la borsetta con le paillettes
Retirando la servilleta de cuadros	Sollevando il tovagliolo a quadretti
Extraña mendiga del pelo blanco	La mendicante dai Capelli Bianchi
Guionista de cine	Sceneggiatore
Ciegos de aburrimiento	Annoiati a morte
La misma cara de diablo	La stessa faccia birichina
Las cortinillas de gasa corridas	Le tendine semitrasparenti abbassate
En la cristalera de separación	Nel vetro divisorio
La niña que había inventado lo del jamón	La bambina che aveva inventato il prosciutto
Quería probar de lo de más allá	Se voleva assaggiare dell'altro
Figuraban ser velas de las que se ponen en las tartas	Volevano rappresentare le candeline che si mettono sulle torte di compleanno
Ya ves lo de la tarta de fresa	Hai visto con la torta di fragole
Nunca he encontrado un quehacer más importante que el de escuchar historias	Non ho mai trovato niente di più importante dell'ascoltare storie
Para disponer de mí	(Chi crede di essere) per usarmi così?
¿Ha salido algo de eso?	È venuto qualcosa?
Tenía que elegir entre lo de fuera y lo de dentro	Doveva scegliere tra quello che sta fuori e quello che sta dentro
(le daba rabia) haberse perdido lo de fuera	(Le faceva rabbia) essersi persa quello che stava fuori
Lo de menguar de tamaño solo lo podía conseguir Alice	Soltanto Alice poteva rimpicciolire
Se imaginaba limpiando de trastos	Immaginava se stessa mentre sbarazzava dalle carabattole
Le hacía gracia que la abuela se defendiera de los sermones de su propia hija	Trovava divertente che la nonna si difendesse dalle prediche di sua figlia
Como caerse de las nubes	Come cadere giù dalle nuvole
Asombrada de su eficacia y de su actividad	Era stupita dalla sua efficienza e dal suo dinamismo
¿Pero de qué te ríes?	Ma perché stai ridendo?

Como era de esperar	Com'era da aspettarselo
El taxista se abstuvo de hacer más comentarios	Il tassista si astenne dal fare commenti
Que de todo haces un folletín	Fai sempre delle tragedie!
Dádoselas de conquistador	Si dava arie da gran conquistatore
Ella se encogía de hombros	Lei si stringeva nelle spalle
Aunque quedara alguna de sobra	Anche se ce n'era qualcuna non utilizzata
La nuera del viejo dueño se pusiera de parto	(Prima che) la nuora del vecchio proprietario avesse le doglie
Obsesionado por culpa de la tarta	Aveva l'ossessione per la torta di fragole
Sabes de sobra	Sai benissimo
La echaba furiosamente de menos	Sentiva furiosamente la sua mancanza
En esto no estoy de acuerdo	Non sono affatto d'accordo
Me has salido una amiga de merengue	Mi è toccata un'amica pappamolle
Las conversaciones de antemano	Le conversazioni in anticipo
No hay que hacer caso de las prohibiciones	Non bisogna mai dar retta ai divieti
Deseando que se haga de noche	Desidero che diventi buio
La mirada parecía de fiar	Lo sguardo sembrava ispirare fiducia
Había material de sobra	Ce n'era in abbondanza
Ya estoy de luz de neón	Ne ho fin sopra i Capelli della luce al neon
Me alegro de que le pille de camino	Sono contenta di essere sulla sua strada
De los viajeros que esperaban en el andén	Tra i viaggiatori che aspettavano sulla panchina
La conversación versaba sobre la comparación de unos manjares con otros	La conversazione verteva sul confronto tra le portate
De todos los empleados	Fra tutti gli impiegati
De postre trajeron unos pastelitos	Come dessert portarono pasticcini
Como de barquillo duro	Simili ai millefoglie
Los testigos presenciales del suceso	I testimoni che avevano assistito al fatto
Que podía servir de dato para relacionar	Poteva mettere in rapporto
Se llevó un poco de susto	Si spaventò un poco
¿Tan rico, que tenía de todo?	Così ricco che aveva tutto?
Aunque no sé si servirá de mucho	Anche se non credo che gli servirà a molto
¿Cómo sabe usted tanto de niños?	Come fa a sapere tante cose sui bambini?
Rodeada de una pelambreira	Circondata da una zazzera
Se iluminaban, nimbadadas de unos efectos	(Si accendevano e si spegnevano) dando origine a effetti
(un vasto reino) surcado de tuberías	Un vasto regno solcato da tubature
(uniforme) estampado de fresas y manzanas	Uniforme, con le fragole e le mele stampate
(andaba encorvada) presa de un súbito desaliento	Camminava curva in preda a un improvviso scoramento
(una nube de fotógrafos) estaba pendiente de la llegada de aquella mujer	Un nugolo di fotografi aspettava ansiosamente l'arrivo
Totalmente impropia de su edad	Assolutamente inadatto alla sua età

(El público) es muy poco amigo de innovaciones	(il pubblico) è poco propenso alle innovazioni
Tocaban unas variaciones del Let it be	Erano variazioni su Let it be
Al borde de la acera	Sul bordo del marciapiede
Los escalones que se separaban unos niveles de otros	Scalini che separavano i livelli
(díselo) sin exageración de focos	Dillo, non devono esserci troppe luci
Con el ayudante de rodaje	Con l'aiuto-regista
El conato de nieve	La timida nevicata
Una pausa del rodaje	Stavamo facendo una pausa
Donde estaba la raya de separación	Dov'è la linea che separa
Evitándole toda clase de peligros	Evitandole ogni rischio
En sus excursiones de pesquisa pastelera	Negli spostamenti per le sue indagini dolciarie
(el automóvil) sortea audazmente toda clase de obstáculos	(L'automobile) schiva audacemente ogni ostacolo
¡Que porquería de vida!	Che vita schifosa!
Las ganas de freírle a preguntas	Il desiderio di fargli domande su domande
Se han perdido las referencias del tiempo	Si sono perduti i riferimenti temporali
Transcurrió sin sentir el viaje de vuelta	Passò in un battibaleno
Escribir guiones de cine	Scrivere copioni per il cinema
De uniforme prestado	Con un'uniforme prestata
(Parecía un barrio) más de pueblo	(Sembrava un quartiere) ma molto più popolare

Tabla 3 – Resultados del rastreo traductológico de la preposición *DE* en la novela *Caperucita en Manhattan*

*DE* se configura como una de las preposiciones de la lengua española más rentables y a partir de la tabla de arriba, se pueden observar numerosos matices semánticos: en la forma *verbo + DE* con verbos que requieren la preposición para introducir el complemento objeto o sus complementos; como marcador modal, especialmente típico del lenguaje coloquial como en *déjate de psiquiatras/lascia perdere gli psichiatri*. Existe también el uso de *DE* en modismos y frases hechas como en *echarle de menos/sentirne la mancanza y saber de sobra/sapere benissimo*. En *de repente/improvvisamente*, *DE* adquiere un matiz temporal o *de reojo/con la coda dell'occhio* un matiz modal, aportando una característica o una especificación a la acción. La expresión fija *de verdad* cuenta con diferentes opciones de traducción en italiano según exprese la modalidad de una acción (*in realtà, sul serio; per davvero*) o ejerce la función de intensificador de un verbo y su predicado (*ha salido lista de verdad/ è davvero sveglia*).

Además, *DE* puede acompañar al verbo 'deber' para expresar probabilidad: *debían de ser cosas buenas/doveva essere qualcosa di bello*. Siempre en el ámbito verbal, *DE* puede construir una perífrasis verbal o verbos incoativos en la formación: *verbo + DE + infinitivo* y eso generalmente se corresponde con la idea de acabar

de realizar una acción, estar a punto de, finalizar una acción, dejar de hacerla, reanudarla después de una interrupción o llegar a iniciar algo. En la oración *como si quisiera abrigarla de peligros/come se volesse proteggerla da pericoli* se nota la traducción de la preposición *DE* en la preposición italiana *da*. Este fenómeno se produce tanto en formas verbales como ‘proteger de’, ‘librar de’, como con adjetivos (*uniforme estampado de fresas y manzanas/uniforme con fragole e mele stampate*). Con verbos que expresan acciones, *DE* puede tener un significado locativo (*había sacado de una bolsita/aveva tirato fuori da un borsellino*) y si es una expresión de tiempo, puede concretarse en italiano mediante algunos sustantivos y preposiciones: *de principio/ai primi di*; *de vísperas/la sera prima*; *de letargo/in letargo*; *de vez en cuando/ogni tanto*. *DE* a veces se traduce en italiano a través de la preposición *per* con un sentido de finalidad (*el sándwich de tu padre/il panino per tuo padre*); en algunos casos conforma una expresión que describe un estado o la manera de ocupar una posición: *de pie, de rodilla, de un salto*. *DE* puede expresar una especificación cuando se refiere a un sustantivo (*las explicaciones del chico de la gorra/le spiegazioni del ragazzo col berretto*) o cumple la función de describir y adjetivar (*un viaje de muchas millas/un viaggio lungo molte miglia*). En el caso de una característica, el *DE* se traduce en mayor medida como *con y a* (*de uniforme prestado/con un uniforme prestato; a quadros/a quadri*).

Existe también la traducción de *DE* como característica con la preposición italiana *di* pero con una frecuencia menor. Se puede notar, como en el caso de las demás preposiciones analizadas, que el español se decanta más por la construcción *sustantivo + DE + sustantivo*, mientras que el italiano lo expresa con una sola palabra (adverbio, adjetivo o nombre), como en *gusano de luz/luciole* y *librería de viejo/libri usati*.

Cuando *DE* se utiliza como partitivo e indica una parte de un conjunto o una tipología, se traduce con *tra*; puede señalar una cantidad (*servirá de mucho/servirà a molto*) y puede indicar un tema y también una selección dentro del tema abordado (*¿cómo sabes tanto de niños?/come sai tanto di bambini?*). Cuando *DE* se utiliza en español en el ámbito de una comparación con referencia a una situación, un estado o una acción, la lengua italiana no necesita preposiciones (*parecía un barrio más de pueblo/sembrava un quartiere ma molto più popolare*). En la construcción *la niña que había inventado lo del jamón/la bambina che aveva inventato il prosciutto* se puede notar un ejemplo del uso frecuente de *pronombre + DE* en español que en italiano se traduce con *quello/quella che* u omitiendo directamente la referencia pronominal (*lo de menguar de tamaño solo lo podía conseguir Alice/soltanto Alice poteva rimpicciolire*).

#### 4. A MODO DE CONCLUSIÓN

El recorrido realizado por la traducción del régimen preposicional entre el español y el italiano mediante el corpus creado a partir de la novela *Caperucita en Manhattan* de la escritora salmantina Carmen Martín Gaité ha sido sumamente interesante, pues nos ha permitido apreciar cómo el proceso traductológico no

es lineal, a pesar de ser el italiano y el español dos lenguas pertenecientes a la familia romance y a tener el latín como tronco común. La mencionada falta de linealidad ha quedado corroborada por los ejemplos proporcionados en el análisis de la traducción de las tres preposiciones examinadas. Siguiendo la propuesta taxonómica de las técnicas de traducción de Hurtado Albir (2011: 269), se han empleado varios procedimientos traslativos, especialmente cambios de categoría gramatical (transposición), de perspectiva o punto de vista (modulación), comprensiones lingüísticas, ya que en varios casos las preposiciones españolas quedan eliminadas en italiano en pro de construcciones simples.

Ahora bien, el análisis efectuado demuestra claramente que no existe una modalidad estándar de llevar a cabo la traducción del régimen preposicional del español hacia el italiano. Además, las consideraciones expuestas pueden resultar válidas también al abordar el proceso de traducción del italiano hacia el español, así como en lo concerniente a otras preposiciones que por motivos de espacio no hemos tenido en cuenta en el presente artículo. Lo que sí ha quedado claro a partir de nuestro estudio es que al realizar las operaciones traductológicas, se producen numerosas disimetrías morfosintácticas<sup>4</sup> que es posible sistematizar según las tipologías y categorías en las que las preposiciones estén enmarcadas, como hemos hecho con anterioridad.

Para finalizar, creemos poder desmitificar el falso tópico según el cual traducir entre lenguas afines es una actividad que se ve favorecida por el parentesco lingüístico y filogenético. La cercanía entre estos dos idiomas se perfila como un gran desafío para el traductor, quien no puede permitirse el lujo de infravalorar ningún elemento de la lengua de la que y hacia la que traduce, pues los calcos y las interferencias lingüísticas siempre están al acecho.

4 También es interesante el análisis de este fenómeno lingüístico en la traducción oral, concretamente, en la interpretación simultánea español-italiano. A este respecto, véase el volumen de Mariachiara Russo (2012) titulado *Interpretare lo spagnolo. L'effetto delle dissimmetrie morfosintattiche nella simultanea*.

- Bermejo Calleja F. (2007) "Preposiciones y adverbios locativos: identidad y contraste en los diccionarios bilingües español-italiano/italiano-español". En Félix San Vicente (ed.), *Perfiles para la historia y crítica de la lexicografía bilingüe del español*, Monza, Polimetrica International Scientifical Publisher, pp. 159-183.
- Bertazzoli R. (2015) *La traduzione: teorie e metodi*, Roma, Carocci.
- Caballero Rubio M. del C. y Corral Hernández G. B. (2006) "Las preposiciones locativas en español e italiano". En *Actas del VIII Congreso Internacional de ASELE*. Recuperado de: [https://cvc.cervantes.es/ensenanza/biblioteca\\_ele/asele/pdf/08/08\\_0193.pdf](https://cvc.cervantes.es/ensenanza/biblioteca_ele/asele/pdf/08/08_0193.pdf) [Consulta: 12.09.2021].
- Cano Gines A. y Flores Ramírez A. (1998) "Instrucciones gramaticales y uso de las preposiciones en castellano". En *ASELE, Actas IX*. Recuperado de: [https://cvc.cervantes.es/ensenanza/biblioteca\\_ele/asele/pdf/09/09\\_0522.pdf](https://cvc.cervantes.es/ensenanza/biblioteca_ele/asele/pdf/09/09_0522.pdf) [Consulta: 10.09.2021].
- Hurtado Albir A. (2011) *Traducción y Traductología. Introducción a la Traductología*, Madrid, Cátedra.
- Katerinov K. (1980) (ed.) *Uso delle preposizioni in italiano e in spagnolo*, Perugia, Guerra.
- Lenarduzzi R. J. (2007) "Preposiciones y modalidad". En Félix San Vicente (ed.), *Partículas/particelle. Estudio de lingüística contrastiva español e italiano*, Bologna, CLUEB, pp. 71-88. Recuperado de: <http://amsacta.unibo.it/6062/1/san%20vicente%20%28ed%29%20Actas%20CLUEB%20Particulas%2017-06.pdf> [Consulta: 15.09.2021].
- Luque Toro L. (2019) "Las locuciones adverbiales introducidas por la terna preposicional *a, de, en*". En *Léxico Español Actual VI*, Venezia, Editrice Cafoscarina, pp. 133-143.
- Martín Gaité C. (1990) *Caperucita en Manhattan*, Madrid, Siruela.
- Martín Gaité C. (2011) *Caperucita en Manhattan (Cappuccetto Rosso a Manhattan)*, Milano, Salani, [trad. de M. Finassi Parolo].
- Minguell A.E. (2009) "La preposición: categoría léxica y funcional. Aportes para la enseñanza del español", *SIGNOS ELE*. Recuperado de: <http://p3.usal.edu.ar/index.php/ele/article/view/1224> [Consulta: 15.09.2021].
- Osimo B. (2010) *Manuale del traduttore* (seconda edizione), Milano, Hoepli.
- Real Academia Española y Asociación de Academias de la Lengua Española (2010) *Nueva Gramática de la Lengua Española, Manual*, Buenos Aires, Planeta.
- Regueiro Rodríguez M. L. (2014) "El desafío preposicional en ELE", *Revista Nebrija de lingüística aplicada a la Enseñanza de las lenguas*, 16, pp. 1-10. Recuperado de: <https://www.nebrija.com/revista-linguistica/el-desafio-preposicional-en-ele.html> [Consulta: 15.09.2021].
- Russo M. (2012) *Interpretare lo spagnolo. L'effetto delle dissimmetrie morfosintattiche nella simultanea*, Bologna, Clueb.
- Solsona Martínez C. (2007) *Análisis de Errores y de la Interlengua en el aprendizaje de las preposiciones italianas por parte de hispanohablantes*, Tesis Doctoral, Universidad de Sevilla.
- Trovato G. (2013) "El régimen preposicional en español e italiano: breve estudio contrastivo e traductológico", *OGIGIA (Revista Electrónica de Estudios Hispánicos)*, 13, pp. 19-34. Recuperado de: [https://www.academia.edu/23099154/El\\_r%C3%A9gimen\\_preposicional\\_en\\_espa%C3%B1ol\\_e\\_italiano\\_breve\\_estudio\\_contrastivo\\_y\\_traductol%C3%B3gico](https://www.academia.edu/23099154/El_r%C3%A9gimen_preposicional_en_espa%C3%B1ol_e_italiano_breve_estudio_contrastivo_y_traductol%C3%B3gico) [Consulta: 20.09.2021].



Trovato G. (2018) *Lingüística española y traducción desde la contrastividad*, Roma, Aracne.

#### WEBGRAFÍA

Diccionario de la Real Academia Española [<https://dle.rae.es/>]

Diccionario CLAVE [<http://clave.smdiccionarios.com/app.php>]

Dizionario Treccani [<http://www.treccani.it/vocabolario/>]

## Sezione/Section 2

Esperimenti di *Distant Reading*.  
Estrazione, analisi e visualizzazione  
di dati linguistici da corpora letterari



# Corpora e storia della lingua

MICHELE A. CORTELAZZO  
Università di Padova  
[cortmic@unipd.it](mailto:cortmic@unipd.it)

## ABSTRACT

During the last twenty years, in particular at the lexical level, the study of the history of the Italian language has benefited greatly from corpora including literary texts, such as “Letteratura Italiana Zanichelli” or “Biblioteca Italiana” (and, perspectivevely, the Italian section of the ELTeC - European Literary Text Collection). However, those corpora were compiled mainly according to the literary canon, and they include a small number of texts compared to the time span they are supposed to cover. Thus, they can provide a limited account of the evolution of literary Italian in time. Today, a large number of texts are available online (to begin with Google Books), covering a large time span and broad variety of genres. The analysis of such large corpora calls for suitable tools and approaches to ensure their effective distant reading. Consequently, the interdisciplinary co-operation of linguists, statisticians and academics of literature may lead to an innovative approach to the study of corpora to assess the evolution of literary language.

## KEYWORDS

Distant reading, corpus linguistics, diachronic corpora, chronological textual data, history of the italian language

Luca Serianni ha pubblicato, nel 2006, un contributo dal titolo *Gli archivi elettronici e la lessicografia storica*, che riprende un intervento tenuto qualche anno prima a Saarbrücken al convegno tenutosi in occasione del settantesimo compleanno di Max Pfister. Questo contributo rappresenta bene l'entusiasmo che la disponibilità dei primi *corpora* digitalizzati aveva generato negli studiosi, ma al tempo stesso la consapevolezza della loro insufficienza per uno studio pieno della storia della lingua di uno strumento costituito da una selezione di testi che, pur con qualche scelta originale, si basava, necessariamente direi, sul canone tradizionale della storia letteraria. Serianni sottolinea la rilevanza di quello che allora era lo strumento principe per lo studio della lingua italiana, la LIZ, *Letteratura italiana Zanichelli*, per lo studio prima di tutto della lingua poetica, "data la tradizionale vischiosità della tradizione lirica italiana e il rilievo esemplare che vi assumono i grandi classici" (Serianni 2006: 42). Ne rileva anche i limiti, consistenti principalmente nella quasi esclusiva letterarietà del canone di riferimento. Non solo sono totalmente esclusi testi di natura non letteraria (come i testi di natura tecnica, con il risultato che "è scarso il contributo alla conoscenza dei linguaggi settoriali, se non in quanto suscettibili di rifluire nella lingua letteraria, come avviene in qualche misura per il lessico giuridico o per quello medico", Serianni 2006: 42), ma anche nella scelta delle opere di autori inclusi nel repertorio (nota, ad esempio, che di un letterato, ma anche scienziato, come Lorenzo Magalotti è raccolta la *Relazione sulla China* e non i *Saggi di naturali esperienze*, di uno storico come Pietro Giannone la *Vita* e non la *Istoria civile del Regno di Napoli*, di un economista come Antonio Genovesi le *Memorie autobiografiche* e non le *Lezioni di commercio o sia di economia civile*). Queste scelte sono ben comprensibili in un'ottica concentrata sulla storia letteraria, ma riducono l'utilità dello strumento per la storia della lingua.

Nonostante questi limiti, la LIZ, nelle sue diverse edizioni, sempre più accresciute, ha costituito un serbatoio enorme anche per chi intendeva studiare la storia della lingua e, soprattutto, il suo lessico. Più di un risultato interessante è proposto proprio da Serianni (2006).

Certo, leggere questo contributo circa un ventennio dopo la sua prima stesura ingenera un senso di nostalgia e tenerezza, pensando alla soddisfazione che provavamo nell'aver a disposizione un repertorio che, nella sua ultima edizione, comprendeva circa mille testi, interrogabili in tempo reale. Sembrava un'enormità. Ma mille testi per quasi mille anni di storia letteraria significa una media di un testo per anno. Troppo poco per studiare in dettaglio l'evoluzione della lingua che, si sa, è caratterizzata, anche nelle sue realizzazioni scritte, da un alto grado di varietà non solo cronologica (quella che è il centro di interesse nel caso di indagini diacroniche), ma anche di genere testuale, di funzione, di provenienza regionale.

Io stesso, in quegli anni, mi ero basato sul corpus della prima LIZ (quindi su poco meno di 400 testi), per ricostruire la protostoria e la storia della perifrasi

progressiva (Cortelazzo 2007) e mi pareva da una parte di annegare in mezzo ai dati, dall'altra di aver fondato la storia del costruito su una base testuale ineguagliabile. Oggi, la base testuale utilizzata solo una quindicina di anni fa appare ben modesta e certamente da incrementare soprattutto per le fasi più recenti (ottonovecentesche) della storia della nostra lingua.

Dal punto di vista quantitativo, i *corpora* nella piena disponibilità degli utenti non sono aumentati in misura considerevole. La più ampia raccolta aperta e filologicamente attendibile di testi della letteratura italiana, la Biblioteca Italiana (sigla BiBit, sito <http://www.bibliotecaitaliana.it>), comprende più di 1600 opere: si tratta certamente di un notevole incremento rispetto alla LIZ, che però raffina solo parzialmente la densità cronologica dei testi. Del resto, l'obiettivo della BiBit (erede di imprese pionieristiche, come quella avviata dal CIBIT, Centro Interuniversitario Biblioteca Italiana Telematica, fin dal 1996) è pur sempre quella di allestire una biblioteca digitale formata, come si legge nel sito, da "testi rappresentativi della tradizione culturale e letteraria italiana dal Medioevo al Novecento". Cioè anche questa iniziativa prevede, di necessità, l'individuazione di un canone (quello che permette di riconoscere la rappresentatività di un testo rispetto a un altro). Lo stesso corpus otto-novecentesco ELTeC (European Literary Text Collection), costituito nell'ambito del progetto europeo (COST Action) *Distant Reading for European Literary History* (Burnard, Schöch, Odebrecht 2021), prevede per ogni lingua la raccolta di 100 opere, collocate cronologicamente di norma tra il 1840 e il 1920 (per l'italiano ne è stata finora costituita una versione preliminare, di testi tra il 1825 e il 1923), con una densità, dunque, del tutto analoga a quella della LIZ.

Indubbiamente *corpora* come quelli che ho ricordato hanno il vantaggio di essere codificati secondo sistemi standard e, almeno in parte, annotati (per es. in relazione alle parti del discorso o ad alcune categorie di nomi: Schoch et alii 2021; ed altri *corpora* sono più finemente annotati, sotto il profilo sintattico o sotto quello semantico o tematico). Sul piano qualitativo si tratta, quindi, di *corpora* più ricchi di informazioni dei semplici testi digitalizzati; ma la codificazione richiede un impegno che rallenta la predisposizione dei *corpora*, obbliga a trovare stretti criteri di selezione dei testi e ci riporta, quindi, alla prefigurazione di un canone.

Ancor più selettivi sono *corpora* nati con finalità specifiche, come il "PIC, Padua Italian Corpus" (così denominato da Savoy 2018), costituito per identificare le similarità tra i testi di Elena Ferrante e quelli di altri autori italiani a lei contemporanei.

## 2. CORPORA ELETTRONICI PER LA STORIA DELLA LINGUA ITALIANA

I *corpora* letterari finora raccolti sono stati di grandissimo aiuto anche per la storia della lingua italiana. Il fatto di avere a disposizione sul proprio computer dai mille ai milleseicento testi dell'intera storia letteraria italiana, e poterli interro-

gare in tempo reale (senza la necessità di consultarli uno per uno alla ricerca di una parola o di un costrutto) ha modificato profondamente il modo di raccogliere i dati per ricostruire pezzi di storia della nostra lingua. Il vantaggio maggiore riguarda la ricostruzione della storia delle parole, ma anche i costrutti grammaticali più facilmente identificabili per mezzo di definiti elementi formali (come è il caso della citata perifrasi progressiva) hanno avuto una più precisa illustrazione storica.

I criteri che (seguo ancora Serrianni 2006) hanno portato alla costituzione dei *corpora* citati, sono ragionevoli, ma non rappresentano sempre la soluzione migliore per lo storico della lingua. Il primo criterio è quello della rappresentatività storico-letteraria o, in alternativa quello della costruzione di un corpus bilanciato, in base a criteri che non sempre corrispondono ai criteri che sono ugualmente validi per tutte le epoche storiche (basti pensare al bilanciamento in base al genere dell'autore, in una tradizione letteraria che è stata prevalentemente maschile). Il secondo criterio è quello della disponibilità di edizioni affidabili, meglio se recenti, in quanto facilmente digitalizzabili (criterio che oggi è caduto in secondo piano, grazie al perfezionamento degli strumenti di lettura automatica dei testi a stampa originali).

Oggi abbiamo la possibilità materiale di uscire dalle strette della rappresentatività esclusivamente letteraria grazie alla enorme disponibilità in rete di volumi scansionati e resi disponibili almeno in formato pdf (in Google Libri, o nel preziosissimo Internet Archive, o in altre biblioteche digitali a libero accesso, come Hathi Trust o ulteriori raccolte di altre Università di oltreoceano, tutte ricche anche di libri italiani). Fondamentali anche, per l'italiano postunitario, gli archivi storici dei giornali, anche se non tutti pienamente utilizzabili ai nostri fini: è il caso della raccolta del "Corriere della sera", che non permette la ricerca per significanti, dal momento che elabora le richieste con criteri semantici. Ma utilissimi sono gli archivi della "Stampa", del "Piccolo", dell'"Avanti!", dell'"Unità", oltre che di "Repubblica" (se lo studio riguarda la storia dell'italiano più recente). Preziosi anche gli archivi storici di Camera e Senato.

In primo luogo risultano comunque fondamentali le raccolte digitali di volumi, che, per la maggior parte dei testi non più protetti dal diritto d'autore, consentono anche di scaricare la versione pdf del testo. Si tratta di una risorsa di grandissima utilità per chi voglia descrivere con maggiore precisione la configurazione dell'italiano scritto degli ultimi secoli e le dinamiche che si sono realizzate, uscendo dai canoni di matrice letteraria che, ancor più da quando sono risultati disponibili i *corpora* di prima generazione come la LIZ, hanno forzatamente guidato la ricerca sistematica sui testi, basata su risorse automatiche di interrogazione, con un riferimento alla lingua letteraria più forte di quello che si verifica nei più tradizionali studi qualitativi.

Ma le grandi raccolte di testi digitalizzati ci danno immediato accesso all'insieme globale di volumi pubblicati in italiano, o alla sua gran parte, con la possibilità di attingere anche alla letteratura non canonica (quella letteratura di con-

sumo dal limitato valore artistico, ma che ha avuto una circolazione di estrema ampiezza tra i lettori), alla letteratura tecnica e scientifica, alle traduzioni. In questi mesi sto sperimentando quanto questa disponibilità di testi ci permetta di meglio datare il lessico degli ultimi due secoli. La lezione che traggio da questi spogli (che si basano su due presupposti di fondo: l'attingibilità di raccolte vaste di testi finora non studiati e l'esistenza di appropriati strumenti di estrazione dei dati) è che questo enorme giacimento di testi digitalizzati ci permette di studiare in maniera più approfondita, e con una grana più fine, la lingua almeno degli ultimi due secoli. Anzi, questa nuova disponibilità di risorse ci impone di farlo. Ma per utilizzare appieno il patrimonio disponibile è necessario, a mio avviso, mutare la prospettiva con cui si guarda ai *corpora*.

La consuetudine di agire su *corpora* ampi, ma ristretti rispetto alla realtà di riferimento, scelti in base a criteri di rappresentatività (e quindi in base a un canone generale o a un canone definito *ad hoc*) e annotati è scarsamente compatibile con l'attuale ampia disponibilità di testi digitalizzati. È anche scarsamente coerente con gli sviluppi del trattamento automatico dei dati. Certamente le raccolte sempre più estese di testi non possono essere definite in senso stretto *big data*, in quanto consistono sì in un volume di dati fino ad oggi impensabile e sono caratterizzate anche da grande varietà, ma sono dati statici, che non variano con la velocità e l'imprevedibilità tipica dei *big data*. Ma la tendenza che si sta imponendo di elaborare grandi quantità di dati non può non estendersi, con le dovute finanze metodologiche, anche allo studio della storia della lingua, quando viene sviluppato non attraverso l'individuale capacità di identificazione delle dinamiche culturali che stanno alla base dei movimenti della lingua, ma attraverso lo spoglio o lo studio quantitativo di *corpora*.

### 3. STRUMENTI PER UNA VISIONE DA DISTANTE DELLA STORIA DELLA LINGUA ITALIANA

Ci troviamo, quindi, davanti a uno scenario nuovo, che ci permette di indirizzarci verso la possibilità di confrontarci con un numero di testi finora impensabile e di farlo nell'ottica del *distant reading* (che si oppone alla classica prospettiva del *close reading*), cioè attraverso quell'analisi indiretta dei testi che avviene attraverso elaborazioni di vario genere (per es. grafici, carte, alberi). Un'ottica che "fa vedere meno dettagli, vero: ma fa capire meglio i rapporti, i *pattern*, le forme" (Moretti 2005: 3) ed è certamente l'unica possibile se vogliamo utilizzare *corpora* di grandi dimensioni. Ma, in una sorta di circolo vizioso, la disponibilità di ampie raccolte di testi non giustifica più l'utilizzo di campioni ridotti, di insieme di testi rappresentativi, scelti sulla base di canoni più o meno condivisi. Questo è particolarmente vero quando l'obiettivo è quello di individuare andamenti diacronici su ampi archi cronologici, per i quali non sono sufficienti testimonianze uniche per ogni porzione temporale (come abbiamo visto accadere un po' in tutti i *corpora* ora disponibili); ed è ancor più valido in relazione a fasi storiche che hanno



visto verificarsi rilevanti processi di evoluzione linguistica (preciso che, anche in coerenza con gli altri interventi pubblicati in questa sede, penso principalmente, all'Otto-Novecento).

Si tratta, allora, di trovare gli strumenti che ci permettano di procedere a una lettura da distante delle ampie raccolte di testi digitalizzati ora disponibili, di individuare modelli e tendenze, in un'ottica che per uno storico della lingua è principalmente esplorativa e cioè tale da suggerire ipotesi di lavoro da verificare poi con analisi di tipo più tradizionale, basate su spogli da vicino dei testi appartenenti alle raccolte disponibili.

A questo fine possiamo affiancare alla ormai consolidata tradizione di costituire *corpora* di testi forniti di annotazioni, più o meno elaborate, una metodologia di analisi quantitativa di dati testuali basati su *corpora* testuali se non grezzi, certamente debolmente normalizzati, ma di dimensioni ben più ampie di quelle tradizionali (ciò significa anche, nel campo della ricostruzione dello sviluppo della lingua, caratterizzati da una rappresentatività dell'evoluzione cronologica molto più fine).

Insomma, se abbiamo la possibilità (e forse anche il dovere) di utilizzare ampie raccolte di testi, siamo costretti a ridurre al più basso livello possibile l'elaborazione finalizzata alla costruzione del dato. I testi oggetto di analisi vengono, sostanzialmente, a corrispondere a testi digitalizzati in forma basica, quindi in semplice formato di testo; i testi vanno certamente ripuliti degli errori che qualsiasi sistema di lettura e digitalizzazione automatica comporta e di eventuali tratti "sporchi" che possono venire introdotti nella transcodificazione da un formato all'altro (tipicamente da pdf a txt); inoltre vanno normalizzate le consuetudini grafiche (a meno che non siano esse stesse oggetto di analisi), per esempio unificando la scelta del tipo di segnacento nelle parole tronche (è noto, ad esempio, che Einaudi utilizza l'accento acuto per tutte le vocali chiuse, comprese *i* ed *u*; pur riconoscendo che, alla base di questa scelta, c'è una motivazione non priva di buone ragioni, mantenerla nei testi significherebbe rilevare a livello quantitativo e attribuire quindi ai testi pubblicati da un editore una diversità che si situa solamente sul piano delle scelte grafiche, ma che finirebbe con l'allontanare tali testi da testi nella sostanza molto simili).

La diminuzione di informazione rispetto a testi annotati deve essere, però, compensata dalla raffinatezza degli strumenti di analisi quantitativa. Da statistici o matematici lo storico della lingua, capace di offrire temi di indagine e insiemi di testi validi per affrontare tali temi, deve aspettarsi di ricevere strumenti di misurazione e di classificazione che permettano quella visione da distante che è tanto più necessaria quanto più ampio è l'insieme dei testi sottoponibile ad analisi.

Dalla statistica ci possono giungere strumenti raffinati per riconoscere e raffigurare l'andamento temporale di forme o lemmi presenti in *corpora* diacronici, cioè in *corpora* di testi capaci di rappresentare l'evoluzione storica di un tipo testuale (Trevisani e Tuzzi 2013) o di uno specifico dominio nozionale, come, ad

esempio, la letteratura specialistica relativa a un preciso ambito scientifico (Tuzzi 2018, Trevisani e Tuzzi 2018, Trevisani e Tuzzi 2015).

In questo numero della rivista *Sciandra*, Trevisani e Tuzzi hanno messo alla prova gli strumenti statistici su un corpus di narrativa italiana; ma l'obiettivo più ambizioso è quello di poter tratteggiare l'evoluzione dell'uso lessicale almeno scritto di un'intera comunità linguistica, guardando da distante una quantità rilevante di testi di varia natura, capaci di rappresentare con ricchezza di dati ciascuna porzione (per esempio annuale) dell'arco temporale indagato. Da un'altra prospettiva, con tecniche ascrivibili al *distant reading* e basate sul *machine learning*, si è cercato di valutare l'efficacia delle periodizzazioni proposte per la rappresentazione della storia della lingua italiana (Cortelazzo et alii in print).

L'utilizzo di *corpora* di ampie dimensioni, costituiti dal maggior numero di testi possibili, indipendentemente da una scelta legati a canoni letterari (per quanto larghi), capaci di collocarsi in una fitta linea del tempo, implica un vero lavoro interdisciplinare tra studiosi di lingua italiana ed esperti di metodi quantitativi di analisi. Ai primi spetta, comunque, il recupero dei testi e la loro pulizia, l'individuazione dei dati più significativi estraibili (almeno potenzialmente) da quei testi, l'interpretazione sul piano della storia culturale dei risultati che emergono dall'elaborazione dei dati; ai secondi di individuare i modi più raffinati, e se possibile più innovativi, per rappresentare l'evoluzione nel tempo dei fenomeni rappresentati dai dati oggetto di analisi (meglio se permettono di individuare tendenze non immediatamente percepibili con la lettura da vicino).

Mi è già capitato (Tuzzi, Cortelazzo 2018) di utilizzare un adattamento di una frase di Baricco (1994) per indicare di cosa è necessario disporre quando ci si appresta a svolgere indagini testuali con metodi quantitativi. I ricercatori orientati verso una considerazione quantitativa dei testi possono dirsi: "non sei fregato veramente finché hai da parte un buon corpus e un metodo con cui analizzarlo". Ciò mi appare particolarmente vero quando si vuole operare in chiave storica, su una grande quantità di testi (e allora occorre ricorrere a metodi particolarmente raffinati); oppure quando si hanno disposizione metodi particolarmente raffinati (e allora servono *corpora* ampi, che permettano di rappresentare nella maniera più densa e fine l'intero arco cronologico oggetto di analisi).

- Baricco A. (1994) *Novecento. Un monologo*, Milano, Feltrinelli.
- Cortelazzo M. A. (2007) “La perifrasi progressiva in italiano è un anglicismo sintattico?”, in *Studi in onore di Pier Vincenzo Mengaldo per i suoi settant'anni, a cura degli allievi padovani*, Firenze, SISMELE. Edizioni del Galluzzo, pp. 1753-1764.
- Cortelazzo M. A., Gatti F. M. T., Mikros G. K. e Tuzzi A. (in stampa) “Does the Century matter? Machine learning methods to attribute historical periods to an Italian literary corpus”, in *Quantitative Approaches to Universality and Individuality in Language. Selected papers of the 11th International Conference on Quantitative Linguistics (QUALICO)*. Ed. di M. Yamazaki et alii, Berlin, De Gruyter Mouton.
- LIZ 1993 = *Letteratura italiana Zanichelli (= LIZ)*, CD-ROM dei testi della letteratura italiana. A cura di P. Stoppelli e E. Picchi, Bologna, Zanichelli.
- LIZ 1995 = *Letteratura italiana Zanichelli (= LIZ)*, CD-ROM dei testi della letteratura italiana. A cura di P. Stoppelli e E. Picchi, Bologna, Zanichelli.
- LIZ 1997 = *Letteratura italiana Zanichelli (= LIZ)*, CD-ROM dei testi della letteratura italiana. A cura di P. Stoppelli e E. Picchi, Bologna, Zanichelli.
- LIZ 2001 = *Letteratura italiana Zanichelli (= LIZ)*, CD-ROM dei testi della letteratura italiana. A cura di P. Stoppelli e E. Picchi, Bologna, Zanichelli.
- Moretti F. (2005) *La letteratura vista da lontano*, Torino, Einaudi.
- Savoy J. (2018) “Is Starnone really the author behind Ferrante?”, *Digital Scholarship in the Humanities*, 33(4), pp. 902-918.
- Schöch C., Patras R., Erjavec T. e Santos D. (2021) “Creating the European Literary Text Collection (ELTeC). Challenges and Perspectives”, *Modern Languages Open*, (1), p. 25.
- Serianni L. (2006) “Gli archivi elettronici e la lessicografia storica”, in *Nuovi media e lessicografia storica. Atti del colloquio in occasione del settantesimo compleanno di Max Pfister*. A cura di W. Schweickard, Tübingen, Niemeyer, pp. 41-58.
- Trevisani M. e Tuzzi A. (2013) “Shaping the history of words”, in *Methods and Applications of Quantitative Linguistics. Selected papers of the VIIIth International Conference on Quantitative Linguistics (QUALICO)*. Ed. by I. Obradović, E. Kelić & R. Köhler, Belgrade, Akademska Misao, pp. 84-95.
- Trevisani M. e Tuzzi A. (2015) “A portrait of JASA: the History of Statistics through analysis of keyword counts in an early scientific journal”, *Quality and Quantity*, 49, pp. 1287-1304.
- Trevisani M. e Tuzzi A. (2018) “Learning the evolution of disciplines from scientific literature. A functional clustering approach to normalized keyword count trajectories”, *Knowledge-based systems*, 146, pp. 129-141.
- Tuzzi A. (2018) (ed.) *Tracing the Life Cycle of Ideas in the Humanities and Social Sciences*, Cham, Springer.
- Tuzzi A. e Cortelazzo M. A. (2018) “What is Elena Ferrante? A comparative analysis of a secretive bestselling Italian writer”, *Digital Scholarship in the Humanities*, 33(3), pp. 685-702.

# Tempi verbali e perifrasi gerundivali in un corpus di italiano letterario (1800-2000)

STEFANO ONDELLI  
Università di Trieste  
[sondelli@units.it](mailto:sondelli@units.it)

## ABSTRACT

The Italian gerundial structures have undergone opposite developments in the last century. The continuous periphrasis [*andare* or *venire* + gerund] was more frequent at the end of the 1800s, whereas the progressive structure [*stare* + gerund] increased sharply during the 1900s. Furthermore, some of its syntactic constraints have weakened and today it can include a broader range of verb tenses and actional values. Based on previous research, both qualitative and quantitative, this study analyses a corpus including novels written by Italian authors and translations from foreign languages into Italian dating from 1811-1896 and 1976-2005. In addition to the overall figures, which confirm the dominant role of the continuous periphrases in literary language at the end of the 1900s, data suggest that translators anticipated the development of the gerundial periphrases in Italian in terms of the selection of auxiliary verbs and verb tenses.

## KEYWORDS

Corpus linguistics, verbal periphrases, translation studies, continuous periphrasis, progressive periphrasis

## 1. INTRODUZIONE E QUESITO DI RICERCA

Questo articolo rende conto di alcuni aspetti dell'impiego delle perifrasi gerundivali continua e progressiva in un corpus di italiano letterario che copre il periodo 1811-2005. Il contributo si articola come segue: il § 2. fornisce il background teorico su cui si basa l'analisi e il corpus da cui sono stati estratti i dati testuali; il § 3. illustra la distribuzione dei tempi verbali e tratteggia alcune caratteristiche dei lemmi coinvolti nei costrutti qui considerati, paragonando i dati quantitativi sia in diacronia che nel confronto con opere di romanzieri italiani e traduzioni da diverse lingue straniere; il § 4 trae le conclusioni dello studio e delinea alcune prospettive di ricerca futura.

Questo contributo rende conto della seconda fase di una ricerca ancora in corso (cfr Ondelli in stampa) che ha l'obiettivo di confermare o smentire alcune delle conclusioni a cui sono giunte ricerche analoghe già svolte in passato, aggiungendo l'elemento di novità rappresentato dalla presenza, nel corpus di analisi, di testi tradotti da altre lingue. L'ipotesi è che le traduzioni possano avere svolto un ruolo di rinforzo dell'incremento della frequenza della perifrasi progressiva e abbiano anticipato l'indebolimento della perifrasi continua nell'italiano letterario e paraletterario.

## 2. PREMESSE TEORICHE, MATERIALI E METODI

Gli studi imprescindibili per la definizione delle perifrasi gerundivali continua (d'ora in avanti: CONT) e progressiva (d'ora in avanti: PROG) in italiano sono comprendono principalmente quelli condotti da Pier Marco Bertinetto (1986, 1990, 1991, 1997, 1998, 2000 e 2003), anche in collaborazione con altri ricercatori in ottica comparatistica (Bertinetto et al. 2000). A questi vanno aggiunte le osservazioni puntuali relative allo sviluppo in diacronia di Giovanna Brianti (2000 e Titus-Brianti 2010) e Mario Squartini (1990), oltre ai dati quantitativi estratti dal LIP (De Mauro et al. 1993) da Andrea Villarini (1999) e all'esplicito confronto tra la PROG italiana e la perifrasi [*to be* + *-ing*] inglese, anche in traduzione, condotto da Chiara Degano (2005). Un quadro riassuntivo, che formula l'ipotesi dell'origine endogena della crescente fortuna della PROG, a prescindere dal ruolo di rinforzo svolto dall'influsso inglese, è stato invece offerto da Michele Cortelazzo (2007).

Non è questo il luogo per riproporre una descrizione esaustiva delle due perifrasi gerundivali qui considerate. Ci limiteremo, piuttosto, a ricordare alcune osservazioni, svolte da chi ci ha preceduti, su cui torneremo in sede d'analisi dei risultati dello spoglio del corpus.

Iniziamo dalla questione dell'organicità sintattica delle perifrasi verbali in genere, evidenziata da Bertinetto (1990: 344), secondo la quale *sta studiando* si presenta, evidentemente, come un costrutto dotato di maggior perifrasticità rispetto a una frase come *sta in camera, studiando accanitamente*, in cui il verbo

modificatore mantiene il proprio significato originario di stato, e la gerundiva si presenta come una subordinata che indica simultaneità. Successivamente Bertinetto (2003: 113 e segg.) si riferisce a queste strutture con l'etichetta di *costrutti pseudoperifrastici*, che risultano significativi non solo per apprezzare il carattere localistico della PROG (*mutatis mutandis*, il discorso si applica anche alla CONT), ma anche per evidenziare le possibili difficoltà insiste nelle procedure di individuazione automatica delle perifrasi gerundivali: la ricerca delle co-occorrenze di gerundio e verbi modificatori negli stessi contesti conduce, inevitabilmente, a un ampio numero di falsi positivi.

Tra le restrizioni d'uso della PROG, Bertinetto (1991: 132 e segg.) ricorda lo scarso impiego dell'infinito e, di conseguenza, dei verbi modali, oltre all'esclusione, nell'italiano contemporaneo, dei tempi perfettivi, ammessi invece nell'italiano antico fino all'Ottocento, seppure con un numero ridotto di forme (*aspettare, parlare e pensare*) al perfetto semplice. Inoltre, risulta rara la presenza della PROG in entrambi i membri di costrutti correlativi e temporali indicanti simultaneità. In particolare, sulla scorta delle osservazioni di Brianti (2000), Bertinetto (2003: 102, nota 16) afferma che fino a Cinquecento la PROG

è stata unicamente impiegata nello schema sintattico 'di contemporaneità' (una forma del quale è, per esempio: *mentre X stava lavorando, Y dormiva*) piuttosto che 'di incidenza' (ad esempio: *quando X arrivò, Y stava dormendo*). A partire dal Seicento compare invece lo schema 'di incidenza', che cresce poi fino a divenire il modulo sintattico esclusivo tra la fine dell'Ottocento e l'inizio del Novecento.

Per quanto riguarda, invece, la CONT, va ricordato che questo costrutto, a differenza della PROG, non prevede un istante di focalizzazione (d'ora in avanti: Tf) ed è facilmente accostabile ad avverbiali di durata. Le due perifrasi risultano però intercambiabili nei casi in cui esprimano un evento che si svolge gradualmente, tendendo a un risultato finale (per una descrizione esaustiva, cfr. Bertinetto 1990: 138 e segg.). Questa "telicità tendenziale" (Bertinetto 1998: 125) è propria della CONT e può essere accompagnata da avverbiali che indicano gradualità (per es., *man mano, a poco a poco*) e incremento (per es., *sempre più*) o espressioni di maniera e verbi che indicano reiterazione, suscettibili di sottolineare la processualità dell'evento ed esprimere l'accostamento al *telos*.

In particolare, gli spogli condotti da Bertinetto (2003: 105) sulla prosa letteraria indicano che, tra i tipi verbali più frequentemente impiegati nella CONT, un ruolo di primo piano è svolto dai *verba dicendi* almeno fino alla metà del Novecento. Il dato dimostrerebbe che, subito prima e in corrispondenza della forte contrazione quantitativa, il costrutto ha subito una stereotipizzazione che lo ha limitato a formule ripetitive di introduzione al discorso riportato, del tipo *andava dicendo o veniva domandando*.

Il corpus utilizzato per questa ricerca è lo stesso preso precedentemente in considerazione per una serie di spogli relativi ai verbi procomplementari (Ondelli 2019a, 2019b e 2020), e comprende romanzi scritti da autori italiani (d'ora in avanti: testi nativi) e tradotti da diverse lingue straniere tra il 1811 e il 2005 (cfr. l'appendice a questo articolo). Per il momento, l'identificazione delle perifrasi gerundivali è stata condotta sui subcorpora relativi a due momenti storici (1811-1896 e 1976-2005) per cogliere le differenze tra gli esordi e la definitiva stabilizzazione del predominio della PROG sulla CONT, mentre si è rimandato a una fase successiva lo spoglio del periodo intermedio.

Rispetto alle ricerche svolte in precedenza da altri studiosi (con l'eccezione di Degano 2005), la novità nei materiali riguarda principalmente la presenza opere paraletterarie, accanto a esempi di letteratura alta, e soprattutto di traduzioni di romanzi stranieri. Tale bilanciamento è volto a valutare il contributo che i traduttori possono aver dato al successo della PROG nel passaggio tra XIX e XX secolo, in particolare nel caso di opere destinate a una fruizione di consumo, quindi (presumibilmente) caratterizzate da una minore attenzione alla qualità linguistica.

L'estrazione delle perifrasi è avvenuta manualmente, dopo aver individuato in maniera automatica tutte le concordanze delle forme gerundivali presenti nel corpus. Poiché i dati quantitativi sono già stati resi noti (cfr. Ondelli in stampa), in questa occasione ci si soffermerà sulla distribuzione dei tempi verbali nelle perifrasi in esame, con qualche annotazione relativa ai lemmi più frequenti o che risultano significativi alla luce delle ricerche precedenti. Poiché la valutazione degli aspetti azionali richiede l'analisi qualitativa dei contesti d'uso, in considerazione della mole dei materiali estratti, in particolare nel secondo dei due periodi considerati, se ne rimanda la classificazione a una fase successiva della ricerca.

### 3. TEMPI VERBALI E PERIFRASI GERUNDIVALI

Questa sezione illustra l'impiego dei tempi verbali a cui sono coniugati i verbi modificatori delle perifrasi gerundivali nel corpus esaminato, oltre a soffermarsi su alcuni fatti linguistici (per es., certi tipi verbali o costrutti pseudoperifrastici) che sono stati evidenziati in precedenza dagli studiosi che si sono occupati dell'argomento. In particolare, i §§ 3.1. e 3.2. tratteggiano l'evoluzione in diacronia delle due perifrasi che ci interessano nei romanzi scritti da autori italiani, mentre dal §3.3 al §3.6 si opera un confronto tra testi nativi e traduzioni.

Per comodità di esposizione, riportiamo qui le abbreviazioni che saranno utilizzate nel testo:

PRE – indicativo presente  
 IPF – indicativo imperfetto  
 PFS – indicativo perfetto semplice  
 PFC – indicativo perfetto composto  
 PPF – indicativo piuccheperfetto  
 FUT – indicativo futuro  
 COND – condizionale semplice  
 CONG PRE – congiuntivo presente  
 CONG IPF – congiuntivo imperfetto  
 CONG P – congiuntivo composto  
 INF – infinito

### 3.1. CONT: EVOLUZIONE IN DIACRONIA NEI TESTI NATIVI

Nativi '800				totale%	andare%	venire%	Nativi '900				totale%	andare%	venire%
	andare	venire	totale					andare	venire	totale			
PRE	18	2	20	13,15789	16,36364	4,761905	PRE	1	0	1	0,60241	0,606061	0
IPF	79	24	103	67,76316	71,81818	57,14286	IPF	138	0	138	83,13253	83,63636	0
PFS	7	8	15	9,868421	6,363636	19,04762	PFS	0	0	0	0	0	0
PFC	0	1	1	0,657895	0	2,380952	PFC	0	0	0	0	0	0
PPF	1	6	7	4,605263	0,909091	14,28571	PPF	23	1	24	14,45783	13,93939	100
FUT	2	0	2	1,315789	1,818182	0	FUT	0	0	0	0	0	0
COND	0	0	0	0	0	0	COND	1	0	1	0,60241	0,606061	0
CONG PRE	0	0	0	0	0	0	CONG PRE	0	0	0	0	0	0
CONG IPF	3	1	4	2,631579	2,727273	2,380952	CONG IPF	1	0	1	0,60241	0,606061	0
CONG P	0	0	0	0	0	0	CONG P	1	0	1	0,60241	0,606061	0
INF	0	0	0	0	0	0	INF	0	0	0	0	0	0
Totale	110	42	152	100	100	100	Totale	165	1	166	100	100	100

Tabella 1 – CONT nei testi nativi nell'Ottocento e nel Novecento. Frequenze assolute e %

#### 3.1.1. CONT CON VENIRE NEI TESTI NATIVI IN DIACRONIA

Come si può vedere nella tabella sopra, tra l'800 e la fine del '900, nei testi nativi la perifrasi continua con *venire* in pratica scompare. L'unico caso emerge in un contesto di evidente innalzamento del registro, in compresenza con un'al-



tra CONT, stavolta costruita con *andare*, che, in un quadro di imperfettività, telicizza gli eventi descritti:

ma intuiva che erano anche cambiati da quando li aveva conosciuti, e che avevano cominciato a vedere le cose con maggiore chiarezza [quadro perfettivo], perché i fatti che accadevano [quadro imperfettivo] *andavano aggiustando* il fuoco della lente davanti ai loro occhi, e le loro immagini *si erano venute accomodando* [lettura incrementale che sfocia nel quadro perfettivo] (C. Sgorlon, *L'armata dei fiumi perduti*)

Nel dettaglio, nel corso dell'800 la perifrasi continua con *venire* presenta una minore specializzazione di quella costruita con *andare* (come avremo modo di vedere al § 3.1.2. qui sotto) nei verbi di parola (per es., *strillare*), mentre registra lo stesso *pattern* per i continuativi che indicano un cambiamento di stato (pronominali o no; per es., *crescere* o *inclinarsi*). Tale mancanza di specializzazione è dimostrata anche dal fatto che, tra i verbi principali della costruzione perifrastica, nessun lemma supera le 2 occorrenze.

Notevole, perché numericamente superiore a quella di *andare* nel XIX secolo, la coniugazione di *venire* ai tempi perfettivi: il PPF (per es., *era venuto interrompendo l'invettiva di Maironi con proteste sempre più violente*), il PFS (per es., *o poveraglia! schifosa e compassionevole piaga sociale! gli venne dicendo il Banchiere sfoderando una tirata filosofica*) e anche un caso di PFC (sebbene dubbio: è possibile che si tratti di un verbo di movimento: *il povero Vigilante s'ebbe un cattivo colpo di bastone dal castaldo, ed è venuto strillando il suo guaio*).

Si conferma dunque la facilità di combinazione della CONT con la lettura perfettiva, anche se restano numericamente dominanti i tempi imperfettivi (IPF, PRE e CONG IPF). Gli avverbiali che indicano gradualità si presentano sia in contesti perfettivi (per es., *la padroncina si venne calmando a poco a poco*) che imperfettivi (per es., *ed anche lo confortava il sentire che l'aria si veniva facendo sempre più calda*); questi ultimi compaiono indifferentemente in periodi complessi in cui una subordinata temporale può indicare simultaneità (per es., *mentre così la fanciulla cominciava a smarrirsi in quelle sue solite fantasticherie, un giovine le si veniva appressando da un capo della strada*) oppure realizzare lo schema incidentale che richiede un Tf (per es., *e costui mi veniva manifestando certi suoi disegni sul futuro governo del paese, quando una masnada di selvaggi ne piombò addosso menando giù con certe loro mazze percosse da confessione*).

### 3.1.2. CONT CON *andare* NEI TESTI NATIVI IN DIACRONIA

Come già nella CONT costruita con *venire*, tra i verbi principali della perifrasi con *andare* nessun lemma supera le due occorrenze. In generale, nell'800 dominano i verbi di pensiero (per es. *notare* e *pensare*) e, soprattutto, di parola, a sottolineare l'aspetto continuo (*bofonchiare*, *brontolare*, *contare* nel senso di 'raccontare', *dimandare*, *dire*, *esporre*, *interrogare*, *maledire*, *predicare*, *rispondere soffiare* – ben 2 occorrenze, nel senso di *dire* –, *suggerire* e *sussurrare*,). Si registra anche un buon numero di

verbi la cui semantica implica reiterazione (anche grazie alla prefissazione; per es., *girare e rigirare*) o la ricostituzione di uno stato precedente (e in questo caso è evidente la telicità del processo, come in *riacquistare, ripigliare, riprendere coscienza*). Inoltre la presenza del *telos* emerge nei trasformativi, siano essi pronominali (*accendersi, aggravarsi, allontanarsi, avvivarsi, cangiarsi, chinarsi, condensarsi, confondersi, consumarsi, diradersi, mutarsi, nascondersi, persuadersi, prepararsi, raffreddarsi, restringersi, risvegliarsi, schiarirsi, velarsi*) o no (*calmare, cangiare, crescere, diventare buio, ingrossare, maturare, morire, scomparire*).

Tra i casi notevoli nel subcorpus relativo all'800 si segnalano due occorrenze di FUT, rispettivamente con uso futurale (*la vita è un fiume che dopo un uragano ha le acque torbide; ma lasciate passare dell'acqua, e a poco a poco il fiume andrà schiarendosi*) e uno che, in realtà, si presenta come un INF retto dal modale *dovere*, in cui emerge una certa carica epistemica: *senza madre, povero bambino, lui che m'amava tanto, che aveva tanto bisogno di me, senza madre, nella miseria, dovrà andare accattando, lui, Marco, Marco mio, che tenderà la mano, affamato!*

Nel confronto con l'800, nell'ultimo quarto del XX secolo e all'inizio del XXI, la CONT costruita con *andare*, pur mantenendo una certa consistenza numerica, vede la scomparsa sia del FUT che del PFS, che nel primo dei due periodi storici qui considerati reggeva verbi principali come *avvivarsi, crescere* (2 occorrenze), *mordere, rispondere, sfiorare* e *smorzare*. Si conferma, inoltre, il dominio incontrastato dell'IPF, e anche il PPF registra un notevole incremento, probabile sintomo di una stereotipizzazione del costruito. Paiono meno frequenti i *verba dicendi* (per es., *chiedere, domandare, sciorinare, sussurrare*), mentre permangono i cambiamenti di stato (per es., *maturare*), i verbi pronominali (per es., *sgonfiarsi*) e quelli indicanti il ripristino di una condizione precedente (per es., *riacquistare* e *rimarginare*).

In entrambi i subcorpora considerati si conferma la tendenza della CONT ad accompagnarsi ad avverbiali che indicano il graduale avvicinamento a un *telos*. Riportiamo alcuni esempi tratti dal corpus che comprendono entrambi i verbi modificatori:

il monaco, *a poco a poco*, s'era venuto placando, e aveva detto le ultime parole quasi col tono di ogni altro cristiano

guadagnando ottanta lire al mese e non spendendo nulla per sé, mandava a casa ogni tre mesi una bella somma, con la quale il marito, che era galantuomo, andava pagando *via via* i debiti più urgenti, e riguadagnando così la sua buona reputazione

*man mano* che rivedeva le case, le botteghe, la gente, i soliti amici, andava ricuperando anche il senso della sua vita solita

Si pose a leggere e in capo alla seconda pagina, udendo il respiro della padrona farsi greve, andò *pian piano* smorzando la voce per un mormorio inarticolato, fino al silenzio

In alternativa, la lettura incrementale può essere affidata alla maggiorazione di gruppi aggettivali:

il concetto religioso che mi si veniva formando *sempre più chiaro* nella mente era questo, in breve: Dio esiste, è anche potente, è anche sapiente, tutto come credi tu; ma che noi lo adoriamo e gli parliamo non gliene importa nulla

così la signora Cecca andò rispondendo al Pasotti *sempre più breve e cauta*, e a quel ma, posto delicatamente dove le doleva, scattò

ed anche lo confortava il sentire che l'aria si veniva facendo *sempre più calda*

Non sono escluse altre risorse, come frasi secondarie in funzione causale o temporale:

perché dunque si sentiva preso da malessere, osservando che, *col decrescere della malattia di Eugenia*, il carattere di lei veniva appunto conformandosi all'idea che egli s'era fatta di un'inalterabile felicità domestica, di una esistenza isolata e quasi fuori del mondo?

In altri casi l'idea dell'avvicinamento a un risultato finale è insita nella semantica del verbo e agli avverbiali è affidata la funzione di veicolare la struttura di un processo continuo che si sviluppa per gradi, tramite la ripetizione di fasi in successione:

ma il malato s'andava *sempre* aggravando

*in ogni sua lettera* il giovane le diceva che il tempo della domanda si veniva *sempre* approssimando, che fra un anno il loro voto si sarebbe compiuto

malgrado che peggiorasse, a lui sembrava che il malato andasse riacquistando *lentamente* un poco d'intelligenza

In assenza del *telos*, l'aspetto continuo può essere sottolineato da avverbi o da espressioni che specificano il modo in cui l'evento ha luogo:

dicono che sia impossibile esprimere l'amore ch'ella ha preso alla Camilla, e le carezze ch'ella va prodigandole *continuamente* fanno piangere di tenerezza

egli andava errando *senza mèta*, come in un fantastico labirinto

Luisa tremava a capo chino, non gli rispondeva in alcun modo ed egli si frenò, le prese il capo come una cosa santa, le andò sfiorando con le labbra, *qua, là*, i capelli bianchi che vedeva

il giorno che poterono lasciar l'albergo, era stata una festa per tutti e due percorrere a braccetto quei lunghi corridoi, visitare le celle che il Padreterno veniva aprendo *una dietro l'altra*, dando rapidi schiarimenti

Entrambi i subcorpora presentano quelle che Bertinetto (2003: 113) definisce “costrutti pseudoperifrastici”: si tratta di strutture in cui *andare* e *venire* mantengono il loro significato di verbi di movimento, mentre alla secondaria gerundiva che segue spetta il compito di esprimere il modo in cui avviene il movimento o un evento ad esso simultaneo. Si tratta, come ovvio, di costruzioni che producono un ampio numero di falsi positivi a seguito dell'estrazione automatica dei dati testuali. Un caso emblematico è quello del proverbio che segue:

*ma l'appetito vien mangiando*, dice il proverbio, e don Blasco non si contentava di quelle poche migliaia d'onze, voleva arricchire per davvero, studiava il modo di batter moneta

Altrove, il GER segue il verbo di movimento, con o senza una virgola di separazione:

e quella spesso si offendeva, *andava via* portandosi il lavoro, come a dire che preferiva starsene a lavorare sola in un'altra camera, e a volte sbatteva perfino l'uscio

e *se ne andò* con gli occhi scintillanti, *fregandosi le mani* nell'aspettazione di una piacevole caccia

Nell'ultimo esempio qui sopra, il fatto che si tratti di un movimento compiuto dal soggetto è reso evidente dalla natura procomplementare del verbo (*andarsene*). Un fenomeno analogo è molto ricorrente nelle strutture in cui a essere accompagnato dal GER è il verbo di stato *starsene*:

Eugenia non levava gli occhi d'addosso a Patrizio che *se ne stava in disparte*, guardando ora la facciata della chiesetta, ora gli ulivi attorno, ora il cielo che cominciava a imbrunire

In diversi altri casi, il verbo di stato è variamente specificato da locativi o altre espressioni con valore modale, con la frase gerundiva che può seguire direttamente la principale o esserne separata da una virgola:

Luisa, che *stava in cucina* sbucciando delle castagne, lo udì passare pel corridoio, *stette un momento in forse*, lottando con se stessa, poi balzò fuori, lo raggiunse mentre stava per scender le scale

il duca, che *stava in un angolo* con Benedetto spiegandogli le proprie idee sull'avviamento della Banca Meridionale, che sotto la direzione di don Lorenzo Giulente doveva «venire all'aiuto dell'incremento industriale e commerciale» e «cooperare l'opera protettrice del governo», sorrise impercettibilmente, scrollando le spalle, alla sfuriata del fratello

sua figlia, donna Carolina, *stava in piedi presso la finestra* discorrendo con una favorita della marchesa, nipote del suo fattore

### 3.2. PROG: EVOLUZIONE IN DIACRONIA NEI TESTI NATIVI

Nativi '800			Nativi '900		
		%			%
PRE	14	19,71831	PRE	819	33,92709
IPF	50	70,42254	IPF	1406	58,24358
PFS	3	4,225352	PFS	0	0
PFC	1	1,408451	PFC	0	0
PPF	0	0	PPF	0	0
FUT	0	0	FUT	19	0,787075
COND	0	0	COND	0	0
CONG PRE	0	0	CONG PRE	32	1,325601
CONG IPF	3	4,225352	CONG IPF	136	5,633803
CONG P	0	0	CONG P	0	0
INF	0	0	INF	2	0,08285
Totale	71	100	Totale	2414	100

Tabella 2 – PROG nei testi nativi nell'Ottocento e nel Novecento. Frequenze assolute e %

Nel periodo 1860-1899 la PROG risulta meno frequente della CONT, anche se *stare*, come verbo modificatore, con 70 occorrenze assume una posizione numericamente intermedia tra *andare* (100 occorrenze) e *venire* (42 occorrenze). In effetti troviamo casi di compresenza delle tre soluzioni nello stesso contesto linguistico: nell'esempio qui sotto, la PROG indica la simultaneità di un evento espresso da un verbo non durativo (*mettere un fazzoletto al collo*) e due durativi (*aspettare e brontolare*), di cui il secondo è verbo di parola. Pur rimanendo valida, l'affermazione di Bertinetto (2003: 102), secondo la quale lo schema di incidenza resta l'unica opzione nel passaggio tra XIX e XX secolo, va dunque circostanziata. Occorre comunque dire che il corpus non ha prodotto alcun contesto di simultaneità espresso dalla compresenza di PROG (del tipo: *mentre X stava parlando, Y stava ridendo*):

nel minuscolo salottino assediato dall'inverno Luisa *stava mettendo*, ginocchioni, un fazzoletto al collo di Maria, Franco *aspettava* col cappuccio di sua moglie in mano e la Cia, la vecchia governante, col cappello in testa e le mani nel manicotto, *andava brontolando* al suo padrone: «Che signore è mai Lei! Cosa vuol fare qui solo a casa?»

La PROG risulta poco compatibile con i tempi perfettivi, e il PFC rappresenta l'unico caso di tempo composto:

Figurati con che ansia *sono stato aspettando* il suo ritorno! con che trasporto di gioia l'ho veduto arrivare!

Raro risulta anche il PFS (sempre con i verbi durativi: *ascoltare, aspettare, guardare*, quando Bertinetto [1990: 132], nel notare la progressiva stereotipizzazione di questo tempo verbale, menziona *aspettare, parlare e pensare*):

rimasta sola, la signora Rigey *stette ascoltando* il rumor dei passi che si allontanavano

allora, piangendo, il ragazzo prese una seggiola, sedette e *stette aspettando*, senza levar gli occhi dal viso di suo padre

dirimpetto a lui un maturo Dandy *stette guardandolo* con qualche meraviglia, e poi alzandosi e stringendoglisi al collo

*Aspettare*, con 6 occorrenze (5 all'IPF e 1 al PFS), insieme *scrivere* (4 occorrenze: 3 all'IPF e 1 al PRE) sono i lemmi più usati; nessun altro supera le due occorrenze. Per il resto non sembra esserci un *pattern* preciso nei verbi coniugati all'IPF: uno spoglio cursorio ne individua di tutti i tipi azionali, anche se le fasi successive di questa ricerca, con controlli più precisi sui contesti, (in particolare nel confronto con il subcorpus del '900, caratterizzato da numeri ben maggiori) potrebbero individuare tendenze di fondo al momento invisibili. La CONT coinvolge anche verbi non durativi (per es., *impiccare*) e trasformativi (per es., *colare a picco* e *morire*), ma un solo pronominale (*fortificarsi*), mentre i verbi di parola si limitano a *ripetere* e, forse, *spiegare* (*parlare* non rientra, infatti, tra i *verba dicendi*).

In ben 25 casi la CONT compare in un costrutto relativo che fornisce il contesto imperfettivo in cui si inserisce il Tf, secondo lo schema di incidenza già più volte menzionato:

Franco, *che stava scendendo* i due scalini appoggiati alla soglia della porta, si fermò e aggrottò le sopracciglia

allora tacque, ritornò, tra soddisfatto e preoccupato, nella stanza *dove don Giuseppe stava spiegando* alla signora Barborin, con gesti appropriati, che la Maria le avrebbe portato qualche cosa da mangiare

In 14 occasioni la CONT compare in un contesto di incidenza con Tf determinato da una subordinata temporale introdotta da *quando* (7 occorrenze) e *mentre* (7 occorrenze; negli esempi qui sotto, si noti la doppia soluzione con CONT e IPF per indicare avvenimenti simultanei):

il professor Gilardoni stava leggendo sul suo belvedere dell'orto, *quando* vide Pasotti che veniva dietro il Pinella, fra le rape e le barbabietole

l'ultimo dì dell'anno, *mentre* Franco stava scrivendo le minutissime istruzioni che intendeva lasciare a sua moglie per il governo del giardinetto e dell'orto, *mentre* lo zio rileggeva per la decima volta la sua favorita Storia della diocesi di Como, Luisa uscì a passeggio con Maria

Nel periodo 1976-2005 la situazione muta radicalmente e vede la frequenza della PROG moltiplicata di quasi 35 volte, anche se cambia di poco la distribuzione dei tempi verbali. Si registra, in effetti, un certo incremento nelle subordinate al CONG, specie se rette da *sembrare* e *parere* o costrutti analoghi (cfr. i costrutti analoghi rilevati nel parlato da Villarini 1999):

adesso sembrava *si stessero preparando* all'arrembaggio

pareva che quello a cui si stavano avvicinando non fosse un loro diritto, ma qualcosa di cui *si stessero appropriando* con la forza e con l'inganno

ma tutti ridevano, come se *stessero preparando* una grossa e divertente burla a qualcuno

Più contenuto (19 occorrenze) il numero di CONT al FUT, sempre con valore epistemico:

i miei amici *si staranno chiedendo* che fine ho fatto

Uno spoglio cursorio, in attesa di una disamina più attenta dei contesti, dei principali verbi coniugati all'IPF individua, in ordine di frequenza, *arrivare* (che, come nota Bertinetto 1990: 132, inserito nella PROG assume valore imminenziale) con 51 occorrenze e i durativi *cercare* (49), *aspettare* e *parlare* (35), mentre anche *diventare* (29 occorrenze) è un trasformativo che porta alla lettura imminenziale.

### 3.3. CONT NEL CONFRONTO TRA TESTI NATIVI E TRADUZIONI NELL'OTTOCENTO

Nativi '800	andare			venire			andare%	venire%	Traduzioni '800	andare			venire		
	totale	totale%	andare%	venire%	totale	totale%				andare%	venire%				
PRE	18	2	20	13,15789	16,36364	4,761905			PRE	47	5	52	30,95238	30,51948	35,71429
IPF	79	24	103	67,76316	71,81818	57,14286			IPF	72	4	76	45,2381	46,75325	28,57143
PFS	7	8	15	9,868421	6,363636	19,04762			PFS	8	3	11	6,547619	5,194805	21,42857
PFC	0	1	1	0,657895	0	2,380952			PFC	1	1	2	1,190476	0,649351	7,142857
PPF	1	6	7	4,605263	0,909091	14,28571			PPF	5	1	6	3,571429	3,246753	7,142857
FUT	2	0	2	1,315789	1,818182	0			FUT	7	0	7	4,166667	4,545455	0
COND	0	0	0	0	0	0			COND	1	0	1	0,595238	0,649351	0
CONG PRE	0	0	0	0	0	0			CONG PRE	3	0	3	1,785714	1,948052	0
CONG IPF	3	1	4	2,631579	2,727273	2,380952			CONG IPF	5	0	5	2,97619	3,246753	0
CONG P	0	0	0	0	0	0			CONG P	1	0	1	0,595238	0,649351	0
INF	0	0	0	0	0	0			INF	4	0	4	2,380952	2,597403	0
Totale	110	42	152	100	100	100			Totale	154	14	168	100	100	100

Tabella 3 – CONT nei testi nativi e nelle traduzioni nell'Ottocento. Frequenze assolute e %

#### 3.3.1. CONT CON VENIRE NEL CONFRONTO TRA TESTI NATIVI E TRADUZIONI NELL'OTTOCENTO

Avendo già illustrato sopra l'andamento delle perifrasi che ci interessano nel subcorpus dei testi nativi, nei §§ dal 3.3.1. al 3.6. ci concentriamo sulla distribuzione nelle traduzioni.

Se il numero totale delle CONT nei due subcorpora qui considerati non presenta una differenza apprezzabile (rispettivamente, 154 occorrenze nei testi nativi, 168 nelle traduzioni), colpisce il netto dovario nell'impiego di *venire* come verbo modificatore, in particolare coniugato all'IPF. Pare infatti che i traduttori optino per quella che evidentemente considerano la variante più canonica della CONT, conformandosi alla tendenza alla normalizzazione ipotizzata dalla teoria degli universali traduttivi (cfr Ondelli 2020). In particolare, nelle traduzioni sembrano più rari gli indicatori di continuità e incremento:

ma, al sopravvenire, della notte, la folla – quasi onda crescente – di minuto in minuto s'era venuta addensando



da poi che la luce *man mano* veniva mancando, mi colse un vago, indeterminato fastidio

*e quanto io più m'appressava*, la prospettiva di quella casa mi veniva riconciliando col mio nuovo infortunio

nell'altro caso, il male si vien svolgendo invece *lentamente, a gradi*, e quasi direi per *fasi regolari*

Permangono, invece, i verbi di parola (per es., *dire, poetare, encomiare*), e si mantengono i trasformativi (per es., *riconciliare*), soprattutto pronominali (per es., *addensarsi, diradarsi, insinuarsi, radicarsi*).

Numerosi anche i casi come gli esempi qui sotto, in cui *venire* può essere inteso come verbo di movimento, di cui la gerundiva specifica in qualche modo lo svolgimento:

mentr'egli *veniva accostandosi* alla taverna, sul dinanzi della quale s' era accolto l'intero villaggio, surse ad un tratto un gridio universale

*veniva scortando* cinque uomini con folti mustacchi, lunghe spade, e che palesavano al par di lui un animo risoluto

la folla si unisce alla folla, ogni strada è serrata, una truppa di nemici gli *viene marciando incontro*

il vecchio sindaco *venne correndo* alla lugubre notizia:

una cotal cosa di sangue, che, a mo' di serpe, *vien tutta torcendosi e ritorcendosi* dall'inosservata, solitaria parte della scena di fondo!

un francesino de bon air, capitano, che *veniva per via saltellando*, mi fe' vedere che la cosa era sì facile che nulla più

Il carattere pseudoperifrastico di questi costrutti risulta più evidente dall'inserimento di materiale linguistico tra *venire* e il GER, come nelle ultime due frasi sopra, o quando siano presenti avverbiali locativi, come nel caso qui sotto:

si rivolse a me, e *venne sciorinandomi innanzi* tutta la pompa della sua scienza

È da escludere la perifrasi nei casi in cui il soggetto del verbo di moto non coincide con quello del GER, come nel costrutto riportato qui sotto, analogo al proverbio "l'appetito vien mangiando" visto in precedenza (§ 3.1.3.):

«Giovanotto, la fortuna viene dormendo»

Nelle traduzioni, la CONT costruita con *andare* registra una maggiore incidenza del PRE, a scapito dell'IPF, forse indice di una maggiore presenza nei dialoghi. Compare anche il PFC con il participio passato nella variante arcaica *ito/a*, utilizzata anche per formare il PPF e assente nei testi nativi:

se non che la poveretta s'è ita sempre aiutando alla meglio per trascinare il carro

L'unico COND rinvenuto esprime il futuro nel passato, secondo un modulo disusato nell'italiano moderno:

Il giorno 11, alle nove e undici minuti antimeridiane, la Luna doveva entrare nel suo ultimo quarto. dopo questo termine essa *andrebbe declinando*, e, quand' anche il cielo si fosse rasserenato, le probabilità dell'osservazione sarebbero in ispecial modo diminuite

Più frequente, nel confronto con i romanzi di autori italiani, il FUT, peraltro mai con modalità epistemica:

uno di quegli schiavi manterrà un fuoco vivo sotto il corpo tuo, e l'altro ti *andrà intanto stropicciando* d' olio quelle grame tue membra, come si unge l'arrosto, affinché non abbruci

Raro l'impiego con l'INF, assente dai testi nativi:

Ser John di Walton, che vide il numero degli Scozzesi *andar diminuendo*, sentì rinascere nel dubbio animo quel sospetto che da lungo tempo alterava la bontà del suo carattere

In tre casi sorge il dubbio che sia possibile lettura come verbo di movimento; in particolare, *andare + errando* pare configurarsi come una sorta di stereotipo:

e me ne accorsi tanto più alla sollecitudine che io avea di *andar errando* per essa ed ai castelli in aria, che io stava facendo, figurandomi che Serafina, tuttochè immersa nel dolore, avesse potuto discernere l'amor mio nascente e che ciò avesse con compiacenza osservato

sappiam alcun che d' un Douglas che sedette in un parlamento tenuto a Forfar dal re Malcolm I, e ne è noto che tanta era la sua premura *d'andar inseguendo* il cervo ne boschi, che si costruì nella foresta d'Etrick una torre, che forse esiste tuttora

m'alzo, e corro per la campagna; m' arrampico su per la costa ripida del monte, sento una pazza gioia *d'andar frugando* pel bosco, e aprirmi, a forza di braccia, un sentiero, attraverso gli intrecci delle fratte e dei cespugli

In effetti, anche con *andare*, nelle traduzioni troviamo diverse CONT in cui il verbo modificatore non risulta desemantizzato (in particolare quando si pre-

sentà come procomplementare: *andarsene*), mentre il GER può precisare il tipo di movimento:

e siccome i soldati incurvati si alzavano e *andavan vagando* qua e là, Alice si persuase che quel giuoco era veramente difficile

ma che anche gli adulti *vadano barcollando* pel mondo, a guisa di fanciulli, e com' essi ignorino d' onde siano venuti e dove s' avviino, com' essi agiscano senza determinato intento, e siano governati dalle chicche e dai confetti e dalla ferula, ecco ciò che nessuno vuol aver l'aria di credere

*andai volando* fin quella casa, dove trovai nella sala bene addobbata

In alternativa, il GER esprime un evento simultaneo al movimento:

ciò detto *se n'andò dicendo* confidentemente alla signora Giacinta ed a me, che si faceva mallevadore della vita dell'ammalato

attiravano l'attenzione di quei che *andavano* a piedi *battendo* il bosco e *stanando* la salvaggina in cui s' avvenivano

Un'altra possibilità è data dal GER che esplicita una conseguenza del movimento espresso da *andare*:

gli altri se ne andarono lasciandolo solo nella stanza

Quanto ai lemmi, con 9 occorrenze *dire* è di gran lunga il più frequente, e conduce il gruppo dei verbi di parola che già abbiamo incontrato in precedenza (*borbottare, predicare, discutere, mormorare e ripetere* – quest'ultimo con 4 occorrenze). Anche i verbi trasformativi abbondano (per es., *cessare* o *morire*), specie se pronominali: *accumularsi, accostarsi, accavallarsi, addensarsi, allontanarsi, cicatrizzarsi, farsi sano e forte, irradiarsi, perdersi, rimpicciolirsi* ecc.

Alla luce dei dati, non paiono registrarsi differenze significative rispetto ai testi nativi, se non nel numero e nella distribuzione dei tempi, più variegata nelle traduzioni. Tuttavia, a differenza della CONT costruita con *venire*, nel caso di *andare*, nel confronto coi testi nativi, nelle traduzioni non sembra registrabile un decremento degli indicatori di gradualità, continuità e incremento:

non già che manchino loro gli affari, chè anzi le faccende si vanno *ad ogni ora* più accumulando

lo sconforto e il tedio s' erano andati *sempre più* addensando nell'anima di Werther, finché, da ultimo, tutto l'esser suo ne fu padroneggiato con ferrea disciplina

se non che il mio canto andò *poco a poco* cessando

ed a me toccava riceverlo con una ripugnanza che andavasi aumentando e *col crescere dell'età*

Laura mi andava insegnando il nome dei comici e delle commedianti *di mano in mano* che si presentavano ai nostri sguardi

e chi avendo occhi per vedere ciò che l'occasione ed il tempo gli vanno *di continuo* mostrando a ogni passo del suo cammino, non trascura nulla di quanto egli può lecitamente toccare!

se non che la poveretta s'è ita *sempre* aiutando alla meglio per trascinare il carro.

la mia ferita si andava cicatrizzando *lentamente*

la servetta non si appigliava solamente ai suoi zii ed ai suoi cugini, ma andava *qualche volta* adescando anco gli estrani ed a fare la vedova nobile in casa della buona vecchia, di cui feci parola

### 3.4. PROG NEL CONFRONTO TRA TESTI NATIVI E TRADUZIONI NELL'OTTOCENTO

Nativi			Traduzioni		
		%			%
PRE	14	20	PRE	60	40,54054
IPF	50	71,42857	IPF	80	54,05405
PFS	3	4,285714	PFS	2	1,351351
PFC	0	0	PFC	0	0
PPF	0	0	PPF	0	0
FUT	0	0	FUT	2	1,351351
COND	0	0	COND	0	0
CONG PRE	0	0	CONG PRE	1	0,675676
CONG IPF	3	4,285714	CONG IPF	1	0,675676
CONG P	0	0	CONG P	0	0
INF	0	0	INF	2	1,351351
Totale	70	100	Totale	148	100

Tabella 4 – PROG nei testi nativi e nelle traduzioni nell'Ottocento. Frequenze assolute e %

Dalla tabella sopra si nota immediatamente la maggior frequenza assoluta della PROG in traduzione, ed essendo i due subcorpora di dimensioni paragonabili, il dato è già di per sé significativo. Come nel caso della CONT, anche qui aumenta la frequenza del PRE, forse a indicare un maggior numero di dialoghi nelle traduzioni (la preponderanza di PRE è stata notata nel LIP da Villarini 1999). Per il resto permangono solo due occorrenze di PFS, peraltro entrambe sospettabili di espri-

mere effettivamente uno stato, data la presenza di specificatori temporali. Nel primo degli esempi qui sotto, la collocazione locativa è espressa implicitamente dal verbo *prendere congedo*, nel secondo è esplicita (*la bettola dove*):

egolino stettero *ancora qualche tempo* celiando su questo proposito, dopo di che Pacheco prese congedo

nè mi occorre già fantasticar lungo tempo, perchè Clarino, servidore di don Alessio, mi si presentò all' impensata, ed io lo menai tosto alla prima bettola dove stettimo *gozzovigliando fino a mezzanotte*

Risultano, inoltre, due casi di FUT, sempre con lettura epistemica, assente invece dai testi nativi:

già, già, il vecchio *starà adempiendo* a qualcuno de' suoi doveri, come dice egli

noi lo troveremo, e tanto più facilmente che, se non si è divertito come voi a soccorrere gli uccelli oppressi, *starà cercandovi* egli pure

Per stabilire con sicurezza le classi azionali, l'analisi dei lemmi richiederebbe la descrizione dei contesti, che a questo punto della ricerca non è stata ancora condotta. Mi limito a segnalare che i due lemmi più frequenti sono, rispettivamente, durativo (*osservare*, 7 occorrenze) e non durativo (*morire*, 6), a testimonianza della maggior versatilità della PROG rispetto alla CONT.

3.5. CONT NEL CONFRONTO TRA TESTI NATIVI E TRADUZIONI NEL NOVECENTO

Nativi ,900	andare			venire			andare%	venire%	Traduzioni 900	andare			venire		
	andare	venire	totale	totale%	andare	venire				totale	totale%	andare	venire	totale	totale%
PRE	1	0	1	0,60241	0,606061	0			PRE	2	0	2	13,33333	14,28571	0
IPF	138	0	138	83,13253	83,63636	0			IPF	10	0	10	66,66667	71,42857	0
PFS	0	0	0	0	0	0			PFS	0	0	0	0	0	0
PFC	0	0	0	0	0	0			PFC	0	1	1	6,66667	0	100
PPF	23	1	24	14,45783	13,93939	100			PPF	2	0	2	13,33333	14,28571	0
FUT	0	0	0	0	0	0			FUT	0	0	0	0	0	0
COND	1	0	1	0,60241	0,606061	0			COND	0	0	0	0	0	0
CONG PRE	0	0	0	0	0	0			CONG PRE	0	0	0	0	0	0
CONG IPF	1	0	1	0,60241	0,606061	0			CONG IPF	0	0	0	0	0	0
CONG P	1	0	1	0,60241	0,606061	0			CONG P	0	0	0	0	0	0
INF	0	0	0	0	0	0			INF	0	0	0	0	0	0
Totale	165	1	166	100	100	100			Totale	14	1	15	100	100	100

Tabella 5 – CONT nei testi nativi e nelle traduzioni nel Novecento. Frequenze assolute e %

Nell'ultimo arco temporale coperto dal nostro corpus, la frequenza della CONT subisce un drastico ridimensionamento nelle traduzioni. Come nei testi nativi, si registra un solo caso di *venire*, peraltro coniugato a un tempo perfettivo, con indicazione di gradualità e pronomi benefattivo (cfr. Bertinetto 1990: 142). La compresenza di questi tre elementi fa propendere nettamente per la raggiunta stereotipia sintattica della struttura:

questo schema *mi si è venuto delineando man mano*

Se consideriamo il costrutto con *andare*, possiamo invece notare che l'IPF è per quasi metà coperto dal verbo *aumentare*, peraltro sempre in presenza degli indicatori di incremento e di specificatori del modo in cui si è verificato l'evento:

al ritmo ossessivo delle oscillazioni che andavano *sempre più aumentando* di velocità

li spingeva in là insieme con il mucchio di terra, che andava *via via aumentando*

c'era una certa tensione nell'aria, che *man mano* che le offerte crescevano andava *aumentando*

nel grigio dietro i vetri la proporzione di nero andava *aumentando impercettibilmente*, infine scese la sera

Alla luce dei risultati, è possibile concludere che, agli esordi del XXI secolo, nelle traduzioni la CONT rappresenta uno stereotipo stilistico scarsamente produttivo.

### 3.6. PROG NEL CONFRONTO TRA TESTI NATIVI E TRADUZIONI NEL NOVECENTO

Nativi			Traduzioni		
		%			%
PRE	819	33,92709	PRE	723	34,70955
IPF	1406	58,24358	IPF	1116	53,57657
PFS	0	0	PFS	0	0
PFC	0	0	PFC	0	0
PPF	0	0	PPF	0	0
FUT	19	0,787075	FUT	17	0,816131
COND	0	0	COND	0	0
CONG PRE	32	1,325601	CONG PRE	62	2,976476
CONG IPF	136	5,633803	CONG IPF	157	7,537206
CONG P	0	0	CONG P	0	0
INF	2	0,08285	INF	8	0,384061
Totale	2414	100	Totale	2083	100

Tabella 6 – PROG nei testi nativi e nelle traduzioni nel Novecento. Frequenze assolute e %

La tabella sopra dimostra che i numeri assoluti e le percentuali relative alla distribuzione della PROG risultano fondamentalmente equivalenti nei due subcorpora, fatta salva una leggera prevalenza del CONG in traduzione, probabile sintomo di un impiego più largo nelle frasi subordinate. Anche per quanto riguarda questo ambito, la fase successiva della ricerca sarà dedicata all'analisi dei contesti in cui compaiono i verbi principali delle perifrasi. Per il momento mi limito a segnalare la maggiore frequenza d'uso dell'INF (sebbene i numeri assoluti siano estremamente ridotti, a conferma della rarità del fenomeno), che nelle traduzioni si configura come una routine traduttiva secondo la formula [*avere la sensazione di*] o analoghe + [*stare* + GER]:

*avevo l'impressione* di stare consumando tempo ed energia

*avevo la sensazione* di stare interpretando un ruolo

provavo un costante disagio, *la deprimente sensazione* di stare sbagliando tutto

malgrado ciò provavo lo stesso *la sensazione* di star parlando dei fatti immaginari di una persona immaginaria

#### 4. CONCLUSIONI E PROSPETTIVE DI RICERCA

In diacronia, nei testi nativi, rispetto all'800 la CONT dimostra una sostanziale tenuta a livello quantitativo, che però è del tutto attribuibile ad *andare* come verbo modificatore, poiché *venire* (che pure è sempre risultato minoritario) passa da 42 occorrenze a una sola. Anche se potenzialmente la CONT accetta tutti i tempi verbali, i dati indicano il dominio dell'IPF, che si rafforza nell'ultimo scorcio del XX secolo, mentre in pratica scompare il PRE. Con tutta probabilità, per quanto concerne questo tempo verbale, la CONT viene a soffrire la concorrenza della PROG (che infatti registra un incremento), verosimilmente in seno ai dialoghi. Il contemporaneo incremento del PPF lascia pensare che, nell'italiano letterario di oggi, la CONT resti relegata a contesti formali, caratterizzati anche da una certa complessità sintattica.

Sempre nei testi nativi, se tra il 1860 e la fine del secolo la PROG risulta meno frequente della CONT e registra rari casi di tempi perfettivi, alla fine del Novecento, in sostanza, non c'è battaglia: il dominio quantitativo della PROG è incontrastato, e si amplia anche la varietà dei tempi verbali, in particolare per quanto riguarda il CONG, a indicare un più frequente impiego nelle subordinate.

Le traduzioni evidenziano sin dall'Ottocento le direttrici secondo cui procederà l'evoluzione generale delle perifrasi gerundivali in italiano: da un lato i traduttori optano più spesso per la forma più canonica della CONT, cioè quella costruita con *andare*; dall'altro fanno un uso più che doppio della PROG rispetto ai romanzieri italiani. In entrambi i casi, si dimostrano più flessibili nell'uso dei tempi verbali. Nell'ultimo quarto del Novecento i traduttori, in pratica, abbandonano la CONT, mentre si comportano in maniera analoga ai romanzieri italiani nell'uso della PROG, sia dal punto di vista della frequenza, sia se si guarda alla distribuzione dei tempi verbali. Sembra dunque che le traduzioni, almeno per i costrutti qui considerati, abbiano in qualche modo anticipato lo sviluppo dell'italiano.

Per poter confermare questo ruolo di anticipazione, occorrerà procedere a un'analisi di tipo qualitativo e valutare le variazioni nell'impiego delle perifrasi gerundivali in relazione ai tempi verbali e alle classi azionali dei lemmi coinvolti. Inoltre, lo spoglio di corpora relativi agli anni 1900-1975 potrà definire la natura graduale o improvvisa del sorpasso della PROG ai danni della CONT. Infine, una



classificazione più fine dei dati potrebbe gettare luce sulle differenze tra le opere del canone e la letteratura di consumo, sull'apporto dei singoli autori e traduttori, nonché sul ruolo eventuale delle lingue di partenza delle traduzioni.

#### APPENDICE - COMPOSIZIONE E CONSISTENZA DEL CORPUS

<i>Autore</i>	<i>Titolo</i>	<i>Anno</i>	<i>Occorrenze</i>
I. Nievo	<i>Il barone di Nicastro</i>	1860	54.453
G. Verga	<i>Storia di una capinera</i>	1871	31.529
C. Collodi	<i>Pinocchio</i>	1881	40.670
R. Fucini	<i>Le veglie di Neri: paesi e figure della campagna toscana</i>	1882	43.829
E. De Amicis	<i>Cuore</i>	1886	84.055
E. De Marchi	<i>Il cappello del prete</i>	1888	57.486
G. D'Annunzio	<i>Il piacere</i>	1889	105.195
L. Capuana	<i>Profumo</i>	1892	59.565
F. De Roberto	<i>I viceré</i>	1894	145.781
A. Fogazzaro	<i>Piccolo mondo antico</i>	1895	112.456
E. Salgari	<i>I pirati della Malesia</i>	1896	58.934
G. Deledda	<i>La via del male</i>	1896	67.091

Tabella 7 – Testi italiani e delle traduzioni 1800-1899. N= 861.044

Autore	Titolo	Anno	Occorrenze	Traduttore
J.P. Claris de Florian	<i>Numa Pompilio secondo re di Roma (Tomo I)</i>	1811	19.795	?
L. Sterne	<i>Viaggio sentimentale di Yorick lungo la Francia e l'Italia</i>	1813	43.500	D. Chierico
F.A. Chateaubriand	<i>Atala ovvero Gli amori di due selvaggi nel deserto</i>	1814	23.900	?
F. Schiller	<i>L'oste del sole ovvero Il delinquente per onore perduto</i>	1831	7.891	C. Belligoni
W. Scott	<i>Ivanhoe (vol. III)</i>	1832	34.612	A. Clerichetti
H. de Balzac	<i>Papà Goriot</i>	1835	85.448	L.M.
V. Hugo	<i>Le ore estreme di un sentenziato a morte</i>	1835	39.367	G.B. Carta
W. Scott	<i>Il castello pericoloso</i>	1843	38.537	L.M.
A.R. Lesage	<i>Storie di Gil Blas di Santillano (vol. I)</i>	1853	137.639	Q. Viviani
A. Dumas	<i>I tre moschettieri (vol. I)</i>	1855	43.844	C. Coriani
C. Dickens	<i>Le campane ovvero Il capo d'anno</i>	1855	50.924	P. Bettoni
W. Goethe	<i>I dolori del giovane Werther</i>	1858	52.044	R. Ceroni
E.A. Poe	<i>Storie incredibili (L'uomo della folla)</i>	1869	55.365	B.E. Maineri
L. Carroll	<i>Alice nel paese delle meraviglie</i>	1872	22.687	T. Pietrocola Rossetti
I.S. Turgenev	<i>Acque di primavera</i>	1872	45.718	S. de Gubernatis-Besobrasoff
J. Verne	<i>Dalla terra alla luna. Tragitto in 97 ore e 20 minuti</i>	1872	52.248	?
V. Hugo	<i>Novantatré (vol. II)</i>	1874	40.412	C. Pizzigoni
I.S. Turgenev	<i>Il primo amore e Assia</i>	1876	39.276	E. Zucchelli
C. Dickens	<i>Cantico di Natale</i>	1888	26.815	F. Verdinois

Tabella 8 – Traduzioni 1800-1899. N= 860.022

Autore	Titolo	Anno	Occorrenze
F. Cialente	<i>Le quattro ragazze Wieselberger</i>	1976	67.486
F. Tomizza	<i>La miglior vita</i>	1977	94.441
A. De Carlo	<i>Uccelli da gabbia e da voliera</i>	1982	70.610
C. Sgorlon	<i>L'armata dei fiumi perduti</i>	1985	88.049
U. Eco	<i>Il pendolo di Foucault</i>	1988	205.964
A. Baricco	<i>Castelli di rabbia</i>	1991	53.052
S. Veronesi	<i>Venite venite B-52</i>	1995	95.519
N. Ammaniti	<i>Fango</i>	1996	33.936
M. Buticchi	<i>Le pietre della luna</i>	1997	120.641
G. Faletti	<i>Io uccido</i>	2002	199.268
M. Mazzantini	<i>Non ti muovere</i>	2002	76.093
V.M. Manfredi	<i>L'impero dei Draghi</i>	2005	121.571

Tabella 9 – Testi italiani 1976-2005. N= 1.226.630

Autore	Titolo	Anno	Occorrenze	Traduttore
A. Christie	<i>Addio miss Marple</i>	1976	55.688	D. Fonticoli
M. Crichton	<i>Congo</i>	1980	89.529	E. Capriolo
S. King	<i>Pet cemetery</i>	1985	132.829	H. Brinis
P. Süskind	<i>Il profumo</i>	1988	81.071	G. Agabio
R. Dahl	<i>Storie impreviste</i>	1993	90.942	A. Veraldi
H. Murakami	<i>Dance dance dance</i>	1998	116.264	G. Amitrano
H. Corbin	<i>Cadaveri senza volto</i>	2000	104.456	F. Sessi
K. Follett	<i>Le gazze ladre</i>	2001	136.296	A. Raffo
I. Allende	<i>Il regno del drago d'oro</i>	2003	86.493	E. Liverani
C. Ruiz Zafón	<i>L'ombra del vento</i>	2004	134.436	L. Sezzi
T. Yamada	<i>Estranei</i>	2005	45.949	A. Martini
P. Coelho	<i>Lo Zahir</i>	2005	81.235	R. Desti

Tabella 10 – Traduzioni 1976-2005. N= 1.155.188

- Armellini G. & Colombo G. (1999) (a cura di) *La letteratura italiana*, Bologna, Zanichelli.
- Bertinetto P.M. (1986) *Tempo, aspetto e azione del verbo italiano. Il sistema dell'indicativo*, Firenze, Accademia della Crusca.
- Bertinetto P.M. (1990) "Perifrasi verbali italiane: criteri di identificazione e gerarchia di perifrasticità". In G. Bernini & A. Giacalone Ramat (a cura di), *La temporalità nell'acquisizione di lingue seconde: atti del Convegno internazionale, Pavia, 28-30 ottobre 1988*, Milano, Franco Angeli, pp. 331-350.
- Bertinetto P.M. (1991) "Il verbo". In L. Renzi & G. Salvi (a cura di), *Grande Grammatica Italiana di Consultazione*, Vol. II, Bologna, Il Mulino, pp.13-162.
- Bertinetto P.M. (1997) *Il Dominio tempo-aspetto. Demarcazioni, intersezioni, contrasti*, Torino, Rosenberg & Sellier.
- Bertinetto P.M. (1998) "Sui connotati azionali ed aspettuati della perifrasi continua ("andare/venire + Gerundio")". In G. Bernini, P. Cuzzolin & P. Molinelli (a cura di), *Ars linguistica: studi offerti da colleghi ed allievi a Paolo Ramat in occasione del suo 60. Compleanno*, Roma, Bulzoni, pp. 109-128.
- Bertinetto P.M. (2000) "The progressive in Romance, as compared with English". In By Ö. Dahl (ed.), *Tense and aspect in the languages of Europe*, Berlin/New York, Mouton de Gruyter, pp. 559-604.
- Bertinetto P.M. (2003) *Tempi verbali e narrativa italiana dell'Otto/Novecento*, Alessandria, Edizioni dell'Orso.
- Bertinetto P.M., Ebert K.H. & de Groot C. (2000), "The progressive in Europe". In By Ö. Dahl (ed.), *Tense and aspect in the languages of Europe*, Berlin/New York, Mouton de Gruyter, pp. 517-558.
- Brianti G. (2000) "Diacronia delle perifrasi aspettuati dell'italiano. Il caso di stare + gerundio, andare/venire + gerundio", *Lingua Nostra*, 61, pp. 35-52; 97-119.
- Cortelazzo M.A. (2007) "La perifrasi progressiva in italiano è un anglicismo sintattico?". In AA.VV. (ed.), *Studi in onore di Pier Vincenzo Mengaldo per i suoi settant'anni, a cura degli allievi padovani*, Firenze, SISMEL. Edizioni del Galluzzo, pp. 1753-1764.
- De Mauro T., Mancini F., Vedovelli M. & Voghera M. (1993) *Lessico di frequenza dell'italiano parlato*, Milano, Etaslibri.
- Degano C. (2005) "Influssi inglesi sulla sintassi italiana: uno studio preliminare sul caso della perifrasi progressiva". In A. Cardinaletti & G. Garzone (a cura di), *L'italiano delle traduzioni*, Milano, Franco Angeli, pp. 85-105.
- Ondelli S. (2019a) "I procomplementari nella prosa letteraria tra diacronia e traduzione", *Rivista Internazionale di Tecnica della Traduzione*, 21, pp. 99-117.
- Ondelli S. (2019b) "Verbi procomplementari, locuzioni idiomatiche e traduzione: uno sguardo in diacronia. Il caso di fare". In O.-D. Balaş, A. Gebäilä & R. Voicu (a cura di), *Fraseologia e paremiologia: prospettive evolutive, pragmatica e concettualizzazione*, Riga, Edizioni Accademiche Italiane (Omniscryptum Group), pp. 343-367.
- Ondelli S. (2020) "Una carrellata di clitici: osservazioni su alcuni possibili verbi procomplementari in un corpus di italiano scritto". In E. Pîrvu (a cura di), *Lingua e letteratura italiana nel presente e nella storia. Atti del X Convegno internazionale di italianistica dell'Università di Craiova, 14-15 settembre 2018*, Firenze, Franco Cesati Editore, Quaderni della Rassegna 177, pp. 139-154.

Ondelli S. (in stampa) *Le perifrasi gerundivali nel romanzo italiano 1800-2000: aspetti quantitativi*.

Squartini M. (1990) "Contributo per la caratterizzazione aspettuale delle perifrasi italiane *andare + gerundio, stare + gerundio, venire + gerundio*. Uno studio diacronico," *Studi e saggi linguistici*, 30, pp. 117-212.

Titus-Brianti G. (2010) "La perifrasi progressiva 'in progress': confronto tra italiano e inglese". In M. A. Terzoli, A. Asor Rosa & G. Inglese (a cura di), *Letteratura e filologia fra Svizzera e Italia, Studi in onore di Guglielmo Gorni, Vol. III Dall'Ottocento al Novecento: letteratura e linguistica*, Roma, Edizioni di Storia e Letteratura, pp. 351-362.

Villarini A. (1999) "Analisi delle occorrenze della perifrasi *stare+gerundio* all'interno di un corpus di italiano parlato". In M. Vedovelli (a cura di), *Indagini sociolinguistiche nella scuola e nella società italiana in evoluzione*, Milano, Franco Angeli, pp. 23-26.

# Premessa agli articoli di Sciandra, Trevisani e Tuzzi e di Sciumbata, Nadalutti e Tringali

FLORIANA CARLOTTA SCIUMBATA<sup>1</sup>  
Università di Trieste  
fsciumbata@units.it

Gli studi illustrati nei due articoli che seguono sono basati sullo stesso corpus di 100 opere di narrativa, scritte originariamente in italiano e pubblicati tra il 1825 e il 1923, per un totale di quasi 7.500.000 *token*, di cui illustreremo i dettagli qui di seguito. Si tratta di una raccolta preparatoria che abbiamo utilizzato per testare metodologie e strumenti da applicare prossimamente nell'ambito del progetto *Distant Reading for European Literary History* (COST Action CA16204). Questo stabilisce precisi criteri<sup>2</sup> per la creazione del corpus da analizzare – denominato *ELTeC* (*European Literary Text Collection*) –, come la percentuale di autrici (inclusa tra il 10 e il 50%), la presenza di massimo tre titoli per ogni autore preso in analisi, la distribuzione di testi di diversa dimensione, l'inserimento di prime edizioni e l'impiego solo di opere non tradotte. Il corpus deve inoltre essere vario, bilanciato e rappresentativo del periodo preso in analisi; deve contenere opere letterarie e paraletterarie; e deve includere solo opere che non sono più protette dal diritto d'autore. Il nostro corpus preparatorio segue solo alcuni di questi criteri ed è da considerarsi come una raccolta ancora in fase di lavorazione, che in futuro sarà adattata alle richieste per i corpora del progetto COST Action.

- 1 Hanno contribuito all'estrazione dei dati usati nella premessa e alla sua redazione anche Michele Cortelazzo, Paolo Nadalutti, Stefano Ondelli, Andrea Sciandra, Matilde Trevisani, Luca Tringali e Arjuna Tuzzi.
- 2 Cfr. [https://distantreading.github.io/sampling\\_\\_proposal.html](https://distantreading.github.io/sampling__proposal.html) (consultato il 10/12/2021).

Le statistiche che illustreremo qui di seguito sono state calcolate dopo una fase preliminare di pulizia automatica del corpus. I file, di pubblico dominio, raccolti da database online, sono probabilmente stati realizzati con l'ausilio di tecnologie OCR (*optical character recognition*), cioè strumenti che permettono la digitalizzazione di documenti cartacei e immagini. Tuttavia, l'uso di OCR può determinare errori, in particolare dovuti alla presenza di invii a capo che spezzano parole o frasi; all'indicazione del numero pagina; alla presenza di intestazioni e piè di pagina. Si tratta di elementi che falsano le statistiche relative al testo (per esempio, al numero totale di *token*); rendono difficile l'utilizzo di software di estrazione automatica dei dati, come lemmatizzatori e *tagger* di parti del discorso; non permettono l'applicazione di strumenti come gli indici di leggibilità. Per la pulizia, è stato creato uno *script* nel linguaggio di programmazione *Python* che corregge in automatico gli errori frequenti tramite espressioni regolari.<sup>3</sup>

Dopo la pulizia, il corpus è stato sottoposto a tokenizzazione per determinare il numero esatto di *token* totali. Sono stati perciò esclusi tutti gli spazi e i segni di punteggiatura per distinguere le parole vere e proprie.

Il corpus che abbiamo utilizzato contiene i titoli elencati qui di seguito, pubblicati tra il 1825 e il 1923, per i quali specifichiamo anche autore e anno di pubblicazione.

autore	titolo	anno
Aleramo, Sibilla	Il passaggio	1919
Aleramo, Sibilla	Una donna	1907
Baccini, Ida	Il principino	1883
Barbieri, Ulisse	Lucifero	1871
Battaglia, Giacinto	La lega lombarda	1834
Bersezio, Vittorio	La carità del prossimo	1868
Bersezio, Vittorio	Povera Giovanna	1869
Boito, Camillo	Il maestro di Setticlavio	1891
Boito, Camillo	Senso, Nuove storielle varie	1883
Capuana, Luigi	Giacinta	1879
Capuana, Luigi	Il marchese di Roccaverdina	1901
Capuana, Luigi	Profumo	1892
Caracciolo, Enrichetta	I misteri del chiostro napoletano	1864
Carcano, Giulio	Damiano	1850
Carcano, Giulio	La nunziata	1849

3 Lo script, a cura di Luca Tringali e Floriana Sciumbata, è disponibile all'indirizzo <https://github.com/flometis/ExSTRA/blob/master/pulisci.py> (consultato il 10/12/2021).

Castellani Fantoni Benaglio, Ines	Mia	1884
Collodi, Carlo	Le avventure di Pinocchio	1883
D'Annunzio, Gabriele	Il fuoco	1900
D'Annunzio, Gabriele	Il piacere	1889
D'Annunzio, Gabriele	Le vergini delle rocce	1895
De Amicis, Edomondo	Cuore	1886
Deledda, Grazia	Canne al vento	1913
Deledda, Grazia	La madre	1920
Deledda, Grazia	La via del male	1896
De Marchi, Emilio	Demetrio Pianelli	1890
De Marchi, Emilio	Giacomo l'idealista	1897
De Marchi, Emilio	Il cappello del prete	1888
De Roberto, Federico	I viceré	1894
De Roberto, Federico	Processi verbali	1890
Di Luanto, Regina	Nuovissimo amore	1903
Di Luanto, Regina	Per il lusso	1912
Dossi, Carlo	La colonia felice	1874
Dossi, Carlo	La desinenza in A	1873
Farina, Salvatore	Fante di picche	1874
Fogazzaro, Antonio	Daniele Cortis	1885
Fogazzaro, Antonio	Malombra	1881
Fogazzaro, Antonio	Piccolo mondo moderno	1901
Garibaldi, Giuseppe	Cantoni il volontario	1870
Garibaldi, Giuseppe	Clelia	1870
Guerrazzi Domenico	L'assedio di Firenze	1863
Guerrazzi Domenico	Pasquale Paoli	1860
Guglielminetti, Amalia	Anime allo specchio	1915
Guglielminetti, Amalia	La porta della gioia	1920
Guglielminetti, Amalia	Le ore inutili	1919
Imbriani, Vittorio	Dio ne scampi dagli Orsenigo	1876
Invernizio, Carolina	I misteri delle soffitte	1901
Invernizio, Carolina	Il bacio di una morta	1886
Invernizio, Carolina	La trovatella di Milano	1889
Manzoni, Alessandro	I promessi sposi	1825-1827 <sup>4</sup>
Manzoni, Alessandro	Storia della colonna infame	1840
Marchesa Colombi	Prima morire	1881
Melegari, Dora	Caterina Spadaro	1908
Neera	Lydia	1888
Neera	Un nido	1903

4 La datazione fa riferimento alla prima datazione dell'opera, pubblicata tra il 1825 e il 1827. Nelle rappresentazioni dei due articoli che seguono l'opera sarà indicata con la data intermedia (1826).



Neera	Una passione	1880
Negri, Ada	Le solitarie	1917
Nievo, Ippolito	Le confessioni di un italiano	1860
Nievo, Ippolito	Il barone di Nicastro	1867
Oriani, Alfredo	Al di là	1877
Oriani, Alfredo	No	1881
Perodi, Emma	Il principe della marsiliana	1891
Petruccelli, Ferdinando	Il sorbetto della regina	1890
Petruccelli, Ferdinando	Memorie di Giuda	1867
Petruccelli, Ferdinando	I moribondi del palazzo Carignano	1862
Pirandello, Luigi	I vecchi e i giovani	1904
Pirandello, Luigi	Il fu Mattia Pascal	1913
Pirandello, Luigi	L'esclusa	1901
Praga, Emilio	Memorie del presbiterio	1881
Righetti, Carlo	La scapigliatura	1862
Rovani, Giuseppe	Cent'anni	1859
Rovetta, Gerolamo	Casta Diva	1903
Rovetta, Gerolamo	Mater dolorosa	1882
Ruffini, Giovanni	Il dottor Antonio	1856
Ruffini, Giovanni	Lorenzo Benoni	1853
Salgari, Emilio	Alla conquista di un impero	1907
Salgari, Emilio	I misteri della giungla nera	1887
Salgari, Emilio	I pirati della Malesia	1896
Serao, Matilde	Cuore inferno	1881
Serao, Matilde	Dal vero	1879
Serao, Matilde	La mano tagliata	1912
Silvola, Rodolfo Giuseppe	Alboino in Italia	1840
Slataper, Scipio	Il mio Carso	1912
Svevo, Italo	La coscienza di Zeno	1923
Svevo, Italo	Senilità	1898
Svevo, Italo	Una vita	1892
Tarchetti, Iginio Ugo	Fosca	1869
Tartufari, Clarice	Il miracolo	1909
Tartufari, Clarice	L'albero della morte	1912
Tartufari, Clarice	Maestra	1887
Tommaseo, Niccolò	Fede e bellezza	1840
Torricelli, Giovan Battista	Oddantonio Feltre	1860
Tozzi, Federigo	Con gli occhi chiusi	1919
Tozzi, Federigo	Tre croci	1920
Vamba	Giamburrasca	1912
Verga, Giovanni	Il marito di Elena	1882
Verga, Giovanni	Tigre Reale	1875

Verga, Giovanni	Una peccatrice	1866
Vertua Gentile, Anna	Il romanzo di una signora perbene	1897
Zuccoli, Luciano	La freccia nel fianco	1913
Zuccoli, Luciano	Roberta	1897

Il corpus presenta le seguenti statistiche:

	<b>numero di autori</b>
<b>donne</b>	16 (29,6%)
<b>uomini</b>	38 (70,4)
<b>totale</b>	<b>54</b>

	<b>numero di opere</b>
<b>donne</b>	30
<b>uomini</b>	70
<b>totale</b>	<b>100</b>

<b>anno</b>	<b>numero di opere</b>
1826	1
1834	1
1840	3
1849	1
1850	1
1853	1
1856	1
1859	1
1860	3
1862	2
1863	1
1864	1
1866	1
1867	2
1868	1
1869	2
1870	2
1871	1
1873	1
1874	2
1875	1
1876	1
1877	1
1879	2
1880	1
1881	5

1882	2
1883	3
1884	1
1885	1
1886	2
1887	2
1888	2
1889	2
1890	3
1891	2
1892	2
1894	1
1895	1
1896	2
1897	3
1898	1
1900	1
1901	4
1903	3
1904	1
1907	2
1908	1
1909	1
1912	5
1913	3
1915	1
1917	1
1919	3
1920	3
1923	1
<b>totale</b>	<b>100</b>

<b>decennio</b>	<b>numero di opere</b>
1820_1829	1
1830_1839	1
1840_1849	4
1850_1859	4
1860_1869	13
1870_1879	11
1880_1889	21
1890_1899	15
1900_1909	13
1910_1920	17
<b>totale</b>	<b>100</b>

# Sulle tracce dell'espressione dell'interiorità: analisi diacronica di un corpus di narrativa italiana del XIX-XX secolo

ANDREA SCIANDRA

Università di Modena e Reggio Emilia  
[andrea.sciandra@unimore.it](mailto:andrea.sciandra@unimore.it)

MATILDE TREVISANI

Università di Trieste  
[matilde.trevisani@deams.units.it](mailto:matilde.trevisani@deams.units.it)

ARJUNA TUZZI

Università di Padova  
[arjuna.tuzzi@unipd.it](mailto:arjuna.tuzzi@unipd.it)

## ABSTRACT

This article concerns an example of statistical analysis of a large diachronic corpus, including 100 works of Italian fiction, with the aim of recognizing prototypical trends in words frequencies and identifying clusters of words that share similar temporal patterns. Starting from the insights of literary scholars, this study tries to test the hypothesis (often brought forward in literary history) that the inner life of the authors and/or their characters gained more importance during the Nineteenth century and represents one of the cornerstones of modernism. Although this hypothesis cannot be rejected, the adopted statistical methods do not highlight clear-cut temporal patterns: some slight trends emerge but a large share of words individually show an undifferentiated behaviour over time. The idea of studying the occurrence in the texts of the inner expression of the writers from a diachronic perspective is certainly very ambitious. This first experiment shows some strengths of the statistical tools but also some limits of the criteria adopted for both the constitution of the corpus and the selection of linguistic features to be observed.

## KEYWORDS

Diachronic corpora, chronological textual data, word embedding, functional data analysis, curve clustering

## 1. INTRODUZIONE

Il progetto europeo *Distant Reading for European Literary History* (COST Action CA16204, <https://www.distant-reading.net/>) è una rete internazionale di studiosi che aspira a sviluppare risorse per leggere la storia della letteratura europea attraverso le lenti dei metodi computazionali. Nell'ambito del progetto sono stati raccolti corpora di testi narrativi di diversi paesi e in diverse lingue con l'obiettivo di studiare diacronicamente fenomeni di interesse linguistico e letterario. In particolare, uno dei gruppi di lavoro del progetto si è concentrato sui problemi della periodizzazione della produzione letteraria con l'idea di verificare l'ipotesi, spesso avanzata nella storia della letteratura europea, secondo cui la rappresentazione della vita interiore dei personaggi abbia assunto maggiore importanza nel corso del XIX secolo e sia diventata uno dei capisaldi del modernismo.

Per poter verificare questa ipotesi attraverso metodi computazionali è necessario innanzitutto tradurre la domanda di ricerca in dati osservabili, cioè in fenomeni che possono essere identificati nei testi con strumenti informatici, e poi adottare strumenti statistici adeguati per la loro analisi e interpretazione.

Per quanto riguarda i fenomeni da isolare per seguirne lo sviluppo nel corso del tempo, il gruppo di lavoro del COST Action ha suggerito due percorsi differenti: (1) ricostruire la costellazione di significati che ruota attorno a sei verbi-seme: *feel, think, believe, know, hope, wish* e (2) selezionare i verbi più frequenti presenti nel corpus, con particolare riferimento a quelli che descrivono la vita interiore dell'autore e/o dei suoi personaggi.

Il presente lavoro illustra alcuni risultati ottenuti adottando i criteri del primo percorso suggerito con l'obiettivo di offrire un contributo allo sviluppo di metodi quantitativi per la storia della lingua che, come spiega Michele Cortelazzo in questo numero della rivista, rappresenta un campo ancora solo parzialmente esplorato ma sta assumendo una rilevanza specifica e un'identità come filone di ricerca autonomo. Sempre in questo numero della rivista, anche i lavori di Floriana Sciumbata, Paolo Nadalutti e Luca Tringali e di Stefano Ondelli si occupano, da prospettive diverse, del problema dello studio di corpora diacronici e presentano risultati che cercano di dare risposte a questioni emerse nell'ambito dello stesso progetto europeo e discusse durante il convegno *Esperimenti di Distant Reading. Estrazione, analisi e visualizzazione di dati linguistici da corpora letterari*, che si è tenuto all'Università di Trieste – IUSLIT/SSLMIT il 22 ottobre 2021.

Il corpus di testi italiani è stato progettato e raccolto come corpus propedeutico a quello previsto dal progetto COST Action e segue alcune linee di indirizzo del *European Literary Text Collection* (ELTeC): è costituito da 100 opere di narrativa pubblicate nel periodo 1825-1923 da 54 autori, 16 donne e 38 uomini. Con un'ampiezza di oltre sette milioni di occorrenze totali, la raccolta si può considerare un corpus di grandi dimensioni.

Una descrizione del corpus analizzato è disponibile in questo numero della rivista come introduzione al presente articolo e quello di Sciumbata, Nadalutti e Tringali (2021).

Il corpus di testi italiani è stato utilizzato nella versione di base, cioè sotto forma di semplici file di testo indipendenti, ripuliti dagli elementi paratestuali (intestazioni, numeri di pagina, ecc.) e privi di taggatura. Gli strumenti adottati per l'analisi statistica dei dati testuali sono stati sviluppati in ambiente R (R Core Team, 2021) sia attraverso procedure scritte ad hoc sia integrando pacchetti esistenti. Nel corso dell'elaborazione è stata creata in R una versione lemmatizzata attraverso il pacchetto `udpipe` (Wijffels 2021, Straka e Straková 2017) e la *Italian Stanford Dependency Treebank* (ISDT).

Dopo aver svolto un riconoscimento delle parole presenti nei testi come sequenze di lettere isolate nel testo per mezzo di spazi e punteggiatura (tokenizzazione), in vista dell'analisi statistica il corpus è stato rappresentato mediante matrici di dati del tipo "parole per testi" (*term-document matrix*) che riportano il numero di occorrenze di ciascuna parola in ciascun testo. Inoltre, dato che si tratta di un corpus diacronico, l'occorrenza di una parola è stata osservata con riferimento all'anno di pubblicazione dell'opera (figura 1) ed è diventata, quindi, una sequenza ordinata nel tempo, rappresentabile graficamente come una traiettoria. In questo caso le occorrenze di opere pubblicate nello stesso anno sono state aggregate.

Per rispondere alle domande di ricerca relative alla periodizzazione del corpus preso in analisi, sono state selezionate le traiettorie di un insieme limitato di parole e sono stati adottati strumenti statistici in grado di trovare regolarità utili a leggere l'andamento temporale della loro occorrenza e a costituire gruppi omogenei di parole che condividono un'evoluzione simile.

	1826	1834	1840	1849	1850	1853	...	1919	1920	1923
adesso	12	0	24	25	75	2	...	10	60	42
almeno	82	4	49	15	46	50	...	27	45	43
amare	20	13	163	21	59	48	...	86	56	144
amico	138	7	59	15	194	146	...	80	101	72
amore	43	27	196	23	62	32	...	89	59	124
ancora	293	24	157	95	228	64	...	177	176	245
andare	1056	16	234	98	315	332	...	267	385	300
anima	30	11	85	27	82	27	...	55	34	7
animo	137	27	143	21	102	84	...	28	28	61
aspettare	148	3	29	34	81	79	...	53	97	77
assai	9	6	70	3	5	80	...	6	12	7
atto	88	13	101	9	42	50	...	22	20	47
avere	3247	187	1650	594	1784	1696	...	1245	1371	3102
avvenire	49	5	36	4	43	38	...	24	18	63
bastare	111	3	38	6	61	31	...	22	41	80
bello	158	39	120	72	160	116	...	100	76	126
bene	557	15	349	103	340	193	...	119	174	245
...	...	...	...	...	...	...	...	...	...	...

Figura 1 – Esempio di matrice di occorrenze “parole per anno” (estratto)

### 3. SELEZIONE DELLE PAROLE OGGETTO DI STUDIO CON *WORD EMBEDDING*

Il gruppo di ricerca del COST Action ha suggerito di ricostruire l’insieme di parole semanticamente riconducibili ai sei verbi-seme (*feel, think, believe, know, hope, wish*) tradotti in italiano (*sentire, pensare, credere, sapere, sperare, desiderare*) e di costruire una lista di parole per le quali rappresentare le occorrenze (normalizzate) su grafici allo scopo di trovare tendenze.

Il metodo suggerito per ricostruire l’insieme di parole collegate ai sei verbi-seme è il *word embedding* (Mikolov et al. 2013). Il *word embedding* è una rappresentazione delle parole di un corpus tramite vettori numerici tale per cui le parole con un significato simile sono rappresentate da vettori simili. Questi vettori identificano quindi le coordinate di ciascuna parola in uno spazio multidimensionale.

Nell’ambito dell’analisi statistica testuale, diverse tecniche si riferiscono a un approccio *bag-of-words* (BoW), che essenzialmente utilizza liste di parole pe-

sate con la loro frequenza. Tra gli svantaggi più importanti del BoW troviamo che l'ordine delle parole viene ignorato e che il BoW generalmente non riesce a cogliere la semantica delle parole. Le tecniche di *word embedding* rappresentano una delle soluzioni più comuni a questi problemi, in quanto generano un numero di dimensioni molto ridotto rispetto al vocabolario di un corpus, ma generalmente in grado di rilevare regolarità linguistiche. I vettori che definiscono le parole possono essere utilizzati anche per identificare analogie come "L'uomo sta alla donna, come il re sta alla regina". Questo risultato si ottiene definendo una finestra di contesto – cioè una finestra di parole che precedono o seguono una parola focale – che viene utilizzata per addestrare un modello di *word embedding*.

In questa applicazione è stato utilizzato il metodo *GloVe* (Pennington, Socher e Manning 2014), uno dei sistemi più comuni per il *word embedding*, che si applica alla matrice globale di co-occorrenze (parole-per-parole) ottenuta da un determinato corpus. In confronto ad altri metodi come *Word2Vec*, *GloVe* mostra solitamente una migliore capacità nell'identificare le relazioni non lineari nello spazio vettoriale e conferisce un peso inferiore alle parole altamente frequenti, evitando che dominino il processo di formazione dei vettori.

Dopo aver svolto alcune operazioni preliminari sui testi che servono a ottenere una lista di frequenza dei lemmi presenti nel corpus (riduzione delle parole in tutto minuscolo, lemmatizzazione con `R udpipe`, rimozione delle parole grammaticali e dei numeri) si è scelto di mantenere nell'analisi solo le parole (lemmi) con almeno 5 occorrenze nel corpus. Per ogni verbo-seme è stata utilizzata l'implementazione del metodo *GloVe* del pacchetto `R text2vec` (Selivanov, Bickel e Wang 2020), definendo una finestra di 5 parole a destra e a sinistra di ogni termine. Successivamente, sono state selezionate le prime 100 parole, ordinate per similarità decrescente, per ciascuno dei sei verbi-seme. Unificando le sei liste di 100 parole ciascuna ed eliminando i duplicati, si ottiene una lista<sup>1</sup> di 134 parole di interesse per l'analisi, in quanto dovrebbe delineare l'insieme di parole che meglio rappresentano l'area semantica dei sei verbi-seme di partenza.

1 adesso, almeno, amare, amico, amore, ancora, andare, anima, animo, aspettare, assai, atto, avere, avvenire, bastare, bello, bene, bisognare, bisogno, buono, capire, caro, caso, cercare, certo, chiamare, chiedere, cominciare, comprendere, conoscere, continuare, cosa, cuore, dare, dimenticare, dio, dire, dolore, domandare, donna, dovere, ecco, essere, famiglia, fanciulla, fare, fatto, figlio, figliuolo, finire, forse, forte, forza, fratello, giovane, grande, idea, insieme, intanto, intendere, invece, lasciare, luogo, madre, mai, male, marito, meglio, metro, mettere, modo, moglie, momento, mondo, morire, morte, mostrare, nome, non, nuovo, ogni, padre, parere, parlare, parola, parte, pensiero, perdere, persona, piacere, pieno, poi, potere, povero, prendere, presto, provare, punto, qualche, qui, quinto, ragione, rendere, resto, riconoscere, ricordare, rimanere, rispondere, riuscire, sembrare, sempre, servire, sicuro, signore, soltanto, sorella, stare, subito, temere, tempo, tenere, toccare, trattare, trovare, uomo, vecchio, vedere, venire, veramente, vero, vita, vivere, volere, volta



L'idea di cercare regolarità nelle traiettorie tracciate da osservazioni temporali di dati discreti è molto studiata in statistica e, con riferimento specifico allo studio dell'andamento temporale delle occorrenze di parole in corpora diacronici, risulta utile fare riferimento ai risultati già ottenuti in precedenza su corpora di letteratura scientifica (Tuzzi 2018, Trevisani e Tuzzi 2018, Trevisani e Tuzzi 2015) e di discorsi istituzionali (Trevisani e Tuzzi 2013).

La procedura adottata per il riconoscimento della dinamica delle traiettorie tracciate dai lemmi identificati attraverso il *word embedding* lavora in un'ottica di apprendimento (*statistical learning*) in tre passi: (1) normalizzazione delle occorrenze, (2) lisciamento della traiettoria (*smoothing*) e (3) raggruppamento delle parole con andamento simile (*curve clustering*).

Le traiettorie analizzate sono 140 (una per ogni verbo-eme e per ogni lemma selezionato con il *word embedding*) e i punti-anno di osservazione sono in tutto 56, cioè tutti gli anni compresi tra il 1825 e 1923 per i quali nel corpus è disponibile almeno un'opera.

L'occorrenza (assoluta) di una parola in un testo non rappresenta una buona misura del suo tasso di presenza in quanto il valore dipende dall'ampiezza del testo. In questo senso una prima trasformazione che tenga conto delle dimensioni in gioco, cioè il calcolo di una frequenza relativa, è un passaggio obbligato. Più in generale il tipo di normalizzazione viene scelto in base alle caratteristiche delle traiettorie e, in questo caso, si è optato per una normalizzazione non lineare che tiene conto sia dell'ampiezza dei testi (la frequenza relativa di una parola  $i$  in un anno  $j$  viene calcolata dividendo l'occorrenza osservata  $n_{ij}$  per il numero di occorrenze totali  $N_j$  nei testi dello stesso anno  $j$ ) sia della popolarità della parola (la frequenza relativa viene ulteriormente normalizzata operando una trasformazione non lineare che, oltre a standardizzare tra 0 e 1 il *range* di frequenza per parola, trasforma ogni valore in una media pesata delle differenze dal minimo e rispettivamente dal massimo, per tenere conto della presenza di asimmetria positiva o negativa caratterizzante la parola; Grilli, Russo e Gismondi 2012). Questo tipo di normalizzazione riesce a tenere conto della diversa ampiezza dei subcorpora associati a ciascun anno di osservazione, a ridurre l'effetto delle parole ad alta frequenza, per evitare che mettano in ombra tutte le altre, e a smussare in qualche misura i bruschi saliscendi (picchi e precipizi) delle traiettorie, dovuti all'asimmetria dello spettro di frequenza specifico di ogni parola.

In particolare, se  $x_{ij} = n_{ij}/N_j$  indica la frequenza relativa (rispetto alla dimensione  $N_j$  dell'anno  $j$ ) della parola  $i$ , la trasformazione non lineare  $y=f(x)$  è tale per cui

$$\frac{y_{ij} - m_i^y}{M_i^y - y_{ij}} = \frac{p_i (x_{ij} - m_i^x)}{(1 - p_i)(M_i^x - x_{ij})}$$

dove  $p_i$  è un parametro di lisciamo, tra 0 e 1, per la differenza tra  $x_{ij}$  e il minimo  $m_i^x$ . Se  $p_i = 0.5$ , il rapporto delle distanze del valore dal minimo e rispettivamente dal massimo è uguale sia nella scala del valore trasformato  $y$  che nella scala del valore originario  $x$ , altrimenti il peso assegnato alla distanza dal minimo è maggiore o minore di quello assegnato alla distanza dal massimo a seconda del tipo di asimmetria (maggiore se asimmetria positiva, minore se negativa). Inoltre, per uniformare il livello delle traiettorie delle parole (l'effetto *popolarità*) si è imposto  $m_i^y = 1$  e  $M_i^x = 1$  per ogni parola.

La ricerca di regolarità nelle traiettorie osservate muove dall'idea che la frequenza normalizzata di una parola in un punto del tempo (anno) sia in grado di rappresentarne la vitalità, come un singolo fotogramma coglie un'immagine in un preciso istante, e che la sequenza delle occorrenze nel corso di diversi punti del tempo ne possa cogliere il ciclo di vita, come una serie di fotogrammi messi in successione ricostruisce un'immagine in movimento.

Dal punto di vista statistico è opportuno in questo contesto adottare un approccio per dati funzionali (*functional data analysis*) e leggere ogni osservazione, cioè la frequenza normalizzata  $y_{ij}$  di una parola  $i$  in un istante di tempo  $j$ , come la realizzazione di una sottostante funzione continua, che ne determina lo sviluppo nel tempo ma non è direttamente osservabile. A partire da una serie di osservazioni, si vuole ricostruire la dinamica generale, cioè la forma della funzione  $x$  che ha generato la traiettoria osservata:

$$y_{ij} = x_i(t_j) + \varepsilon$$

tenendo conto che l'osservazione è affetta da un rumore di fondo (una componente stocastica indicata come un errore  $\varepsilon$ ) così come i fotogrammi risultano un po' sfocati perché l'immagine non è statica.

La figura 2 rappresenta l'insieme delle 140 traiettorie normalizzate oggetto di analisi (sei verbi-*seme* e 134 parole estratte attraverso il *word embedding*).

La sfida dei passaggi successivi consiste nel riuscire a estrarre da un ammasso indifferenziato di traiettorie gruppi di parole che condividono uno sviluppo temporale simile.

Attraverso un'operazione di filtraggio (*smoothing*) le traiettorie originali vengono lisiate in maniera da eliminare le asperità che rendono illeggibile l'andamento di fondo. Il lisciamo delle traiettorie viene svolto grazie a una scomposizione della funzione in una combinazione lineare di funzioni più semplici che, in questo esempio, è costituita da *B-spline* (Ramsay e Silverman 2005).

Attraverso un approccio di stima che penalizza l'irregolarità delle curve (*roughness penalty*) si ottiene un lisciamo ottimale regolando opportunamente i parametri che definiscono le curve. Come esempio, la figura 3 riporta l'andamento delle frequenze normalizzate dei sei verbi-*seme* e, in rosso, la versione liscata della traiettoria che ne evidenzia la tendenza di fondo.

Le traiettorie lisce vengono quindi raggruppate in insiemi (*clusters*) di parole aventi un andamento temporale simile (*curve clustering*) tramite un algoritmo (*k-means*) che utilizza una distanza euclidea tra curve come misura di similarità. L'algoritmo viene reiterato 20 volte (variandone l'inizializzazione) per ogni potenziale numero di gruppi (da un minimo di 2 a un massimo di 26 gruppi), generando così 20 partizioni possibili per ogni numero.

Per determinare il numero di gruppi ottimale si confrontano circa 50 diversi criteri di qualità del *clustering* al fine di trovare il miglior compromesso tra tutti i suggerimenti ottenuti (figura 4). Stabilito il numero ottimale, si individuano tra le 20 partizioni generate quelle scelte dai criteri di qualità e, tra quest'ultime, quella che massimizza il grado di sovrapposizione con le altre tramite un indice (*Rand index*) che rappresenta la stabilità e coerenza dei gruppi.

In questa fase, per scegliere la configurazione di *clustering* definitiva, è opportuno inserire tra i criteri di valutazione anche la lettura qualitativa da parte di esperti, per dare senso ai gruppi di parole identificati dalla procedura e avviare da qui l'interpretazione.

I calcoli necessari sono stati operati con il supporto di librerie e pacchetti R, *fda* (Ramsay et al. 2021), *clusterCrit* (Desgraupes 2018), *cclust* (Dimitriadou 2021), *clusterSim* (Walesiak e Dudek 2020), *km1* (Genolini et al. 2005), integrati da codice ad hoc.

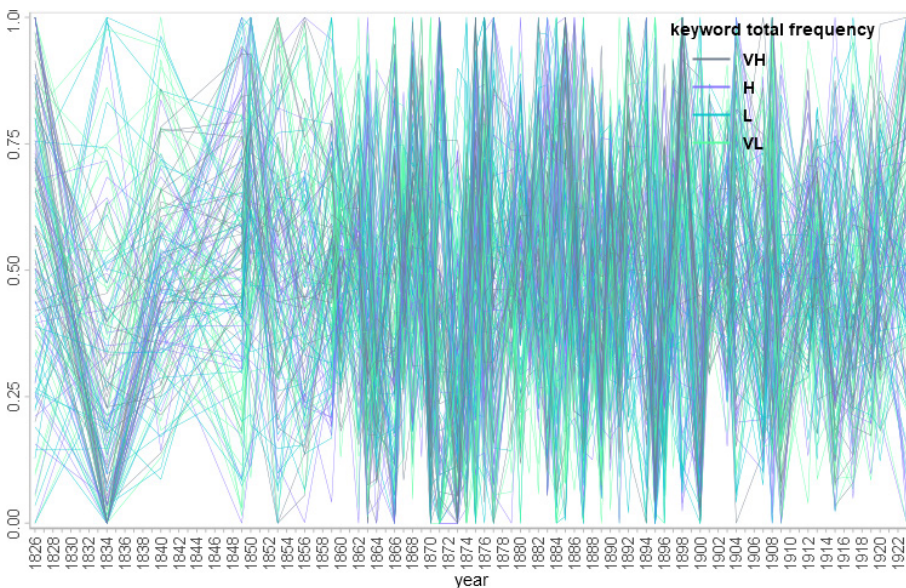


Figura 2 – Rappresentazione delle traiettorie normalizzate delle 140 parole oggetto di analisi. Parole con frequenza molto elevata (VH), elevata (H), bassa (L), molto bassa (VL)

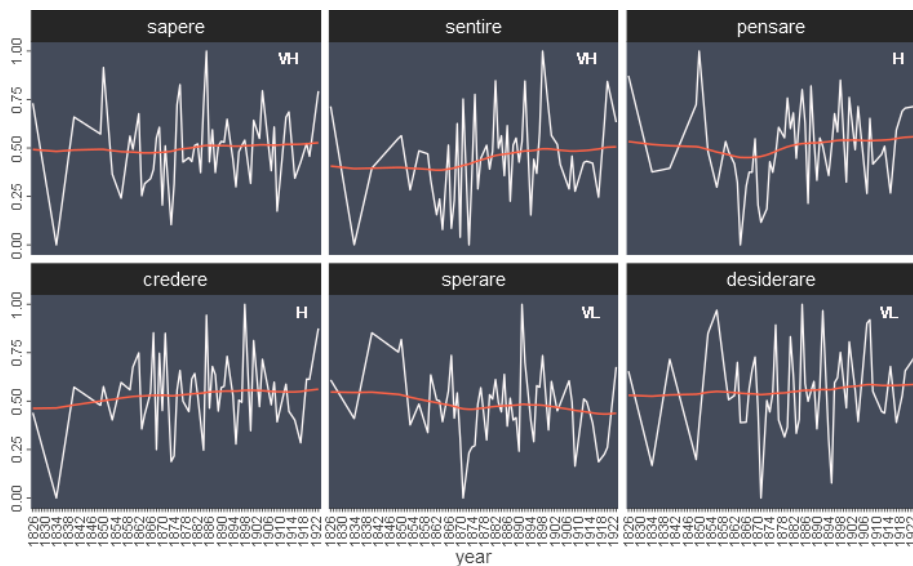


Figura 3 – Rappresentazione delle traiettorie normalizzate dei sei verbi-seme

## 5. RISULTATI

Il numero ottimale di gruppi suggerito dall'analisi dei criteri di qualità si può considerare quello pari a 6 (fig. 4) in quanto è fra i valori più scelti assieme a 3, 4 e 5, ma ha il vantaggio di creare partizioni più fini e di essere stato spesso indicato come prima o seconda soluzione. I metodi statistici impiegati non evidenziano andamenti temporali molto netti: alcune timide tendenze emergono, ma molte parole mostrano individualmente un comportamento indifferenziato nel corso del tempo. Sebbene il numero di gruppi e il numero di parole siano limitati, si può cercare di leggere congiuntamente alcuni andamenti interessanti a livello di *cluster* (figura 5).

Per quanto riguarda la collocazione dei sei verbi-seme, si può osservare come la frequenza di *sentire* (cluster B) e *pensare* (cluster C) abbia sperimentato una lieve crescita nel corso del tempo; *sapere* e *sperare* (cluster A), *credere* e *desiderare* (cluster F) si trovano invece in due cluster dove le parole mostrano una tendenza costante (molto numeroso e leggermente decrescente A, leggermente crescente F). Anche se è molto difficile trarre conclusioni sui temi trattati a partire da un numero così limitato di parole, i legami familiari e personali sembrano avere andamenti interessanti: *moglie* (B), *sorella* e *marito* (C) sono crescenti, restano presenze costanti *madre* e *amico* (F) *fratello*, *figlio* e *figliolo* (A) e risultano, invece, in diminuzione i riferimenti a *famiglia* (D) e *padre* (E).

Cresce nel tempo la frequenza di *donna* (B) mentre cala quella di *uomo* (D).

Il *dolore* e la *morte* mostrano nel periodo di tempo osservato una presenza pressoché costante (A) così come *morire* (F); mentre *vivere*, *piacere*, *continuare*

(B), *volere e vita* (C) sono crescenti. Sono crescenti anche *amare* (C), *anima e amore* (B), mentre i riferimenti al *cuore* e all'*animo* (D) sembrano in declino. Anche *Dio, male, buono* (E) e *bene, bello, ragione* (A, assieme a *sperare e sapere*) mostrano una tendenza in diminuzione o costante e rappresentano spunti interessanti per futuri approfondimenti specifici. Alcuni verbi appartenenti al cluster C, come *parlare, rispondere, chiedere, domandare, ricordare, capire e comprendere* risultano avere una tendenza di fondo in crescita e potrebbero indicare una maggiore attenzione all'interazione e al dialogo sia con altre persone sia a livello interiore.

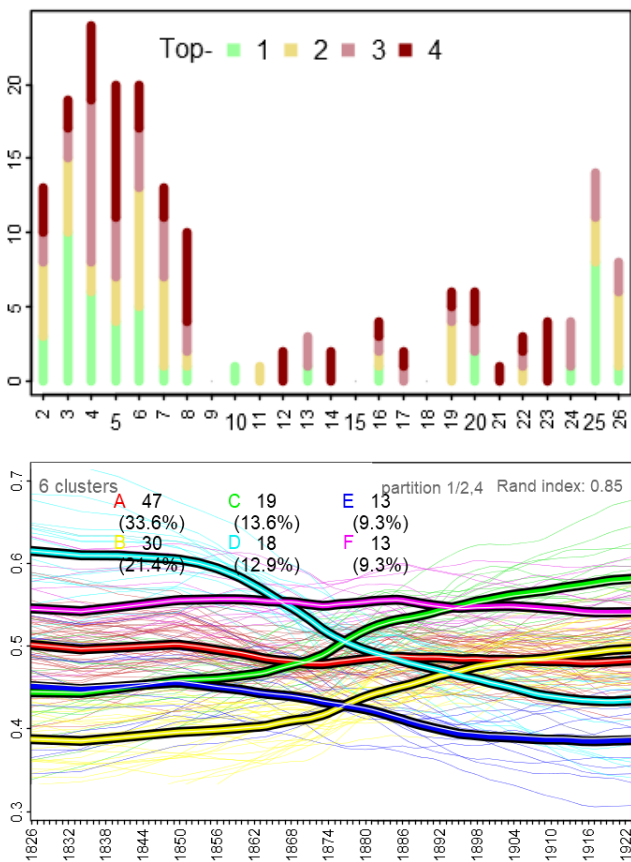


Figura 4 - Sopra, confronto tra le scelte del numero ottimale di gruppi suggeriti dal basket dei criteri: numero di criteri (in ordinata) che assegnano la preferenza al numero di gruppi (in ascissa) come prima (1 verde), seconda (2 giallo), terza (3 rosa) o quarta (4 vinaccia) scelta. Sotto, soluzione con 6 gruppi dove viene evidenziata la curva che rappresenta l'andamento medio e vengono riportate numerosità e percentuale di parole raccolte in ciascun gruppo

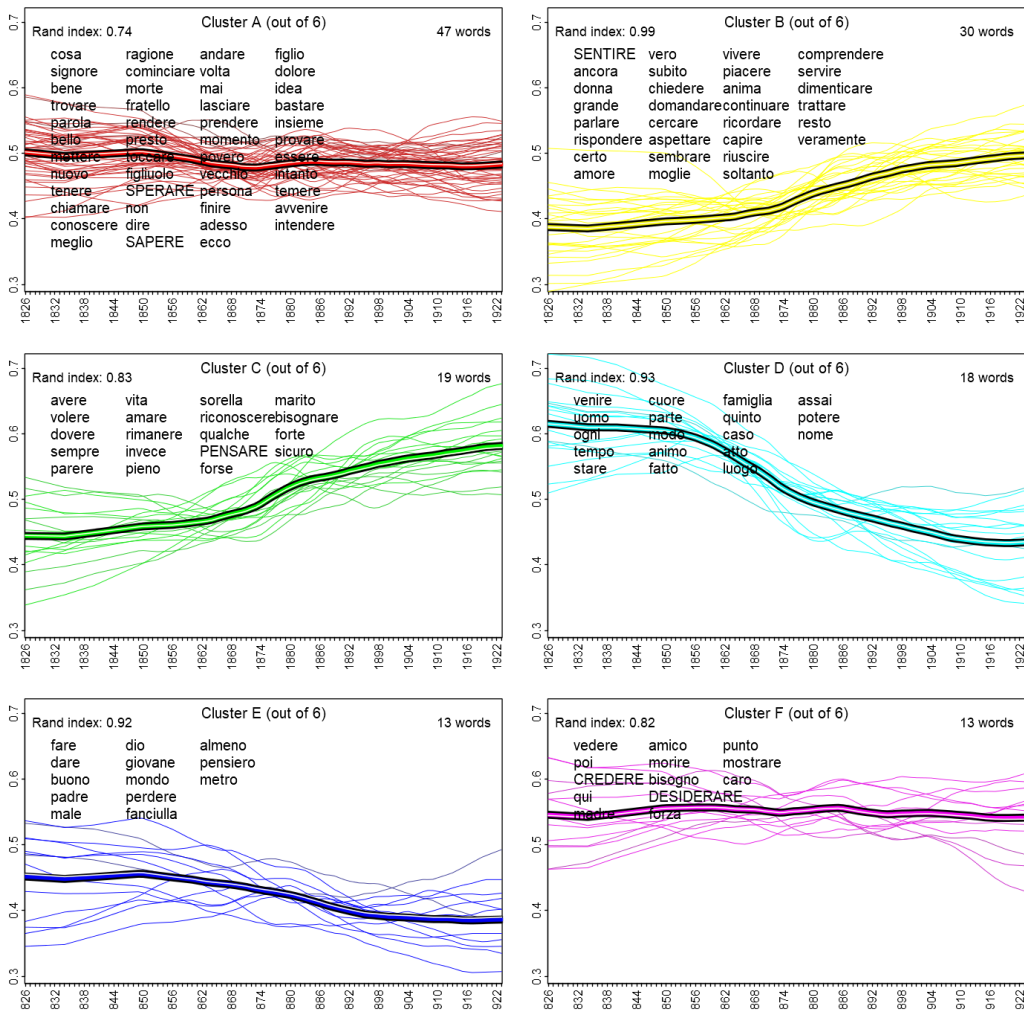


Figura 5 – Rappresentazione delle traiettorie delle parole inserite dalla procedura nei 6 gruppi. La curva evidenziata rappresenta la tendenza media del gruppo

## 6. CONSIDERAZIONI CONCLUSIVE

L'idea di studiare con prospettiva diacronica la presenza nelle opere dell'espressione interiore degli autori e dei loro personaggi è sicuramente molto ambiziosa e questo primo esperimento mostra alcune potenzialità degli strumenti statistici impiegati, ma anche alcuni limiti dei criteri adottati, sia per la costituzione del corpus sia per la scelta dei fenomeni linguistici da osservare.



Sebbene non si possa rigettare l'ipotesi che l'espressione interiore abbia assunto maggiore importanza nel corso dell'Ottocento, le analisi statistiche dei dati testuali non evidenziano con chiarezza le tendenze prospettate. I metodi statistici impiegati per l'analisi e il raggruppamento delle traiettorie disegnate nel tempo dalle occorrenze normalizzate evidenziano tendenze temporali deboli e la gran parte dei verbi seme e delle parole selezionate individualmente mostra, dopo il lisciamento, un andamento poco differenziato nel corso del tempo e solo dall'analisi dei clusters ci vengono alcune suggestioni riguardo a possibili tendenze in aggregato. In questo senso lo strumento statistico riesce solo parzialmente a essere d'aiuto e la riflessione deve essere rivolta più alla fase di selezione del corpus e di costruzione dei dati testuali che a quella di elaborazione. Nonostante le problematiche emerse, alcune timide tendenze, specie a livello di *cluster* e nel contesto specifico della costellazione di espressioni del sentire, dell'interiorità, dei modelli e dei ruoli sociali, delle relazioni e dei legami familiari, permettono di cogliere segnali di crescita e diminuzione, evidenziando globalmente le potenzialità dei metodi utilizzati.

Probabilmente la scelta di partire da sei verbi-seme, che forse nella lingua italiana sono troppo polisemici per essere delle buone *proxy* dell'espressione interiore, e da un insieme limitato di parole semanticamente legate a questi sei verbi-seme non porta a cogliere la complessità del fenomeno che si desidera indagare. La presenza nei testi dell'espressione interiore potrebbe essere effettivamente aumentata come suggerisce l'ipotesi, ma questa tendenza non può essere visibile nella frequenza delle parole se, parallelamente, sono cambiate nel corso del tempo le strategie discorsive per rappresentarla nel testo. Per esempio, nel testo narrativo moderno potrebbero essere frequenti monologhi interiori senza che siano presenti veri e propri marcatori (*credo che, lei pensò che*) che, viceversa, erano più presenti in passato.

Il metodo adottato per selezionare l'insieme di parole (*word embedding*) ha sicuramente grandi potenzialità ma si scontra con alcuni problemi tecnici, che riguardano sia la dimensione limitata del corpus (100 opere) sia l'addestramento dell'algoritmo. Con riferimento alla dimensione del corpus, si specifica che, nonostante il numero di *token* appaia elevato, generalmente l'addestramento dei *word embedding* utilizza corpora molto più estesi (solitamente almeno 5-10 volte la dimensione del corpus analizzato in termini di occorrenze) e questo può aver limitato la capacità di cogliere la semantica presente nei testi. In mancanza di un corpus molto esteso, sono disponibili anche vettori pre-addestrati su grandi corpora, come i contenuti di Wikipedia e di Twitter. Esistono in questo senso *word embedding* specifici per la lingua italiana, ma si è deciso di usare lo stesso corpus che abbiamo analizzato per l'addestramento poiché il contesto (pagine web e social media) in cui sono stati creati questi vettori pre-addestrati non appare coerente per analizzare testi narrativi dell'Ottocento. Anche per il processo di lemmatizzazione valgono considerazioni simili, in quanto le *tree-bank* disponibili non sono state create sulla base di testi di narrativa, ma prin-

cialmente a partire da articoli di quotidiani e pagine di Wikipedia come nel caso di ISDT.

L'ultima considerazione riguarda i criteri adottati per la costituzione del corpus utilizzato rispetto agli obiettivi perseguiti dal progetto europeo. Se messo a confronto con il caso dell'analisi di corpora di testi di genere molto diverso (raccolte di messaggi pubblicati nei social network, articoli di giornale, articoli scientifici, ecc.) sembra più difficile identificare tendenze di fondo nelle frequenze sperimentate dalle parole nelle opere di narrativa. Si tratta, infatti, di testi più dispersivi, che includono contesti molto differenti e trattano temi in maniera più diluita, incidendo in tal modo sulla probabilità di associazione tra le parole su cui si basa il metodo di selezione con *word-embedding*. Avere a disposizione 100 opere significa sicuramente lavorare con un corpus di dimensioni di tutto rispetto. Tuttavia, queste dimensioni risultano insufficienti se le opere devono diventare la base di studi cronologici su un arco temporale molto ampio (un secolo). Nel caso del corpus esaminato, le 100 opere si collocano su 56 punti-anno diversi e, di conseguenza, ogni anno di osservazione è rappresentato da una sola opera o, comunque, da un numero molto ridotto di opere. Per poter rappresentare ogni punto-anno in maniera adeguata è necessario disporre di corpora di dimensioni molto più grandi e, soprattutto, costituiti da testi più numerosi e meglio distribuiti nel periodo di osservazione.



- Cortelazzo M. A. (2021) "Corpora e storia della lingua", *RITT, Rivista Internazionale di Tecnica della Traduzione*, 23, pp. 179-186.
- Desgraupes B. (2018) *clusterCrit: Clustering Indices*, R package version 1.2.8 <https://CRAN.R-project.org/package=clusterCrit>.
- Dimitriadou E. (2021) *cclust: Convex Clustering Methods and Clustering Indexes*. R package version 0.6-23. <https://CRAN.R-project.org/package=cclust>.
- Durum Wheat Varieties, *Quaderno DSEMS*, 23/2007.
- Genolini C., Alacoque X., Sentenac M. e Arnaud C. (2015) "kml and kml3d: R Packages to Cluster Longitudinal Data", *Journal of Statistical Software*, 65(4), pp. 1-34. <http://www.jstatsoft.org/v65/i04/>.
- Grilli, L., Russo, M. A. e Gismondi, R. (2012) "Methodological Proposals for a Qualitative Evaluation of Italian Durum Wheat Varieties", *Journal of Applied Economic Sciences*, 7, Issue 2 (20), summer 2012, pp. 103-122.
- Mikolov, T., Chen, K., Corrado, G. S. e Dean, J. (2013) "Efficient Estimation of Word Representations in Vector Space", in *1st International Conference on Learning Representations, ICLR 2013, Scottsdale, Arizona, USA, May 2-4, 2013, Workshop Track Proceedings*. arXiv:1301.3781.
- Onelli S. (2021) "Tempi verbali e perifrasi gerundivali in un corpus di italiano letterario (1800-2000)", *RITT, Rivista Internazionale di Tecnica della Traduzione*, 23, pp. 187-212.
- Pennington J., Socher R. e Manning C.D. (2014) "GloVe: Global Vectors for Word Representation", in *Empirical Methods in Natural Language Processing (EMNLP)*. Association for Computational Linguistics, Doha, pp. 1532-1543 <http://www.aclweb.org/anthology/D14-1162>.
- R Core Team (2021) *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.
- Ramsay J. e Silverman B. W. (2005) *Functional data analysis*, Springer series in Statistics, New York, Springer.
- Ramsay J.O., Graves S., Hooker G. (2021). *fda: Functional Data Analysis*, R package version 5.5.1. <https://CRAN.R-project.org/package=fda>.
- Sciumbata F.C., Nadalutti P., Tringali L. (2021) "Trovare lavoro' in un corpus di narrativa del XIX-XX secolo. Procedure, aspetti e problemi di creazione, estrazione e rappresentazione dei dati", *RITT, Rivista Internazionale di Tecnica della Traduzione*, 23, pp. 235-268.
- Selivanov D., Bickel M., Wang Q. (2020). *text2vec: Modern Text Mining Framework for R*. R package version 0.6. <https://CRAN.R-project.org/package=text2vec>.
- Straka M. e Straková J. (2017) "Tokenizing, POS Tagging, Lemmatizing and Parsing UD 2.0 with UDPipe", in *Proceedings of the CoNLL 2017 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies*. Association for Computational Linguistics, Vancouver, pp. 88-99.
- Trevisani M. e Tuzzi A. (2013) "Shaping the history of words", in *Methods and Applications of Quantitative Linguistics: Selected papers of the VIIIth International Conference on Quantitative Linguistics (QUALICO)*. Ed. by I. Obradović, E. Kelih e R. Köhler, Belgrade, Akademska Misao, pp. 84-95.
- Trevisani M. e Tuzzi A. (2015) "A portrait of JASA: the History of Statistics through analysis of keyword counts in an early scientific journal", *Quality and Quantity*, 49, pp. 1287-1304.

Trevisani M. e Tuzzi A. (2018)  
“Learning the evolution of  
disciplines from scientific  
literature. A functional clustering  
approach to normalized keyword  
count trajectories”, *Knowledge-based  
systems*, 146, pp. 129-141.

Tuzzi A. (2018) (ed.) *Tracing the Life  
Cycle of Ideas in the Humanities and  
Social Sciences*, Cham, Springer.

Walesiak M. & Dudek A.  
(2020) “The Choice of Variable  
Normalization Method in Cluster  
Analysis”, in *Education Excellence  
and Innovation Management: A  
2025 Vision to Sustain Economic  
Development During Global  
Challenges. Proceedings of the 35th  
International Business Information  
Management Association Conference  
(IBIMA)*, 1-2 April 2020 Seville, Spain.  
Ed. by K. S. Soliman, IBIMA, Seville,  
pp. 325-340.

Wijffels J. (2021) *udpipe:  
Tokenization, Parts of Speech Tagging,  
Lemmatization and Dependency  
Parsing with the ‘UDPipe’ ‘NLP’  
Toolkit*. R package version 0.8.8.  
[https://CRAN.R-project.org/  
package=udpipe](https://CRAN.R-project.org/package=udpipe).



# “Trovare lavoro” in un corpus di narrativa del XIX-XX secolo. Procedure, aspetti e problemi di creazione, estrazione e rappresentazione dei dati

FLORIANA CARLOTTA SCIUMBATA  
Università di Trieste  
fsciumbata@units.it

PAOLO NADALUTTI  
Gruppo Interdisciplinare di Analisi Testuale  
paolo.nadalutti@gmail.com

LUCA TRINGALI  
tringalinvent@libero.it

## ABSTRACT

This paper illustrates methods, tools, and preliminary results of a study aimed to create a list of job titles and analyze them in a corpus of 100 Italian canonical and non-canonical fictional prose works published between 1825 and 1923. Job titles are interesting because they reflect socio-economic changes, as well as giving important information on how literary settings and genres changed. After a short introduction on job titles, we will discuss tools and methods used for the creation of a list of words, data extraction, and data representation, both from a linguistic and programming point of view. Some preliminary results will be shown and discussed with a statistical approach: data do not suggest significant patterns over time: whereas job titles appear more or less consistent from a chronological point of view. Finally, some advantages and limitations will be examined. The goal of this study is to develop a set of tools and methods that can be easily reproduced to build complex lexical lists, find their items, and represent data for corpora of any genre and size in a simple and effective way.

## KEYWORDS

Job titles, fictional prose, chronological analysis, lexical extraction, automatic text analysis

## 1. INTRODUZIONE

Il presente contributo<sup>1</sup> presenta metodologie, strumenti e risultati parziali, ma anche problemi e potenzialità, di uno studio<sup>2</sup> multidisciplinare realizzato con strumenti linguistici, informatici e statistici. La ricerca ha come obiettivo la messa a punto di un sistema per l'estrazione automatica di dati riguardanti i nomi di mestieri e professioni sul corpus di narrativa, che raccoglie testi pubblicati nel secolo che va dal 1825 al 1923.<sup>3</sup> I lavori analizzati sono quindi rappresentativi di un periodo caratterizzato da profondi mutamenti, sia dal punto di vista storico sia da quello letterario, e particolarmente interessante da osservare.

L'estrazione automatica dei nomi di professioni in un corpus non è però di facile esecuzione poiché non sono disponibili liste elettroniche da impiegare per applicazioni automatiche. Anche le caratteristiche morfologiche e sociolinguistiche dell'italiano pongono alcuni ostacoli che devono essere tenuti in considerazione per perfezionare l'estrazione dei dati.

Di seguito saranno illustrate le metodologie scelte per:

- individuare le fonti e creare le liste di nomi di mestieri (§ 3);
- estrarre i dati dal corpus (§ 4);
- rappresentare i risultati (§ 5).

Le questioni saranno osservate anche dal punto di vista informatico e della programmazione (§ 6). Saranno poi commentati alcuni dati ottenuti (§ 7), che sono da considerare come risultati preliminari, qui mostrati per testare le funzionalità del sistema che abbiamo messo a punto. Infine, saranno discussi aspetti, problemi e prospettive di ricerca future (§ 8).

Per raggiungere l'obiettivo di questo studio, abbiamo scelto di combinare più metodi e di creare alcuni strumenti *ad hoc*. In particolare, per automatizzare le procedure, abbiamo messo a punto *ExSTRA* (*Extracting social Status Ranks Automatically*), cioè un insieme di diversi *script* nei linguaggi di programma-

- 1 Il lavoro di ricerca e i testi che lo illustrano sono il frutto di un approccio interdisciplinare che applica metodi linguistici, statistici e informatici e che ha visto la piena collaborazione dei tre autori sotto tutti i punti di vista. Ai soli fini dell'attribuzione, specifichiamo che Floriana Carlotta Sciumbata ha redatto i §§ 1, 2, 3, 4, 5 (con Paolo Nadalutti), 7, 7.1 e 8; Paolo Nadalutti, oltre che del § 5, si è occupato del § 7.2; Luca Tringali ha scritto il § 6.
- 2 Lo studio che descriveremo è stato avviato nell'ambito del progetto Distant Reading for European Literary History (COST Action CA16204), che ha finanziato una missione scientifica breve, intitolata "Tackling Research Questions from WG3 to WG2: analysing reported speech, exploring historical places and meeting social groups in the Italian and English ELTeC corpora", svolta da Floriana Carlotta Sciumbata presso l'Università di Fiume (Croazia) tra gennaio e marzo 2020.
- 3 Per la descrizione del corpus rimandiamo alla premessa a questo articolo e a quello di Scian-dra, Trevisani e Tuzzi.

zione *Python* e *R*,<sup>4</sup> che rappresentano una vera e propria *pipeline*, una catena di montaggio che rende le operazioni facili e veloci, ma soprattutto facilmente replicabili in altre ricerche, che possono partire da liste o *corpora* diversi da quello qui utilizzato, e addirittura da lingue diverse dall'italiano.

## 2. I NOMI DI MESTIERI E PROFESSIONI

I nomi di mestieri e professioni sono senza dubbio un dato interessante poiché sono la cartina al tornasole di cambiamenti avvenuti in ambito economico e sociale nel corso del tempo, che si riverberano anche sulla lingua. Una dimostrazione dell'importanza dell'argomento sono, ad esempio, le numerose discussioni sui nomi di professioni al femminile che animano il dibattito di studiosi, giornalisti e parlanti ormai da qualche decennio (Burr 1995; Cardinaletti e Giusti 1991; Gheno 2019; Robustelli 2012; Sabatini 1987). L'analisi di mestieri e professioni è rilevante anche per chi s'interessa di statistica e sociologia (Gallo e Scalisi 2013), ma anche di storia o di letteratura, poiché riflettono aspetti come l'evoluzione dell'ambientazione in letteratura (contrapposizione tra ambiente urbano e rurale) oppure l'andamento e la rappresentazione delle diverse classi sociali.

Dal punto di vista linguistico, i nomi di mestieri e professioni rientrano nella categoria dei nomi d'agente, etichetta utilizzata nel campo della morfologia per indicare una "persona che fa, ha fatto o farà, una o più volte, per occasione accidentale, per abitudine e/o per funzione un'azione connessa con l'elemento di base" (Lo Duca 1990: 31), dove per "elemento di base" si intende il nome o il verbo che costituiscono il lessema di base della parola derivata. I nomi di agente possono essere quindi creati a partire da nomi, verbi e altri elementi combinati con un'ampia varietà di suffissi. Perciò, costituiscono una categoria estremamente ricca e variegata che ricorre a un'ampia gamma di risorse morfologiche e lessicali, con apporti dalle lingue straniere e da quelle classiche; cambiamenti di denominazione nel corso del tempo; convivenza di definizioni uguali che fanno riferimento a lavori differenti; presenza di diverse varietà diatopiche, con geosinonimi e geomonimi (D'Achille e Grossmann 2017: 147). L'interesse per l'argomento è dimostrato anche dalla pubblicazione di numerosi repertori più o meno recenti (tra cui, a titolo esemplificativo, Barbieri 1874; Fissi 1983; Garzoni 1584; Medici 1967; Proietti 1991; riportati anche in D'Achille e Grossmann 2017).

4 Gli script, i *corpora* e i dati estratti si trovano all'indirizzo <https://github.com/flometis/ExS-TRA/> (consultato il 5/12/2021). Sempre a meri scopi di attribuzione, l'ideazione, la progettazione e la programmazione degli script in *Python* sono a cura di Luca Tringali e Floriana Carlotta Sciumbata. Lo script *R* per i grafici a barre è a cura di Luca Tringali, quello per la creazione interattiva dei grafici è a cura di Paolo Nadalutti.

### 3. CREAZIONE DELLA LISTA DI NOMI DI PROFESSIONI

Uno dei problemi in cui siamo incorsi durante lo studio è la creazione della lista da cui partire per l'estrazione dei nomi dal corpus. Nonostante la letteratura e i repertori esistenti, non ci sono liste elettroniche e sistematiche di nomi di professioni,<sup>5</sup> necessarie per l'estrazione automatica. Abbiamo quindi deciso di stilare una lista apposita per lo studio, la cui creazione ha richiesto la combinazione di diverse fonti estratte con approcci diversi.

In primo luogo, abbiamo consultato dizionari ed enciclopedie generici online per cercare di estrarre in automatico i lemmi rilevanti per la nostra ricerca. Dizionari ed enciclopedie hanno senza dubbio il vantaggio di contenere numerosi lemmi con le relative definizioni, ma si sono rivelati poco utili per la loro scarsa sistematicità e perché contengono dati non strutturati. Per esempio, nel dizionario De Mauro si trovano le seguenti definizioni:

- fornaio: nei forni, addetto alla lavorazione e cottura del pane | chi vende il pane; proprietario o gestore di un forno o di un negozio per la vendita di pane;
- insegnante: chi insegna, spec. per professione;
- medico: chi pratica la medicina avendo conseguito il titolo accademico e l'abilitazione all'esercizio della professione;
- muratore: chi è addetto alla costruzione di opere in muratura;
- oculista: medico specialista nella cura delle malattie degli occhi.

Sebbene alcune regolarità, come la specifica dell'aspetto professionale in alcuni casi e la presenza del pronome *chi* in altri, le definizioni non sono sufficientemente regolari per poter distinguere in automatico i lemmi pertinenti solo a partire dalle definizioni stesse: la creazione di regole per l'estrazione automatica sarebbe perciò estremamente dispendiosa, così come lo spoglio manuale delle voci. Negli altri dizionari sono riscontrabili problemi simili, come nel vocabolario Treccani, che presenta le seguenti definizioni per gli stessi lemmi:

- fornaio: operaio addetto alla lavorazione e cottura del pane; padrone di un forno per il pane; chi ha negozio per la vendita del pane (con o senza il forno annesso) e, talvolta, anche di altri generi alimentari;
- insegnante: chi si dedica all'insegnamento, chi esercita la professione d'insegnare;
- medico: chi professa la medicina;
- muratore: operaio addetto alla costruzione di opere murarie;
- oculista: medico specializzato nello studio e nella cura delle malattie dell'occhio.

Anche le enciclopedie sono caratterizzate da scarsa sistematicità, e le entrate riferite a mestieri e professioni sono più scarse rispetto ai dizionari. Per esempio, la ricerca nell'enciclopedia Treccani non restituisce voci per *fornai*o e *oculista*. La ricerca sulle enciclopedie è stata comunque fruttuosa: ci ha permesso di trovare una pagina dell'enciclopedia *Sapere* di DeAgostini che raccoglie proprio tutte le

5 Cfr. anche Gallo e Scalisi (2013: 43), che lamentano l'assenza di un repertorio nazionale delle professioni.

voci su mestieri e professioni,<sup>6</sup> da *abbadatore* a *zoologo*. La pagina è stata sottoposta a una procedura di *scraping*, cioè all'estrazione automatica dei suoi contenuti: il risultato è una tabella contenente 1.211 lemmi singoli, non sono quindi presenti locuzioni più complesse. Le entrate sono state modificate perché contenevano lettere accentate per indicare apertura e chiusura delle vocali toniche di ogni parola. Le lettere accentate non previste dalla grafia sono perciò state sostituite con gli equivalenti senza accenti poiché avrebbero impedito la corretta individuazione nel corpus, dove le parole sono normalmente scritte senza particolari indicazioni sulla pronuncia.

Un'altra fonte utilizzata per la realizzazione della lista è *WikiData*, cioè la banca dati alla base dei progetti collegati a *WikiMedia*, di cui fa parte anche l'enciclopedia libera e collaborativa *Wikipedia*. *WikiData* presenta diversi vantaggi: oltre ad avere una licenza di pubblico dominio che la rende liberamente utilizzabile, contiene dati strutturati, cioè organizzati per campi predefiniti, che facilitano l'estrazione automatica delle informazioni. Infine, *WikiData* è un database multilingue e ogni voce è collegata ai suoi corrispettivi in tutte le traduzioni disponibili. Tale caratteristica è particolarmente utile perché rende flessibile e replicabile anche in altre lingue il sistema utilizzato in questo studio, che, in prospettiva, si inserisce in un progetto internazionale.<sup>7</sup>

La banca dati contiene una sezione chiamata "Profession", usata per categorizzare le voci riferite a persone di *Wikipedia*. Anche le informazioni di *WikiData* sono state estratte in automatico e inserite in una tabella simile a quella già descritta per *Sapere*, che contiene 1.969 voci composte da lemmi (es. *custode, muratore, sarto*), locuzioni complesse (es. *insegnante di inglese, data scientist, musicista di strada*) ed espressioni polirematiche (es. *agente di commercio, giudice di pace*). La lista va da *abbé* a *zootecnico*.

Le due liste ottenute da *Sapere* e *WikiData* sono state combinate in un'unica tabella: dopo aver unificato le voci doppie, il risultato finale comprende 2.947 entrate, che riportano la forma base del lemma, della locuzione o della polirematica (maschile singolare nella maggior parte dei casi, femminile singolare in alcuni, come nel caso di *ballerina, massaia, mercantessa*); la fonte (*Sapere* o *WikiData*); l'URL alla pagina della voce; e infine, per alcune voci, una breve definizione data dalla fonte stessa. Le voci di *Sapere* contengono anche informazioni su genere e numero del sostantivo.

Nella lista finale si trovano sia nomi di mestieri più antichi e ormai desueti, spesso legati all'ambito del lavoro manuale (es. *buffatore, cisario, lapicida*); sia riferi-

6 <https://www.sapere.it/sapere/enciclopedia/storia-e-societ%C3%A0/economia-e-statistica/generale/mestieri-e-professioni.html> (consultato il 30/11/2021).

7 *WikiData* contiene anche versioni in alcuni dialetti, tra cui siciliano, lombardo, veneto, romagnolo eccetera, potenzialmente interessanti per analizzare autori dialettali o tentare di individuare dialettismi. Bisogna tuttavia considerare diverse limitazioni, come la disomogeneità nella grafia dei lavori in dialetto, le diverse varietà locali, e ovviamente il numero più limitato di voci, probabilmente redatte da un gruppo molto ristretto di autori.



menti a mestieri più contemporanei (*data scientist, social media manager, influencer*): ai primi contribuisce soprattutto la lista di *Sapere*, ai secondi quella di *WikiData*. Sebbene per il corpus che abbiamo preso in analisi siano senza dubbio più utili le forme più arcaiche, la lista ottenuta è piuttosto ampia e versatile e può essere utilizzata su testi di tipologia diversa e prodotti in diversi periodi. In *WikiData* sono anche presenti alcuni forestierismi non molto comuni in italiano (es. *akyn, ama, muhtasib, opričniki, yobidashi*), oltre a diversi anglicismi (oltre a quelli già citati, altri esempi sono *business analyst, character designer, controller, general manager*).

Infine, una terza fonte che abbiamo preso in considerazione è la lista di mestieri e professioni usata dall'ISTAT per le sue rilevazioni, che contiene 800 voci. Oltre a quelle generiche come *avvocati, gelatai, pellettieri*, si trovano diverse locuzioni molto complesse che fanno riferimento a definizioni tecniche (*addetti alla preparazione, alla cottura e alla vendita di cibi in fast food, tavole calde, rosticcerie ed esercizi assimilati; piastrellisti e rivestimentisti in pietra e materiali assimilati; tagliatori di tessuti*), probabilmente poco utili nell'ambito di una ricerca su un corpus di narrativa o, in generale, su testi non tecnici. Inoltre, come si può notare, questa lista contiene le forme al plurale delle parole, diversamente dalle altre due fonti, e avrebbe dovuto essere rimaneggiata per il metodo di estrazione adottato che vedremo nella sezione successiva. Per questi motivi, abbiamo deciso di escludere la lista dell'ISTAT.

#### 4. ESTRAZIONE DEI DATI

Una volta estratte le liste, gli elementi non sono stati individuati direttamente nel corpus perché ciò avrebbe determinato due problemi non trascurabili. Il primo si ricollega al fatto che le voci della lista sono forme base (maschile o femminile singolare): cercarle direttamente nei testi avrebbe inevitabilmente escluso tutte le forme al plurale. Il secondo problema riguarda i falsi positivi: si pensi alla parola “domestico”, che può essere classificata come aggettivo, ma anche come sostantivo, e solo nel secondo caso corrisponde a un mestiere.

Per ovviare a entrambi i problemi, abbiamo deciso di trattare il corpus usando il software *UdPipe 1* (Straka et al. 2016), che converte i testi in formato CoNLL-U,<sup>8</sup> cioè li lemmatizza e aggiunge informazioni sulla parte del discorso (*PoS, part of speech*), le caratteristiche morfologiche e il ruolo sintattico dei singoli *token*.<sup>9</sup>

8 Cfr. <https://universaldependencies.org/format.html> (consultato il 10/12/2021).

9 Per il training di *UDPipe*, abbiamo provato a mettere insieme quattro *corpora* di training attualmente disponibili per l'italiano: *ISDT, PARTUT, POSTWITA* e *VIT* (cfr. <https://universaldependencies.org/treebanks/it-comparison.html>); necessari a “insegnare” al sistema a riconoscere le caratteristiche dei *token*, per tentare di ovviare alla scarsità di risorse. Occorre tuttavia tenere in considerazione che tutti i *corpora* di training sono basati su testi non letterari e in italiano contemporaneo: potrebbero perciò presentarsi diversi errori nella lemmatizzazione e nella classificazione delle caratteristiche linguistiche.

Abbiamo quindi elaborato uno *script* che, dopo aver convertito il testo in formato CoNLL-U, cerca i nomi di mestieri solo tra i *token* lemmatizzati e classificati come sostantivi, così da ottenere risultati più precisi e pertinenti.<sup>10</sup> Per avere un'idea sui risultati, riprendiamo l'esempio di *domestico*, presente nella nostra lista di mestieri. Le parole lemmatizzate come *domestico* occorrono 516 volte: 211 sono classificate dal *PoS tagger* come aggettivi, una come avverbio (dal *token domesticamente*), una erroneamente indicata come nome proprio e una non classificata, e infine 302 come sostantivi. Per il nostro conteggio di nomi di mestieri abbiamo preso in considerazione solo questi ultimi, riducendo quindi il numero di falsi positivi.

Lo *script* che abbiamo realizzato, che si occupa di individuare i lemmi nelle opere che compongono il corpus, restituisce sia una tabella delle occorrenze dei lemmi per singolo testo, sia una tabella riassuntiva con tutti i dati estratti. Per ogni lemma, oltre al numero di occorrenze, sono indicati anche i dati del testo (file, autore, titolo, sesso dell'autore, anno di pubblicazione, decennio, lista di provenienza). Tali indicazioni permettono di creare filtri complessi per l'estrazione dei dati: le occorrenze possono infatti essere calcolate per periodo, ma anche per singolo autore o in base al sesso. È inoltre possibile affiancare altre categorie, come il genere letterario, la provenienza geografica dell'autore e così via in un'ulteriore tabella, utile a incrociare diversi dati e quindi a ottenere informazioni complesse.

## 5. RAPPRESENTAZIONE

Per la rappresentazione dei dati, abbiamo creato due sistemi. Il primo è uno *script* in R che permette di ottenere grafici in modo automatico dalle tabelle con le occorrenze. Tali diagrammi a barre consentono in modo rapido e intuitivo di dare rappresentazione del numero di occorrenze nel corpus di ognuna delle forme grafiche individuate.

Un secondo metodo di rappresentazione è invece dinamico e interattivo, cioè l'utente può scegliere i dati da confrontare e il tipo di analisi da applicare. Il sistema è basato sull'analisi delle corrispondenze (Greenacre e Blasius 1994) e consente di cogliere in un unico grafico il sovra- o sotto-utilizzo delle forme grafiche all'interno dei singoli testi.<sup>11</sup>

10 Per questo studio, abbiamo limitato la ricerca solo alle entrate della lista di nomi di mestieri composte da lemmi singoli, ma è allo studio un filtro per individuare anche locuzioni complesse e forme polirematiche.

11 Rimandiamo al § 7.2 per un approfondimento sul funzionamento del sistema dinamico.

## 6. DAL PUNTO DI VISTA INFORMATICO

In questa sezione rivedremo dal punto di vista informatico alcune delle questioni già illustrate in precedenza, approfondendo in particolare alcuni aspetti della programmazione delle diverse funzioni di *ExSTRA*. Tali questioni risultano interessanti sia perché hanno richiesto l'applicazione di diversi approcci sia per le loro potenzialità anche in vista di applicazioni a materiale e ambiti diversi.

### 6.1. LA PULIZIA DEI TESTI

Dal punto di vista della programmazione, il progetto *ExSTRA* si è rivelato complessivamente piuttosto semplice: ottenuta la lista dei lemmi da cercare (il vocabolario) e la versione lemmatizzata dell'intero corpus, si tratta solo di eseguire un ciclo su tutti i lemmi del vocabolario per cercarli uno a uno all'interno dei testi. Come già accennato, per risolvere il problema della lemmatizzazione abbiamo deciso di ricorrere a *UdPipe*, uno strumento basato su un algoritmo di intelligenza artificiale. Questo perché scrivere un algoritmo che racchiuda tutte le possibili regole di lemmatizzazione di una lingua complicata come l'italiano avrebbe comportato un consumo di tempo notevole, per poi probabilmente terminare comunque con una percentuale di affidabilità non troppo lontana da quella di *UdPipe* (tra il 95% e il 97% con i principali modelli per l'italiano).<sup>12</sup> Tra i vari *tagger* presi in considerazione, quando abbiamo iniziato ad approcciarci al problema, *UdPipe* era l'unico che soddisfacesse facilmente i requisiti: avere modelli sufficienti per la lingua italiana, supportare un buon numero di lingue europee, essere sufficientemente leggero da poter funzionare anche su macchine dotate di poche risorse, e quanto più semplice possibile da integrare in *Python*.

L'unica condizione da rispettare, per utilizzare *UdPipe* o un qualsiasi altro *tagger*, consiste nel ripulire i testi. Molti dei testi, infatti, risultavano "sporchi", chiaramente scansionati da libri cartacei e tradotti in testo digitale da software OCR. Pochi erano i veri ebook, nati come testi digitali e dunque perfettamente puliti. La maggioranza dei testi aveva, quindi, gravi problemi di disallineamento tra colonne ed errori nell'impaginazione, per esempio: capitava spesso di ritrovare il numero di pagina all'interno del testo stesso, mentre è ovvio che per un'analisi in cui si è interessati solo al contenuto di un'opera la numerazione delle pagine e dei paragrafi è irrilevante o, peggio, controproducente. A causa di questi errori dell'OCR e della successiva rimozione degli invii a capo non desiderati (dovuti soltanto ai vincoli spaziali della pagina cartacea), poteva capitare che una parola venisse spezzata nel mezzo dal numero di una pagina o di un capitolo. Se per un umano è ovvio che "gua42rdare" sia la parola "guardare" interrotta dal numero di

12 Cfr. i dati riportati sul manuale del software: <https://github.com/ufal/udpipe/blob/master/MANUAL> (consultato il 10/12/2021).

pagina 42, per *UdPipe* si tratterebbe di un lemma non riconoscibile. Seppure alcuni errori possano essere rimasti, un controllo a campione ci permette di stabilire che la maggioranza di questi problemi è stata identificata con una serie di semplici espressioni regolari e corretta automaticamente. Abbiamo sviluppato un apposito *script* che lancia tutte le espressioni regolari richieste, scritte per pulire gli errori più comuni, in sequenza sui vari testi, in modo da rendere la procedura applicabile anche nel caso si voglia lavorare su un corpus differente.

## 6.2. LO SCRAPING DI SAPERE.IT

L'aspetto più interessante, dal punto di vista informatico, è legato alla creazione del vocabolario di *ExSTRA*, la lista dei mestieri. Abbiamo deciso di sviluppare una metodologia facilmente adattabile ad altri contesti, quindi il vocabolario è realizzato come una tabella che per ogni lemma contiene anche una serie di altre informazioni. Il lemma non è univoco: è infatti logico pensare che, avendo diverse fonti, molti lemmi siano presenti in ciascuna di esse. Inoltre, è possibile che uno stesso lemma appaia due volte perché dispone di significati differenti. Per questo motivo abbiamo optato per l'inserimento di un'altra colonna con la sorgente del lemma e l'eventuale descrizione: questa non è disponibile per tutti i lemmi, ma può aiutare a distinguere i lemmi duplicati durante un controllo manuale. Le due fonti di lemmi attualmente implementate sono l'enciclopedia *Sapere* e la base dati di *WikiData*. Abbiamo deciso di partire da una lista di lemmi estratta dal sito web dell'enciclopedia *Sapere* per avere una fonte autorevole: in questo modo chi desidera cercare soltanto lemmi "sicuri" può filtrare il vocabolario in base alla fonte dei lemmi e scegliere soltanto questi. Questa enciclopedia non è pensata per una ricerca programmata: non offre alcun tipo di API di ricerca e le informazioni devono essere estrapolate dalle pagine HTML. Abbiamo dunque dovuto sviluppare un algoritmo di *scraping*, cioè di estrazione automatica dei contenuti, per leggere il codice HTML di ciascuna pagina e estrarre le informazioni sui lemmi in base agli elementi del DOM.<sup>13</sup> Per fortuna, il sito *sapere.it* è almeno stato progettato per fornire *Rich Snippets*<sup>14</sup> ai motori di ricerca: ciò significa che è possibile identificare il lemma grazie al tag *title* e la descrizione grazie al tag meta chiamato *description*, secondo il fac simile

```
<title="lemma">
<meta content="Testo della descrizione" name="description">
```

13 Il DOM è l'insieme dei vari tag HTML che compongono la pagina.

14 I *Rich Snippets* sono un formato standard di tag riconosciuti da tutti i motori di ricerca per estrarre le informazioni principali da una pagina web. Sono in sostanza l'anteprima che il motore di ricerca mostra della pagina web. Cfr. <https://backlinko.com/hub/seo/snippets> (consultato il 10/12/2021).

Questo ha facilitato l'estrazione delle informazioni: è stato sufficiente scorrere tutta la homepage della categoria delle professioni (nella quale sono presenti i link ai vari lemmi censiti) per ottenere gli URL delle varie pagine nelle quali trovare gli *snippet* da estrarre. Naturalmente, lo svantaggio dello *scraping* è che, se il sito web venisse modificato, sarebbe necessario rifare da capo l'algoritmo. Tuttavia, un'enciclopedia tradizionale come Sapere è da considerarsi "immutabile", visto che gli aggiornamenti sono minimi e sarebbe più facile aggiungere manualmente eventuali nuovi lemmi, quindi il problema non sarebbe troppo difficile da risolvere.

### 6.3. WIKIDATA: UNA FONTE APERTA

Per avere una fonte più flessibile, multilingue, e in continuo aggiornamento, abbiamo optato per l'utilizzo di *WikiData*. È necessario fare una precisazione: *WikiData* ha una propria redazione e un meccanismo di controllo e revisione più rigido rispetto a *Wikipedia*, quindi, seppure sia sempre possibile un errore, la si può considerare una fonte affidabile di informazioni, soprattutto per lingue molto diffuse e con molti autori.<sup>15</sup> Il vantaggio di *WikiData* è che, essendo una sorgente di dati strutturati, è possibile fare ricerche mirate e incrociare i dati per aumentare l'affidabilità dei risultati. Inoltre, è possibile migliorare l'affidabilità dei risultati controllandoli attraverso altre banche dati. Per esempio, le principali voci di *WikiData* hanno un riferimento all'ID del thesaurus della Biblioteca Nazionale Centrale di Firenze,<sup>16</sup> che potremmo usare in futuro per controllare e relazionare i lemmi.

Il database di *WikiData* può essere interrogato tramite *query* SPARQL, grazie a comode API HTTP.<sup>17</sup> Occorre tuttavia impostare in modo ottimale la *query* per evitare di far finire la richiesta in *timeout*. La logica di *WikiData* è facile da comprendere leggendo la scheda di una qualsiasi entità: tutti gli oggetti censiti in *WikiData* sono infatti chiamati *entità*, e ogni entità ha anche *proprietà*, tutte accessibili con codici univoci. Per esempio, l'entità <https://www.wikidata.org/wiki/Q901> (*scienziato*) ha come proprietà P279 ("sottoclasse di") le altre entità *ricercatore* e *erudito* e come proprietà P31 ("istanza di") l'entità *professione*. Questo perché fare lo *scienziato* è una *professione* svolta dalle *persone erudite* o, se vogliamo vedere la questione da un'altra angolazione, uno *scienziato* è un tipo di *ricercatore*.

15 Per esempio, gli autori attivi per la lingua italiana sono mediamente 3.000, mentre per il siciliano sono 8 e per il veneto sono 15 ([https://stats.wikimedia.org/#/it.wikipedia.org/contributing/active-editors/normal|line|2-year|\(page\\_\\_type\)-content\\*non-content|monthly](https://stats.wikimedia.org/#/it.wikipedia.org/contributing/active-editors/normal|line|2-year|(page__type)-content*non-content|monthly), consultato il 10/12/2021).

16 Per esempio <https://thes.bncf.firenze.sbn.it/termine.php?id=3509> per il lemma "scienziato").

17 <https://query.wikidata.org> (consultato il 10/12/2021).

Il metodo più efficiente per ottenere le varie professioni nella lingua italiana consiste nell'eseguire una *query* *SELECT* prelevando tutti gli *item* che hanno la proprietà istanza di (P31) uguale al tipo di entità che si vuole ottenere (per esempio Q28640, che indica la *professione*). L'*item* è l'intero insieme di dati che *WikiData* possiede per quella entità, e nelle ricerche viene rappresentato dal suo *ID* univoco (i codici con lettera Q o P che abbiamo già citato). Nel caso di *ExSTRA*, tra tutti questi dati è in realtà utile principalmente la *label* (l'etichetta, cioè il lemma), e al massimo la descrizione. Queste due entità si possono ottenere come *itemLabel* e *itemDescription*. La lingua predefinita è l'inglese, ma è possibile selezionare una qualsiasi lingua tra quelle supportate da *WikiData*. La selezione della lingua è poco intuitiva, perché non basta specificare il parametro *language*: bisogna ricorrere al servizio *wikibase:language*, e dunque serve una sintassi articolata. Però nel complesso la *query* è abbastanza leggibile. Per esempio, per ottenere tutte le professioni:

```
SELECT ?item ?itemLabel ?itemDescription WHERE {
  SERVICE wikibase:label { bd:serviceParam wikibase:language "it". }
  ?item wdt:P31 wd:Q28640.
}
```

Le *label* sono interessanti perché, oltre a essere declinate in base alla lingua desiderata, possono anche essere declinate in base al genere (maschile o femminile). Attualmente nel vocabolario di *ExSTRA* sono presenti principalmente forme maschili, ma stiamo valutando l'ampliamento della lista anche con le forme femminili. Questo consentirebbe, per esempio, di capire con precisione quando compaia una forma femminile di una professione.

Non tutti i lemmi possiedono una forma femminile (proprietà P2521), quindi la lettura di questo dato deve essere opzionale, e la *query* verrebbe modificata in questo modo:

```
SELECT ?item ?itemLabel ?flabel ?itemDescription WHERE {
  OPTIONAL { ?item wdt:P2521 ?flabel . BIND(lang(?flabel) as ?lang) . FILTER(?lang="it")}
  SERVICE wikibase:label { bd:serviceParam wikibase:language "it". }
  ?item wdt:P31 wd:Q28640.
}
```

Un altro tipo di risultato già fornito dall'attuale *query* (e presente nel vocabolario *ExSTRA*) sono gli *ngrams*, cioè sequenze di parole: *WikiData* offre intere locuzioni come *professore universitario*, utili per distinguere gli ambiti (un *professore* generico da uno che insegna specificamente in un'università). Al momento ci siamo limitati alla ricerca dei lemmi singoli, ma il meccanismo che utilizziamo può già essere modificato per la ricerca degli *ngrams*.

La lista di lemmi ottenuta va ripulita, ma *WikiData* offre già una buona base di partenza. È ovvio che alcune professioni potrebbero non avere la *label* in lingua

italiana, perché la lista che si ottiene rappresenta ogni singola professione recensita su una qualsiasi edizione internazionale di Wikipedia. Possono quindi essere presenti professioni totalmente sconosciute nel nostro Paese e mai tradotte in lingua italiana. Per esempio, il lemma giapponese *ama* (le pescatrici giapponesi) esiste anche per l'italiano. Invece, il lemma portoghese *repentista* (un particolare tipo di poeta) non possiede alcuna *label* in lingua italiana perché la parola non è nota nell'ambiente italiano.

#### 6.4. UTILITÀ DELLE *query* COMPLESSE

Per questo primo studio abbiamo scelto di cercare esclusivamente le “professioni”, ma abbiamo già previsto nello *script* la possibilità di integrare per ricerche future la lista di lemmi tramite le entità “professioni storiche” (anche note come “professioni abbandonate”, Q16335296), oppure i titoli nobiliari e le particelle onorifiche (Q355567 e Q355505).

Per ottenere ulteriori risultati, è già stato implementato anche un altro meccanismo: la lista che si ottiene tramite le istanze dell'entità “professione” è infatti una lista parziale di veri e propri mestieri considerati da *WikiData* come tali. È tuttavia possibile che esistano altri mestieri particolarmente nuovi o rari, oppure che non sono propriamente dei mestieri ma che in alcuni casi potrebbero essere considerati tali. Perciò, non dispongono di una propria pagina dedicata su *WikiData*, ma possono essere identificati perché qualche persona li ha esercitati nell'arco della Storia. Per identificarli, la soluzione scelta consiste nel prendere un intervallo di anni e scorrere l'elenco di tutti gli esseri umani nati in quell'intervallo, andando a vedere cosa venga dichiarato nel loro campo *professione* (P106, *occupation*). In buona parte si tratta di mestieri già identificati tra le entità *professione*, ma con questo meccanismo è possibile identificare qualche termine *borderline*. Per esempio, il lemma “coltellinaio” non è presente come istanza di una *professione*, probabilmente perché troppo specifico, ma potrebbe tornare utile in alcuni particolari *corpora*, ed è stato rintracciato grazie a questo meccanismo perché è indicata come professione per James Black,<sup>18</sup> coltellinaio statunitense nato il 1° maggio 1800. In questo specifico esempio, la professione è stata identificata anche nell'enciclopedia *Sapere*. Invece, il mestiere di coppiere (servitore dedicato a versare il vino durante i banchetti,<sup>19</sup> non è presente nemmeno nella lista di *Sapere*, e non viene ritrovato con la semplice ricerca delle professioni su *Wikipedia* perché è troppo specifico. Può però essere scoperto cercando i mestieri svolti dagli esseri umani nati tra il 1300 e il 1400 (segnalato come mestiere per Gaston of Foix-Béarn):<sup>20</sup>

18 <https://www.wikidata.org/wiki/Q6129818> (consultato il 10/12/2021).

19 <https://www.wikidata.org/wiki/Q1075791> (consultato il 10/12/2021).

20 <https://www.wikidata.org/wiki/Q3758791> (consultato il 10/12/2021).

```

SELECT ?item ?itemLabel ?occupationLabel ?genderLabel ?citizenshipLabel ?
birth
WITH {
SELECT ?item ?gender ?occupation ?citizenship ?birth WHERE {
  ?item wdt:P31 wd:Q5; #P31 means we are looking for items that are
instances of the entity
  wdt:P21 ?gender;
  wdt:P106 ?occupation;
  wdt:P27 ?citizenship;
  wdt:P569 ?birth. hint:Prior hint:rangeSafe true.
  filter (?birth > "1300-00-00"^^xsd:dateTime && ?birth < "1400-00-
00"^^xsd:dateTime)
}
} AS %results
WHERE {
INCLUDE %results.
?item rdfs:label ?itemLabel. FILTER( LANG(?itemLabel)="it" )
?gender rdfs:label ?genderLabel. FILTER( LANG(?genderLabel)="en" )
?occupation rdfs:label ?occupationLabel. FILTER( LANG(?
occupationLabel)="it" )
?citizenship rdfs:label ?citizenshipLabel. FILTER( LANG(?
citizenshipLabel)="en" )
}

```

La soluzione scelta per evitare il *timeout* consiste nel paginare le ricerche in base all'anno di nascita, facendo quindi una diversa *query* per ogni anno di nascita delle persone. Tra l'altro, naturalmente le persone censite aumentano man mano che ci si avvicina all'età contemporanea quindi, se tra il 1300 e il 1400 si può tranquillamente fare una ricerca unica, ha senso limitare la singola *query* alle persone nate tra il 1800 e il 1801 con questo filtro per data:

```

filter (?birth > "1800-00-00"^^xsd:dateTime && ?birth < "1801-00-
00"^^xsd:dateTime)

```

Un altro meccanismo che potrebbe tornare utile per restringere il campo consisterebbe nell'applicare filtri agli esseri umani ricercati in base alle loro proprietà.<sup>21</sup> Per esempio, se si cercano specificamente professioni svolte da donne può avere senso restringere il campo ai soli esseri umani di genere femminile. È un meccanismo utilizzabile anche per fare una ricerca inversa e trovare le persone che hanno svolto una determinata professione. Per esempio, la *query*

```

SELECT ?item ?itemLabel ?occupationLabel ?genderLabel ?citizenshipLabel ?
birth
WITH {

```

<sup>21</sup> Si vedano "properties for this type": <https://www.wikidata.org/wiki/Q5#P1963> (consultato il 10/12/2021).



```

SELECT ?item ?gender ?occupation ?citizenship ?birth WHERE {
  ?item wdt:P106 wd:Q36180;
  wdt:P21 ?gender;
  wdt:P106 ?occupation;
  wdt:P27 ?citizenship;
  wdt:P569 ?birth. hint:Prior hint:rangeSafe true.
  filter (?birth > "1800-00-00"^^xsd:dateTime && ?birth < "1801-00-
00"^^xsd:dateTime)
}
} AS %results
WHERE {
INCLUDE %results.
?item rdfs:label ?itemLabel. FILTER( LANG(?itemLabel)="it" )
?gender rdfs:label ?genderLabel. FILTER( LANG(?genderLabel)="en" )
?occupation rdfs:label ?occupationLabel. FILTER( LANG(?
occupationLabel)="it" )
?citizenship rdfs:label ?citizenshipLabel. FILTER( LANG(?
citizenshipLabel)="en" )
}

```

troverebbe tutte le persone che hanno esercitato la professione di scrittore<sup>22</sup> tra chi è nato nell'anno 1800. La condizione di essere umano (Q5) è omessa perché si dà per scontato che solo gli esseri umani esercitino professioni, e quindi non serve specificare che l'entità ricercata debba anche essere una persona. Questo consente di velocizzare la *query*.

Potremmo utilizzare le *query* avanzate anche per fare una statistica delle professioni censite entro un certo periodo storico, in modo da confrontare la distribuzione “reale” dei mestieri con la distribuzione di questi mestieri nello stesso corpus. Per esempio, con la *query*

```

SELECT ?occupationLabel (COUNT ( ?occupationLabel) AS ?count)
WITH {
SELECT ?item ?gender ?occupation ?citizenship ?birth WHERE {
  ?item wdt:P31 wd:Q5; #P31 means we are looking for items that are
instances of the entity
  wdt:P21 ?gender;
  wdt:P106 ?occupation;
  wdt:P27 ?citizenship;
  wdt:P569 ?birth. hint:Prior hint:rangeSafe true.
  filter (?birth > "1820-00-00"^^xsd:dateTime && ?birth < "1821-00-
00"^^xsd:dateTime)
}
} AS %results
WHERE {
INCLUDE %results.
?item rdfs:label ?itemLabel. FILTER( LANG(?itemLabel)="it" )
?gender rdfs:label ?genderLabel. FILTER( LANG(?genderLabel)="en" )

```

22 <https://www.wikidata.org/wiki/Q36180> (consultato il 10/12/2021).

```

?occupation      rdfs:label      ?occupationLabel.      FILTER(      LANG(?
occupationLabel)="it" )
?citizenship     rdfs:label      ?citizenshipLabel.     FILTER(      LANG(?
citizenshipLabel)="en" )
} GROUP BY ?occupationLabel
ORDER BY DESC(?count) ?occupationLabel

```

si scopre che i mestieri più esercitati da persone nate nel 1820 (e dunque presumibilmente in attività dopo il 1840) censite su *WikiData* erano *politico, pittore, avvocato, scrittore, giornalista, professore universitario, ufficiale, giudice, medico, poeta*. Mentre nel corpus utilizzato per questo studio vediamo che i mestieri più citati nelle opere pubblicate nel 1840 sono *re, giudice, servitore, medico, scrittore, contadino, barbiere, mercante, deputato, e poeta*. Possiamo riassumere e visualizzare i risultati con un semplice diagramma di Venn (Figura 1):

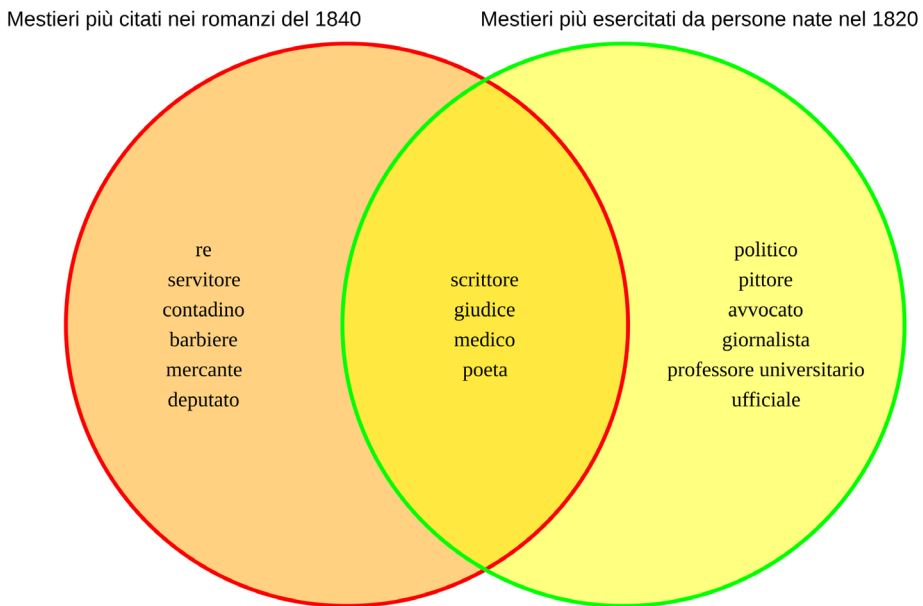


Figura 1 – Diagramma di Venn dei mestieri più citati nelle opere del corpus del 1840 e di quelli più esercitati da persone nate nel 1820 secondo *WikiData*

Chiaramente una ricerca su un singolo anno non ha rilevanza statistica, ma suggerisce che sarebbe possibile fare una classifica più ampia e capire se le opere di narrativa parlino di mestieri in voga nell'epoca in cui sono stati scritti o se tendano a parlare d'altro (e, soprattutto, dopo quanti anni dalla prima comparsa di un nuovo mestiere nella società si inizi a parlarne nelle opere letterarie). La singola *query* va eseguita su un solo anno al massimo, per evitare che sia troppo lunga: il *group by* viene eseguito dopo l'ottenimento dei risultati e non aiuta il *query planner* a trovare

una strada più rapida, quindi non riduce il tempo di esecuzione della *query*. Tuttavia, è possibile eseguire più *query* e sommare i risultati in un'unica tabella.

È chiaro che il limite di questo meccanismo di ricerca sia il fatto che le persone censite in un'enciclopedia siano solo persone in qualche modo "famoso", quindi non si troveranno molti mestieri "umili", che però possono essere facilmente identificati con il meccanismo di ricerca descritto in precedenza e attualmente impiegato per il vocabolario di *ExSTRA*, cioè le istanze dell'entità *professione*.

## 7. RISULTATI PARZIALI

In questa sezione vedremo a titolo esemplificativo alcuni risultati ottenuti grazie alle procedure che abbiamo illustrato in precedenza. Oltre a dare un'idea di come sono cambiati i mestieri nel periodo rappresentato dal corpus preso in analisi, i grafici che abbiamo realizzato permettono di testare la funzionalità e di evidenziare alcuni problemi dovuti a strumenti e metodi che abbiamo scelto di impiegare.

### 7.1. MESTIERI E PROFESSIONI PIÙ FREQUENTI

In primo luogo, osserviamo le occorrenze dei 20 lemmi riferiti a mestieri e professioni più frequenti in tutto il corpus (Tabella 1).

<b>Lemma</b>	<b>Occorrenze</b>
capo	6032
re	3276
guardia	1856
medico	1500
cavaliere	1278
maestro	1276
poeta	1220
avvocato	1056
giudice	686
lettore	651
contadino	617
stella	508
ingegnere	454
servitore	450
regina	438
cameriere	413
serva	406
artista	402
cacciatore	393
segretario	376

Tabella 1 – Lista dei 20 lemmi più frequenti del corpus

La tabella non dà informazioni sui cambiamenti temporali, ma restituisce già l'immagine preliminare dei lavori menzionati più spesso nel corpus. I risultati sono piuttosto prevedibili: si trovano infatti mestieri comuni, come *medico, maestro, avvocato, ingegnere, segretario, servitore, contadino e giudice*.

Dall'altra parte, i lemmi nella tabella evidenziano problemi dovuti alla composizione alla lista da cui sono estratti. Emergono infatti parole come *stella* e *capo*: è ovvio però che si tratta di casi non pertinenti a mestieri e professioni. Nel primo caso (esempi 1, 2, 3 e 4), i lemmi osservati in gran parte dei contesti fanno riferimento agli astri e non alle dive dello spettacolo; nel secondo (esempi 5, 6, 7 e 8), alla parte anatomica e non alla persona che comanda qualcosa.

1. "Il fiume portava seco la notte con tutte le sue stelle." (D'annunzio, *Il Fuoco*)
2. "[...] e mentre il profumo dolce della violacciocca si mesce all'odore acre dell'euforbia, le prime stelle salgono sopra il Monte." (Deledda, *Canne al vento*)
3. "E Forestina chiedèa: - Babbo, e lassù, di là dalle stelle, che c'è?" (Dossi, *La colonia felice*)
4. "Si fermò sull'ultimo limite di questo, quando non vide più dinanzi a sé che il mare bruno ed immenso, su cui scintillavano le stelle." (Verga, *Una peccatrice*)
  
5. "Egli si fece rosso, e scosse energicamente il capo." (Castellani Fantoni Benaglio, *Mia*)
6. "E il burattino fece col capo e colle mani un segno come dire: «Non ne ho»." (Collodi, *Pinocchio*)
7. "Ella rimane a capo chino." (Rovetta, *Casta diva*)
8. "L'orgogliosa e la grande Inghilterra, dominatrice di mezzo mondo, ha piegato talvolta il capo dinanzi a noi, tigri di Mompracem." (Salgari, *Alla conquista di un impero*)

Inoltre, occorre una riflessione su cosa siano un mestiere o una professione:<sup>23</sup> *stella* e *capo*, anche nel caso del calco semantico di *star* o del ruolo, sono considerabili lavori? Altri casi dubbi dello stesso tipo riguardano titoli come *re* e *regina*, che, benché rappresentino attività continuative, non sono comunemente associati a lavori, ma forse piuttosto a status sociali, oltre che a titoli nobiliari. Infine, sebbene il *lettore* possa essere "l'insegnante incaricato di esercitare praticamente gli studenti in una lingua straniera" (Vocabolario Treccani), in realtà le opere del corpus si riferiscono piuttosto al loro pubblico (esempi 9, 10, 11 e 12).

9. "C'era una volta... - Un re! - diranno subito i miei piccoli lettori" (Collodi, *Pinocchio*)
10. "Che importa a me d'avere, per esempio, cento lettori nell'isola dei Sardi ed anche dieci ad Empoli e cinque, mettiamo, ad Orvieto?" (D'Annunzio, *Il piacere*)
11. "Ricorderanno i lettori che siamo nella seconda quindicina di febbraio" (Garibaldi, *Clelia ovvero il governo dei preti*)
12. "E ora i nostri giovani lettori si trasportino con la fantasia a Collebianco [...]" (Baccini, *Il principino*)

23 Si tratta di una questione complessa che può essere affrontata da diversi punti di vista. Per esempio, per la definizione a fini sociologici e statistici, rimandiamo a Gallo e Scalisi 2013: 45-48.

Per avere un quadro<sup>24</sup> su come sono cambiati i mestieri nel corso del tempo, è possibile estrarre dati riguardanti periodi delimitati grazie ai filtri che abbiamo descritto nel § 4. I grafici in Figura 2, Figura 3 e Figura 4, che inseriamo qui a titolo esemplificativo,<sup>25</sup> contengono una rappresentazione dei 20 lemmi relativi a mestieri e professioni che occorrono più spesso nei decenni 1840-1849, 1880-1889<sup>26</sup> e 1910-1920 (e oltre), cioè il primo rappresentativo,<sup>27</sup> l'ultimo periodo, e il decennio a metà tra i due.

exstra\_dictionary\_COMPLETE 1840\_1849

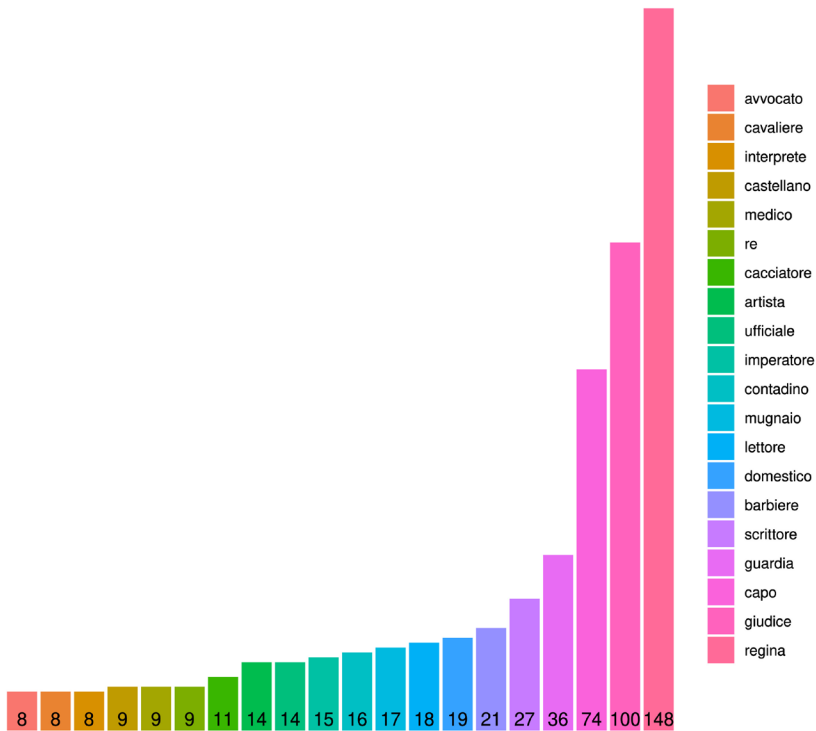


Figura 2 – Occorrenze dei 20 lemmi più frequenti nel decennio 1840-1849

24 Le rappresentazioni che seguono non sono significative dal punto di vista statistico perché non tengono conto della dimensione dei subcorpora, nella distribuzione dei lemmi nei singoli testi eccetera.

25 I grafici di tutti gli altri decenni sono disponibili sul repository di ExSTRA all'indirizzo <https://github.com/flometis/ExSTRA> (consultato il 10/12/2021).

26 Si tratta anche del decennio con il maggior numero di opere (21).

27 Nel corpus sono presenti anche opere pubblicate in precedenza, ma si tratta solo di due casi: l'analisi non sarebbe quindi rappresentativa.

Escludendo i risultati poco pertinenti di *capo* e *lettore*, che abbiamo già commentato e che compariranno anche nei due grafici successivi, gran parte dei risultati della Figura 2 rientrano sostanzialmente in tre categorie: quella lavoratrice più legata a mestieri manuali (*barbiere, domestico, mugnaio, contadino, cacciatore*), quella borghese (*giudice, medico, avvocato*), e infine compaiono titoli nobiliari (*regina, re, imperatore*) o altri titoli in qualche modo a essi connessi (*cavaliere, castellano*). *Interprete* è invece un altro caso di parola polisemica che non ha a che vedere con la professione legata all'intermediazione tra lingue diverse. Ciò emerge osservando i contesti d'uso del lemma (tutti contenuti in *Storia di una colonna infame* di Manzoni), in cui appare chiaro che il significato è piuttosto legato a colui o colei che spiega qualcosa, per lo più in ambito giuridico, come negli esempi 13 e 14:

13. “La questione dev'esser dunque, se i criminalisti interpreti (così li chiameremo, per distinguerli da quelli ch'ebbero il merito e la fortuna di sbandarli per sempre) sian venuti a render la tortura più o meno atroce di quel che fosse in mano dell'arbitrio [...]”
14. “Anzi la ragione di quelle precauzioni, la ricavavano gl'interpreti dalla legge medesima [...]”.

extra\_dictionary\_COMPLETE 1880\_1889

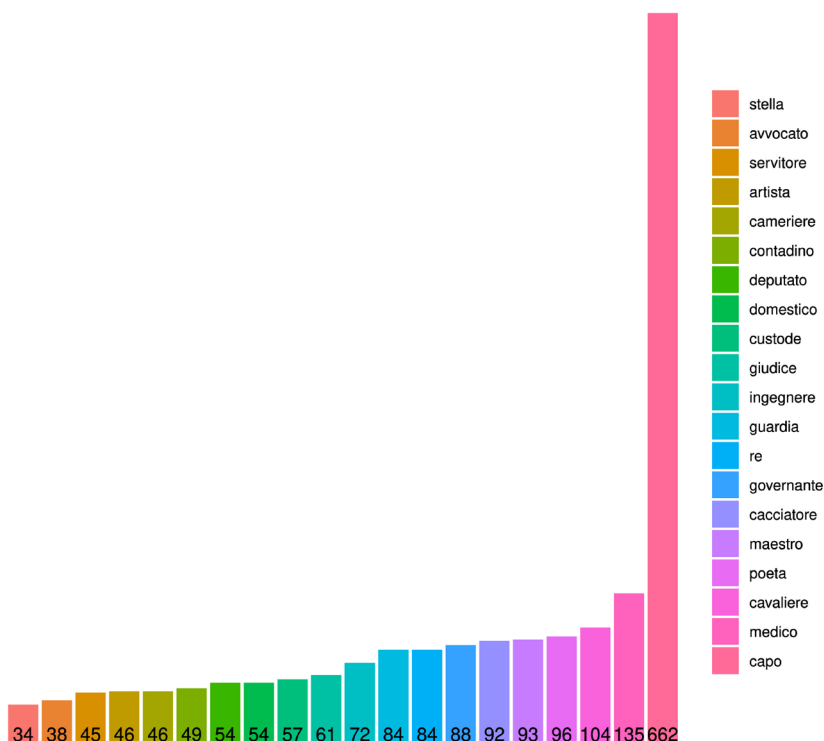


Figura 3 – Occorrenze dei 20 lemmi più frequenti nel decennio 1880-1889

La Figura 3 non evidenzia particolari cambiamenti nel decennio 1880-1889 rispetto al decennio 1840-1849: anche in questo caso si distinguono tre classi principali e, al di là della scomparsa di *regina* e della presenza di altri mestieri considerati più “umili” (*domestico*, *cameriere*) e borghesi (*medico*, *ingegnere*, *giudice*), si nota l’ingresso di mestieri collegati all’ambito artistico e culturale: oltre a *maestro*, compaiono *artista* e *poeta*.

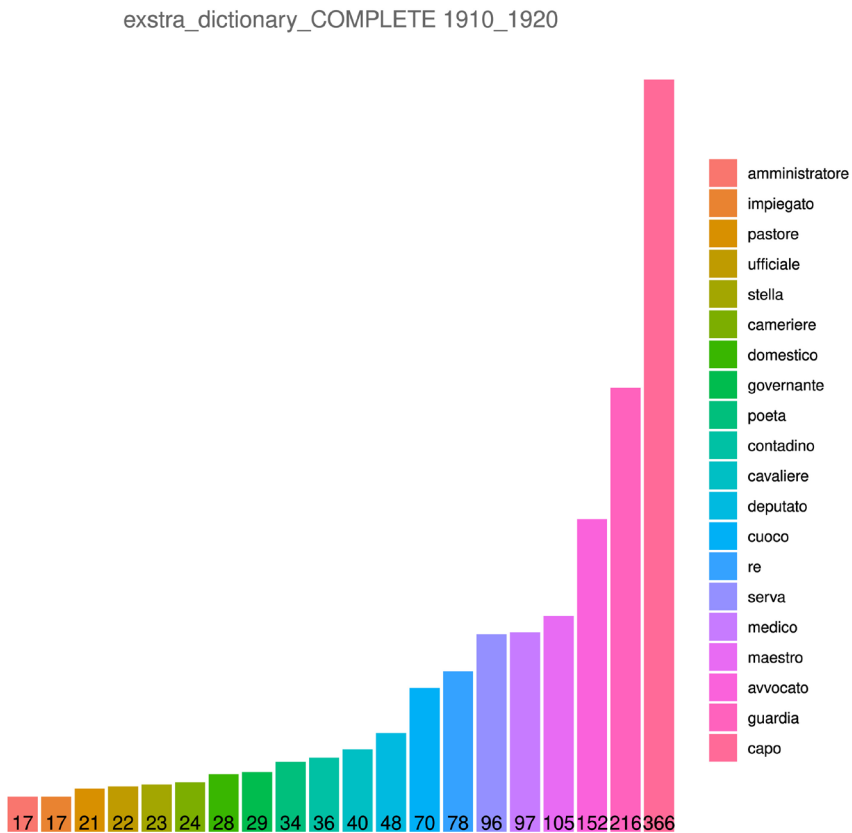


Figura 4 – Occorrenze dei 20 lemmi più frequenti nel periodo 1910-1920 (e oltre)

Anche l'ultimo periodo, rappresentato nella Figura 4, non vede grandi cambiamenti, a parte l'ascesa di *avvocato*. Lo scenario di mestieri e professioni nella narrativa italiana sembra quindi piuttosto costante nel tempo per quanto riguarda i mestieri descritti, con qualche variazione nel numero di occorrenze e nell'alternanza tra classi sociali.

Un altro filtro permette di confrontare i mestieri citati più spesso da uomini e donne (Figura 5 e Figura 6).

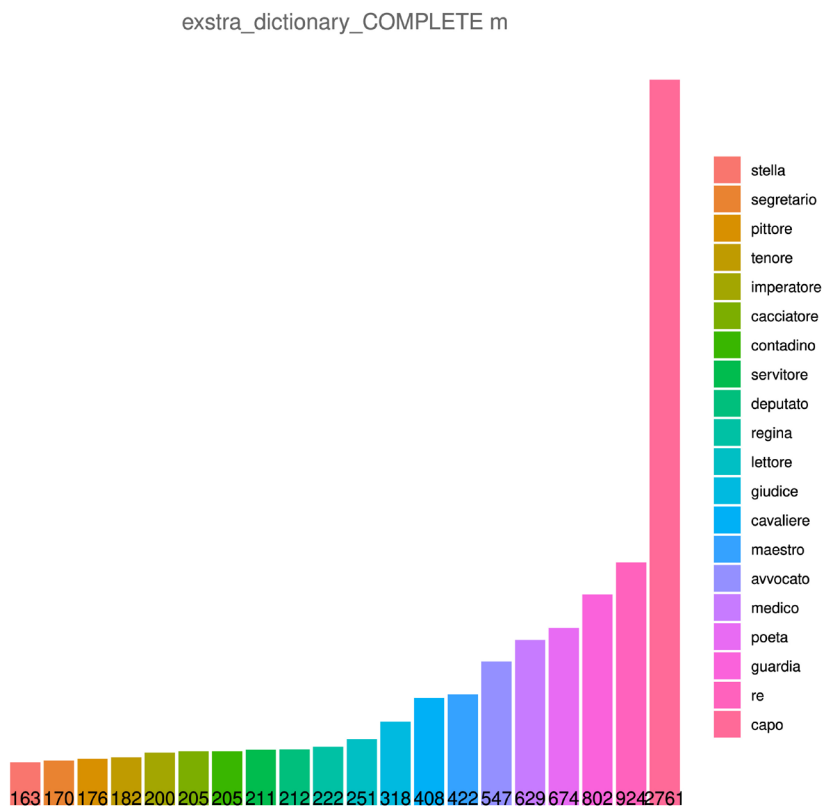


Figura 5 – Occorrenze dei 20 lemmi più frequenti nelle opere scritte da uomini



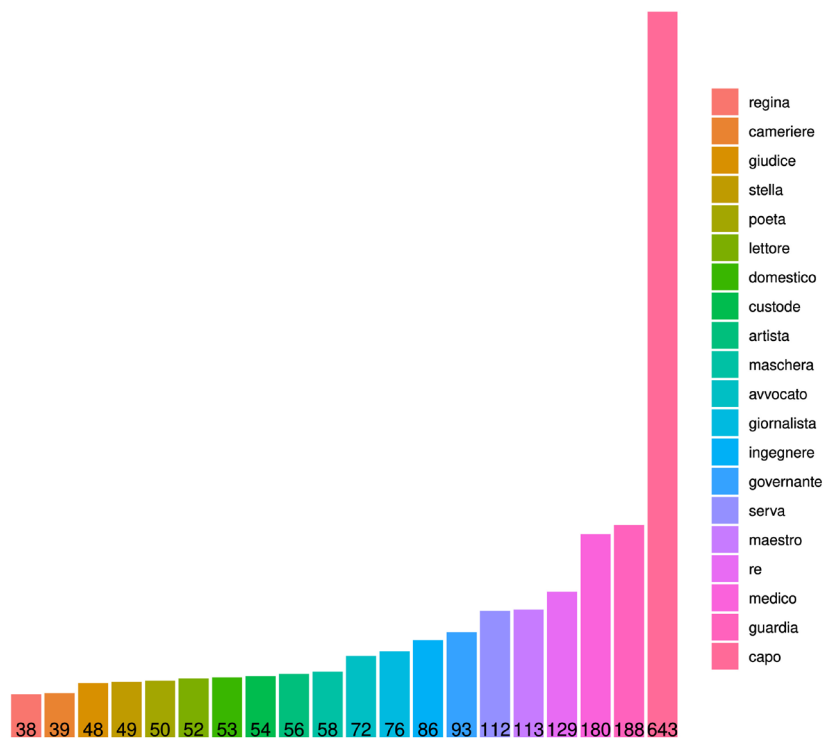


Figura 6 – Occorrenze dei 20 lemmi più frequenti nelle opere scritte da donne

Anche in questo caso non si notano particolari divergenze tra le due categorie, benché si possa osservare una maggiore tendenza nelle opere di donne a citare professioni più legate all'ambito domestico (*serva*, *governante*, *domestico*, *cameriere*). Non si tratta tuttavia di divergenze sostanziali. La parola *maschera* è evidentemente un falso positivo, poiché i lemmi non sono collegati alla figura che accompagna il pubblico in cinema e teatri, ma all'accessorio che copre il viso o ai travestimenti in generale (esempi 15, 16, 17 e 18).

15. “Lo stesso giovane che l’ultima notte di carnevale si è qui ricoverato in costume da maschera.” (Invernizio, *La trovatella di Milano*)
16. “[...] Nora aveva seguito il veemente ed appassionato sfogo del banchiere: in quel momento le appariva trasformato come se l’enorme mucchio d’oro smosso dalle sue parole lo coprisse di una prodigiosa maschera abbagliante di riflessi gialli [...]” (Di Luanto, *Per il lusso*)
17. “[...] il piccolo domino nero si voltò ai due giovani e li guardò a traverso i fori della maschera.” (Serao, *La mano tagliata*)
18. “Così nera, col viso scuro come una maschera, aveva davvero qualche cosa di diabolico [...]” (Deledda, *La madre*)

Confrontiamo infine due autori a ulteriore verifica della flessibilità dei filtri. Vediamo le differenze tra i lavori di Neera (Figura 7) e quelli di Emilio Salgari (Figura 8), un'autrice e un autore entrambi vissuti a cavallo di Otto e Novecento, ma che si occupavano di generi diversi: la prima rappresentava la condizione femminile, il secondo è invece celebre scrittore di romanzi di avventura ambientati in luoghi esotici. Nel corpus che abbiamo analizzato si trovano tre romanzi di ciascuno.

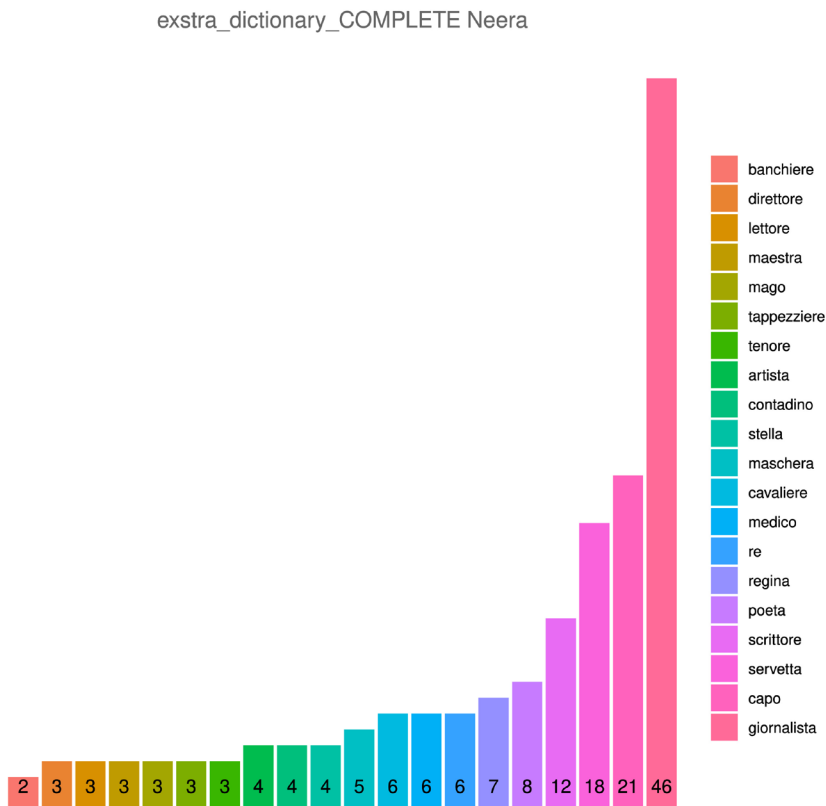


Figura 7 – Occorrenze dei 20 lemmi più frequenti nei romanzi di Neera

I mestieri citati da Neera si rifanno a diversi ambiti: quello di corte (*re, regina*); quello artistico (*scrittore, poeta, artista, tenore*); quello borghese (*giornalista, medico*); e quello legato alla classe lavoratrice (*servetta, contadino, tappezziere*). Si notano anche mestieri chiaramente al femminile, come *servetta, regina* e *maestra*.<sup>28</sup>

<sup>28</sup> Rimandiamo al § 8 per una riflessione sui mestieri al femminile nella lista che abbiamo utilizzato nel nostro studio.

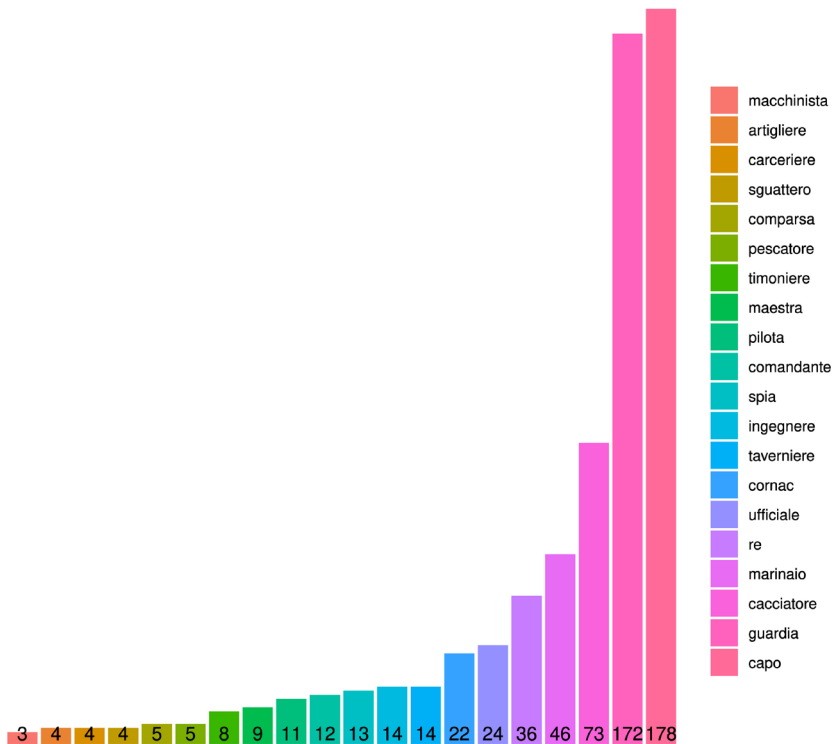


Figura 8 – Occorrenze dei 20 lemmi più frequenti nei romanzi di Emilio Salgari

In Salgari, invece, spiccano mestieri legati al mondo del mare e all'avventura (*marinaio, ufficiale, spia, comandante, pilota, timoniere, pescatore*) che ci si aspetterebbe di trovare nei romanzi dell'autore. L'ambientazione esotica si collega invece a *cornac*, parola che designa il conducente indiano di elefanti.

## 7.2. ANALISI MULTIVARIATA

L'analisi delle corrispondenze consente di individuare dei fattori impliciti che sottendono a tutti i corpus potenzialmente coinvolti nel corpus, andando perciò, ad es. a evidenziare i mestieri comuni a più testi e i mestieri che invece sono tipici unicamente di un testo solo. Intuitivamente, come si può vedere da Figura 9, il mestiere "cocchiere" è comune a *Il principino* e *La città del prossimo*, ma non compare ne *Il passaggio*. Il mestiere "poeta", invece, compare in *Il passaggio* e *La carità del prossimo*, ma non in *Il principino*. L'analisi delle corrispondenze considera le frequenze standardizzate, e dunque non è sensibile alla potenziale lunghezza dei testi (sempre intuitivamente: se un testo è molto lungo, potenzialmente

un mestiere è citato più spesso). Dunque, facendo sempre riferimento a Figura 9, possiamo ragionevolmente dire che il mestiere “cocchiere” sia citato più frequentemente in *Il principino*, rispetto quanto sia in *La carità del prossimo*, anche ipotizzando che il secondo sia (molto) più lungo del primo.

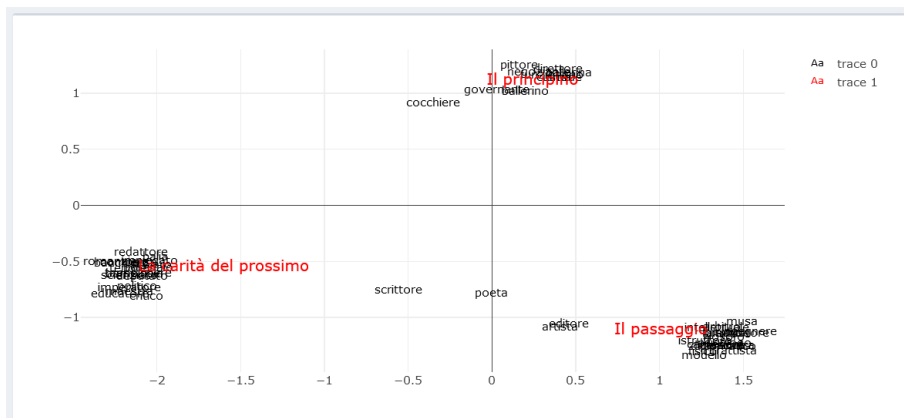


Figura 9 – Analisi delle corrispondenze per i romanzi *Il passaggio*, *Il principino* e *La carità del prossimo*

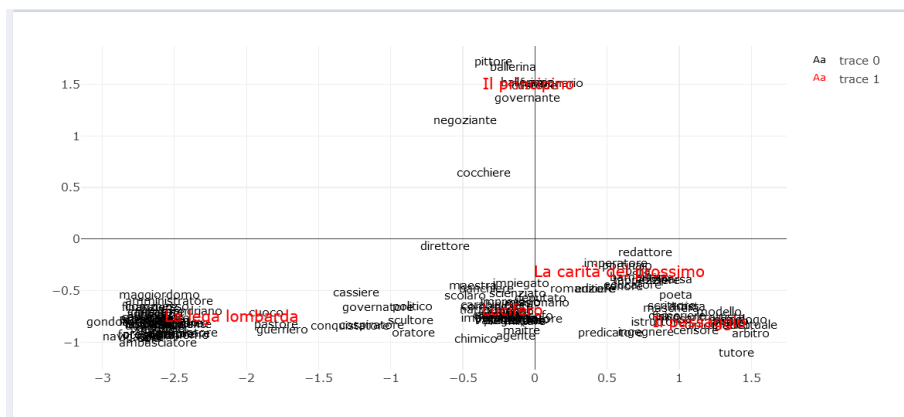


Figura 10 – Grafico delle analisi delle corrispondenze, con un numero più alto di romanzi

L’analisi delle corrispondenze consente anche di analizzare e comparare tra loro contemporaneamente più opere o subcorpora, superando i tradizionali limiti imposti da analisi tabellari o diagrammi a barre o diagrammi a dispersione. Questo consente anche di individuare opere che sono più “simili” tra loro nel nominare i vari mestieri. Come possiamo vedere in Figura 10, *Lucifero*, *La carità del*

*prossimo* e *Il passaggio* risultano essere molto simili tra loro, perché condividono un set di mestieri. La similarità tra testi può essere sempre dedotta dalla vicinanza grafica delle loro etichette: più due opere si trovano vicine tra loro nel grafico, più essi condividono un insieme di mestieri.

Nell'interpretare questi grafici bisogna sempre tenere a mente che lemmi identificati erroneamente come mestieri possono influenzare le analisi. È per questo motivo che il grafico consente di eliminare manualmente questi casi, come illustrato in Figura 11. Nel prossimo paragrafo si illustrano i comandi che consentono di interagire con i grafici.

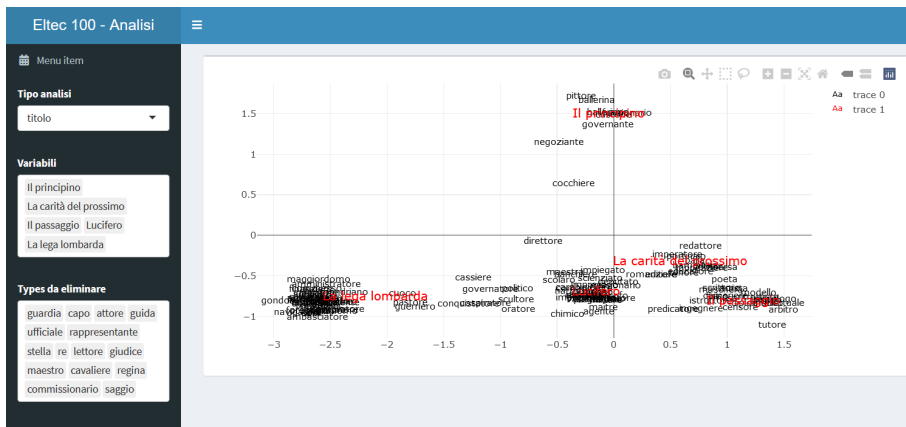


Figura 11 – Illustrazione dell'interfaccia grafica

L'interfaccia grafica consente agli utenti di interagire con diverse caratteristiche del corpus. Dal menù di navigazione a sinistra della pagina si possono selezionare diverse opzioni di visualizzazione, nella fattispecie:

- il subcorpus da analizzare (titolo, sesso, anno, decennio, autore);
- elementi da considerare (variabili, ovvero i singoli elementi che compongono il subcorpus selezionato al passo precedente);
- *types* da eliminare (quali *types* sono da eliminare dalle analisi).

Ad esempio, volendo confrontare tra loro diversi decenni, in modo da isolare i mestieri che caratterizzano gli anni tra il 1840 e 1849 dagli anni che vanno dal 1870 al 1879 e quelli che vanno dal 1900 e il 1909, dalla voce "Tipo di analisi" si dovrà selezionare "decennio". Dalla voce "Variabili" si dovrà selezionare "1840-1849", "1870-1879" e "1900-1910". La voce "Types da eliminare" consente di togliere dai grafici i mestieri identificabili come falsi positivi.

Il risultato di questa selezione è visibile in figura 12.

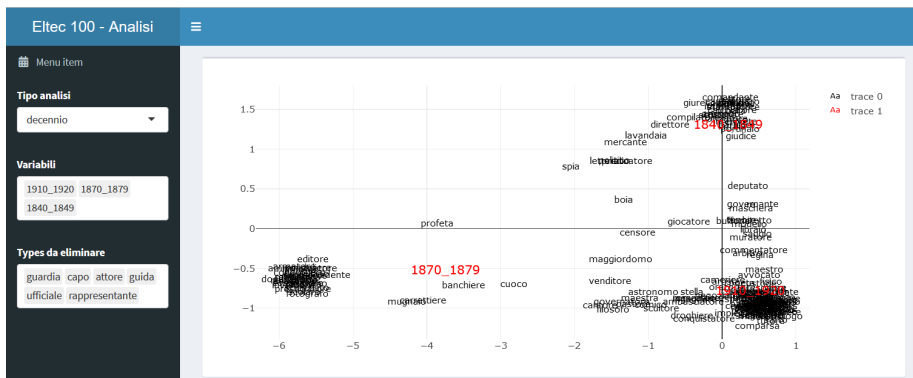


Figura 12 – Analisi per decennio

Volendo invece confrontare tra loro gli scrittori in base al loro sesso, dalla voce “Tipo di analisi” si dovrà selezionare “sesso”, mentre dalla voce “Variabili” si dovranno selezionare le voci “m” e “f”. Si veda l’esempio di figura 13. Si noti anche come, data la presenza di due uniche variabili, il grafico non rappresenti più l’esito di un’analisi delle corrispondenze, ma un semplice diagramma a dispersione delle frequenze standardizzate dei *types* nei due subcorpora. Sempre facendo riferimento a Figura 13, ad es. il *type* “avvocato” compare con una frequenza relativa di 0,004 nel subcorpus maschile, e una frequenza relativa di 0,002 nel subcorpus femminile.

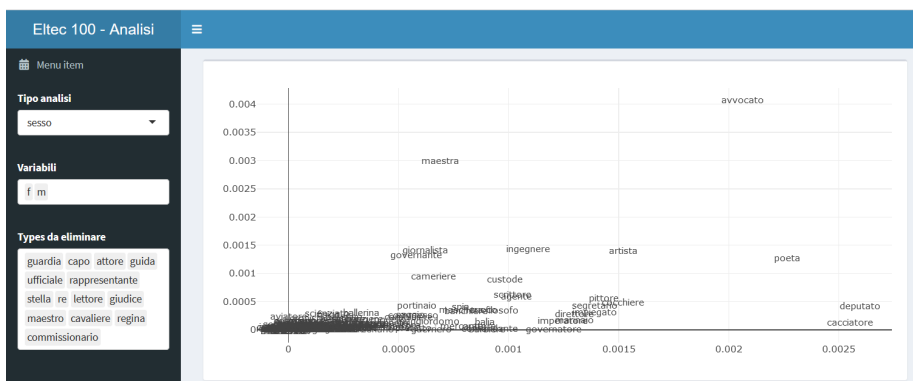


Figura 13 – Analisi per sesso

Come si può notare dalle immagini, a volte si individuano delle “nuvole di parole” così dense da risultare poco leggibili. Le “nuvole di *types*” sono caratterizzate da profili di frequenza estremamente simili tra loro, verosimilmente frequenze basse. Tali *types* a volte sarebbero perfettamente sovrapposti tra loro (si pensi agli



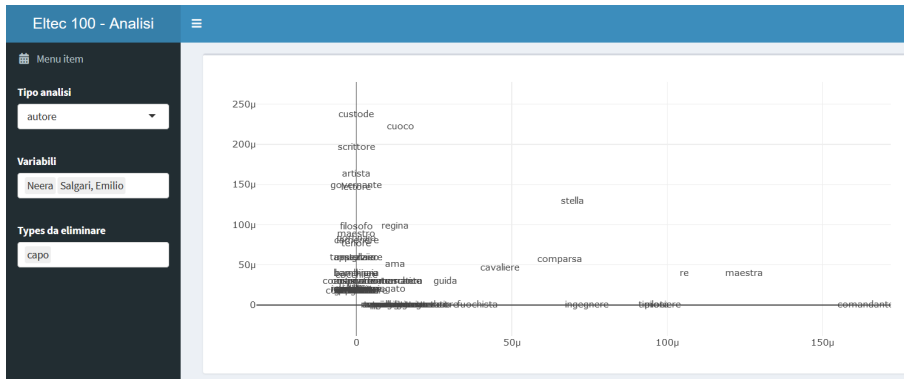


Figura 16 –Zoom sull’analisi per autore dei romanzi di Neera e Salgari

Rispetto ai dati riassunti nella Figura 8, sono ancora più chiare le differenze tra i due autori e trovano conferma alcune ipotesi già avanzate nel § 7.1. In Neera, infatti, si ripropongono i mestieri domestici (a cui si affiancano anche *cuoco* e *custode*, assenti nella Figura 8), e quelli del settore culturale, ai quali si aggiunge *filosofo*. In Salgari si affermano l’ambito navale e militare, con un’ulteriore conferma in *fuochista*, non presente nella Figura 8. Potrebbe invece sorprendere la maggiore presenza di *maestra*, che però fa riferimento all’ambito navale e non al mestiere (esempi 19 e 20):

19. “[...] attraversò quattro quattro il ponte, urtando forte col gomito un indiano che stava chiudendo il boccaporto di maestra.” (Salgari, *I misteri della giungla nera*)
20. “Si slanciarono verso i paterazzi, s’inerpicarono fino alla estremità dell’albero di maestra.” (Salgari, *I pirati della Malesia*)

## 8. CONCLUSIONI: ASPETTI, PROBLEMI, SOLUZIONI E PROSPETTIVE DI RICERCA

Dai dati preliminari su mestieri e professioni, che abbiamo individuato nel corpus per testare il nostro sistema, si può desumere che i riferimenti al lavoro presenti nelle opere prese in analisi sono prevedibili, cioè sono citati mestieri comuni, alcuni dei quali diffusi ancora oggi. Questi sono tutto sommato costanti nel periodo rappresentato dal corpus, con alcune leggere variazioni nel numero delle occorrenze. Alcune differenze emergono dal confronto tra scrittori e scrittrici, che vede la prevalenza di mestieri legati all’ambito domestico nelle opere femminili. La comparazione di singoli autori mette invece in luce divergenze più marcate: il paragone tra Neera e Salgari mostra che i lavori menzionati dall’autrice e dall’autore sono piuttosto diversi, ed evidenziano anche i diversi generi letterari in cui si inseriscono i due scrittori. Si tratta tuttavia di dati ancora parziali e grezzi che saranno ulteriormente filtrati e approfonditi, soprattutto dal punto di vista statistico, con metodi che permettano di valutare la distribuzione



nel tempo, come ad esempio il *curve clustering* utilizzato da Sciandra, Trevisani e Tuzzi (2021) in questo stesso volume.<sup>29</sup>

La trattazione e i dati che abbiamo analizzato suggeriscono che il sistema che abbiamo proposto funziona e presenta diversi vantaggi: permette infatti di estrarre e incrociare diversi dati, a partire da liste consistenti, in *corpora* anche di grandi dimensioni, impossibili da spogliare a mano, una prerogativa dei metodi automatici di estrazione dei dati testuali. I metodi che abbiamo proposto sono inoltre replicabili e le diverse fasi potranno, in futuro, essere applicate a progetti di ricerca con finalità simili ma basati su diverse tipologie testuali, per ottenere dati utili ad ampliare le conoscenze linguistiche, storiche, sociologiche, statistiche e letterarie attualmente disponibili sia in prospettiva diacronica, come nel caso di questo studio, sia in prospettiva sincronica. La struttura flessibile del sistema, il ricorso a una fonte multilingue come *WikiData* e l'uso di uno strumento come *UdPipe*, che è disponibile per un'ampia gamma di lingue per le quali sono usati sempre gli stessi sistemi di annotazione, predispongono inoltre l'applicazione a progetti che riguardano lingue diverse dall'italiano, anche a scopo di confronto, per esempio in ambito traduttivo.

Per contro, nonostante gli aspetti positivi, occorre fare alcune considerazioni e vagliare alcuni problemi metodologici e legati agli strumenti che abbiamo impiegato, che sono emersi durante la fase di progettazione del sistema e durante l'analisi dei dati. Tali problemi sono utili a contestualizzare e interpretare meglio le informazioni ottenute, ma soprattutto a migliorare il sistema in vista di prossimi utilizzi.

In primo luogo, sebbene la lista di professioni e mestieri ottenuta dalla fusione dei dati di *Sapere* e *WikiData* sia piuttosto nutrita, anche grazie all'integrazione delle due fonti, presenta senza dubbio alcuni difetti. Un problema è sicuramente rappresentato dalla presenza di falsi positivi, come *stella* o *capo*, piuttosto ricorrenti nel nostro corpus. Tuttavia, si tratta di un problema di facile soluzione: sarebbe sufficiente uno scarto manuale delle forme indesiderate (già realizzabile tramite il sistema proposto nel § 7.2), che compaiono nei risultati preliminari.

La lacuna più evidente della lista è dovuta alla prevalenza di voci al maschile singolare, che limita i lemmi al femminile ad alcuni lavori, in alcuni casi quelli storicamente condotti da donne (es. *asolaia*, *balia*, *cardatrice*, *indossatrice*, *levatrice*). La mancanza di nomi al femminile è resa ancora più problematica dal processo di lemmatizzazione. Infatti, quest'ultimo solitamente riduce i *token* alla forma base maschile e singolare, che dovrebbe corrispondere alla forma presente anche nella lista. Tuttavia, la lemmatizzazione è incostante per la natura dell'intelligenza artificiale su cui si basa: prendiamo di nuovo il caso di *domestico*, i *token* al femminile *domestica* e *domestiche* vengono lemmatizzati come *domestico* in alcuni casi, mentre in altri come *domestica*. Quest'ultima forma non figura però nella nostra lista di mestieri e professioni, quindi il lemma non viene conteggiato corretta-

29 Cfr. anche Trevisani e Tuzzi 2013; 2015; 2018.

mente. Casi simili si presentano con *mugnaia*, *squaterra*, *pastora*, *ortolana*, presenti in diversi testi, ma che il lemmatizzatore lascia alla forma femminile, assente nella lista.

L'attuale sistema di individuazione non rende inoltre possibile la distinzione tra nomi di mestieri maschili e femminili, che può però essere interessante da analizzare per individuare il momento in cui sono emersi i nomi di determinati mestieri al femminile, la loro eventuale scomparsa e la loro distribuzione del tempo. Una soluzione, che sarà oggetto di nostri studi futuri, può prevedere la creazione di un dizionario di tutte le forme del lemma (maschile, femminile, singolare e plurale) tramite sistemi automatici, come la progettazione di un sistema che interviene sulla morfologia in base alle desinenze della parola e/o che consulta dizionari online strutturati in campi, per esempio il *Wikizionario*,<sup>30</sup> un altro progetto targato *WikiMedia*, che contiene le forme flesse delle voci.<sup>31</sup> Il *Wikizionario* contiene anche una sezione di sinonimi che possono essere facilmente estratti in modo automatico: potremmo valutarne l'aggiunta alla nostra lista di lemmi. Occorre tuttavia ponderare attentamente i risultati e valutare un'attenta pulizia manuale, poiché si potrebbe aumentare l'incidenza di falsi positivi. Un approccio basato su un dizionario delle forme flesse permetterebbe anche di evitare la lemmatizzazione per ridurre gli errori a essa correlati dovuti a un corpus di *training* inadatto per quantità e tipologia dei testi.

La lista di professioni e mestieri e gli errori del lemmatizzatore causano anche un altro problema: la difficoltà di individuazione di forme marcate, innanzitutto in diacronia, particolarmente rilevanti nel nostro studio. Sebbene la lista di *Sapere* riporti molte forme antiquate, è probabile che siano assenti molti lavori non più praticati. Lo stesso vale per *WikiData*, che, come già spiegato, tende a focalizzarsi su personaggi famosi, quindi può mancare di mestieri meno recenti e più umili. Questi potrebbero eventualmente essere integrati tramite la digitalizzazione di repertori di mestieri più antichi, che, come abbiamo visto, al momento non sono disponibili in formato elettronico. Inoltre, al momento non abbiamo ancora inserito la sezione "Professioni abbandonate" di *WikiData*, che potrebbe ampliare la gamma di lavori ormai desueti attualmente a nostra disposizione.

Un ulteriore problema dovuto alla variazione nel tempo si collega alle diverse regole ortografiche dell'italiano, ancora instabili nel periodo preso in analisi. Il corpus presenta infatti grafie antiquate: un esempio è la forma *lampionajo* che si trova in *Il fu Mattia Pascal* di Pirandello, che non è stata individuata perché nella lista si trova solo la versione che rispetta le regole di scrittura attuali (*lampionaio*). Questo problema può essere risolto con un'operazione semplice (benché poco filologica) come la normalizzazione<sup>32</sup> a monte della grafia dei testi, che permet-

30 [https://it.wiktionary.org/wiki/Pagina\\_\\_principale](https://it.wiktionary.org/wiki/Pagina__principale) (consultato il 10/12/2021).

31 Si veda ad esempio la pagina relativa alla parola "scienziato" <https://it.wiktionary.org/wiki/scienziato> (consultato il 10/12/2021).

32 Si vedano anche le considerazioni di Cortelazzo (2021) in questo volume.

terebbe di ottenere un numero maggiore di risultati pertinenti. Anche le variazioni diatopiche, particolarmente comuni in italiano e tuttora vitali,<sup>33</sup> possono generare forme difficilmente individuabili. Pensiamo al caso dei diversi nomi usati in Italia per il fruttivendolo: *fruttarolo*, *fruttaiolo*, *verduraio*, *verduriere vendi-frutta*, *ortolano*, *besagnin(o)*, *erbaiolo*, *erbivendolo*, *venditore di frutta*, *quello della frutta*, *ortofrutticolo*, *ortofrutta*, *frutta e verdura*, *alimentari* (D'Achille e Gassman 2017: 147-148). Infine, le forme marcate a causa di variazioni diafasiche possono influenzare il numero di risultati: un caso emblematico può essere individuato nelle decine di sinonimi – determinati, oltre che dal registro linguistico, anche dal ricorso a gerghi, dalla variazione diatopica o dalla diacronia – relative al mestiere di prostituta, assenti nella nostra lista di lavori. Anche in questi casi, potrebbe essere utile attingere a diversi repertori già esistenti e a dizionari strutturati come il *Wikizionario* per l'integrazione di nuove forme.

Infine, la lista che abbiamo proposto per questo lavoro può essere la fonte per studi con metodi diversi, soprattutto in ambito informatico, come ad esempio il *word embedding*,<sup>34</sup> un metodo che permette di individuare in automatico parole con significato simile grazie al contesto in cui appaiono e che sarebbe quindi utile a individuare nel corpus nuovi mestieri non previsti nella lista. Anche in questo caso, tuttavia, visto il numero elevato di elementi già in nostro possesso, il rischio è di incorrere in un alto numero di risultati poco pertinenti. Sarebbe tuttavia possibile impostare dei filtri, per esempio basati su dizionari strutturati, per ottenere dati più puliti.

33 Cfr. D'Achille e Grossmann 2017: 147-148.

34 Cfr. anche Sciandra, Trevisani e Tuzzi (2021).

- Barbieri M. (1874) *Nomenclatura italiana figurata e corredata da un'appendice di 750 nomi d'arti e mestieri ad uso della gioventù e delle scuole primarie d'Italia*. Terza edizione riveduta ed aumentata, Torino, Paravia.
- Burr E. (1995) "Agentivi e sessi in un corpus di giornali italiani", in *Dialettologia al femminile. Atti del Convegno Internazionale di Studi, Sappada/Plodn (Belluno)*, 26.-30.06.1995. A cura di G. Marcato, Padova, CLUEB, pp. 349-365.
- Cardinaletti A. e Giusti G. (1991) "Il sessismo nella lingua italiana. Riflessioni sui lavori di Alma Sabatini (Sexism in the Italian Language: Reflections on the Works of Alma Sabatini)", *Rassegna Italiana di Linguistica Applicata*, XXIII, pp. 169-189.
- Cortelazzo M. A. (2021) "Corpora e storia della lingua", *RITT, Rivista Internazionale di Tecnica della Traduzione*, 23, pp. 179-186.
- D'Achille P. e Grossmann M. (2017) "I nomi dei mestieri in italiano tra diacronia e sincronia", in *Per la storia della formazione delle parole in italiano: Un nuovo corpus in rete (MIDIA) e nuove prospettive di studio*. A cura di P. D'Achille e M. Grossmann, Firenze, Franco Cesati, pp. 145-181.
- Enciclopedia Sapere DeAgostini, <https://www.sapere.it/> (consultato il 30/11/2021).
- Enciclopedia Treccani, <https://www.treccani.it/enciclopedia/> (consultato il 30/11/2021).
- Fissi A. (1983) "I nomi di mestiere a Firenze fra '500 e '600", *Studi di lessicografia italiana*, 5, pp. 53-192.
- Gallo F. e Scalisi P. (2013) "Rappresentare il lavoro che cambia. Una lettura diacronica dell'osservazione statistica delle professioni", *Sociologia del lavoro*, 129, 1, pp. 40-62.
- Garzoni T. (1584) *Piazza universale di tutte le professioni del mondo*, Venezia, Somasco.
- Greenacre M. J. e Blasius J. (1994) *Correspondence analysis in the social sciences: Recent developments and applications*, Londra-San Diego, Academic Press.
- Il Nuovo Dizionario De Mauro, <https://dizionario.internazionale.it/> (consultato il 30/11/2021).
- Lo Duca M. G. (1990) *Creatività e regole. Studio sull'acquisizione della morfologia derivativa dell'italiano*, Bologna, Il Mulino.
- Medici M. (1967) *Nuovi mestieri e nuove professioni*, Roma, Armando.
- Proietti D. (1991) *Nuovi mestieri, nuove professioni: Come orientarsi per scegliere la propria occupazione*, Roma, Sovera Multimedia.
- Robustelli C. (2012) "L'uso del genere femminile nell'italiano contemporaneo: Teoria, prassi e proposte", in *Atti della X Giornata REI (Roma, 29.11.2010) 'Politicamente o linguisticamente corretto? Maschile e femminile: usi correnti della denominazione di cariche e professioni*. A cura di M. A. Cortelazzo, Bruxelles, DG Traduzione, pp. 1-18.
- Sabatini A. (1987) *Il sessismo nella lingua italiana*, Presidenza del Consiglio dei Ministri, Direzione generale delle informazioni della editoria e della proprietà letteraria, artistica e scientifica.
- Sciandra A., Trevisani M. e Tuzzi A. (2021) "Sulle tracce dell'espressione dell'interiorità: analisi diacronica di un corpus di narrativa italiana del XIX-XX secolo", *RITT, Rivista Internazionale di Tecnica della Traduzione*, 23, pp. 219-233.
- Straka M., Hajič J. & Straková J. (2016) "UdPipe: Trainable Pipeline for Processing CoNLL-U Files Performing Tokenization, Morphological Analysis, POS Tagging and Parsing", in *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*, pp. 4290-4297. <https://aclanthology.org/L16-1680>.

Trevisani M. e Tuzzi A. (2013)  
“Shaping the history of words”,  
in *Methods and Applications of  
Quantitative Linguistics: Selected  
papers of the VIIIth International  
Conference on Quantitative  
Linguistics (QUALICO)*. Ed. by I.  
Obradović, E. Kelih & R. Köhler,  
Belgrade, Akademska Misao, pp.  
84-95.

Trevisani M. e Tuzzi A. (2015)  
“A portrait of JASA: the History  
of Statistics through analysis  
of keyword counts in an early  
scientific journal”, *Quality and  
Quantity*, 49, pp. 1287-1304.

Trevisani M. e Tuzzi A. (2018)  
“Learning the evolution of  
disciplines from scientific  
literature. A functional clustering  
approach to normalized keyword  
count trajectories”, *Knowledge-  
based systems*, 146, pp. 129-141.

Vocabolario Treccani, <https://www.treccani.it/vocabolario/> (consultato il 30/11/2021).

Wikizionario, [https://it.wiktionary.org/wiki/Pagina\\_principale](https://it.wiktionary.org/wiki/Pagina_principale) (consultato il 10/12/2021).