

Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN
TRADUZIONE, INTERPRETAZIONE E INTERCULTURALITÀ

Ciclo XXXIII

Settore Concorsuale: 10/G1 Glottologia e Linguistica

Settore Scientifico disciplinare: L-LIN/02 Didattica delle Lingue Moderne

PROGETTAZIONE E SPERIMENTAZIONE
DI UNITÀ DIDATTICHE SUL NEOSTANDARD
PER L'INSEGNAMENTO DELL'ITALIANO A STRANIERI

Presentata da: Anna Zingaro

Coordinatrice Dottorato

Raffaella Baccolini

Supervisore

Marco Mazzoleni

Co-supervisore

Adriano Ferraresi

Esame finale anno 2021

La libertà nella lingua deve venire dalla perfetta scienza e non dall'ignoranza. La quale debita e conveniente libertà manca oggigiorno in quasi tutti gli scrittori. Perché quelli che vogliono seguire la purità e l'indole e le leggi della lingua, non si portano liberalmente, anzi da schiavi. Perché non possedendola interamente e fortemente, e sempre sospettosi di offendere, vanno così legati, che pare camminino fra le uova.

Giacomo Leopardi, Zibaldone, [704] 28 febbraio 1821

RINGRAZIAMENTI

Un sincero ringraziamento va ad Adriano Ferraresi, Marco Mazzoleni, Silvia Bernardini, Michele Prandi e Gloria Bazzocchi per avermi sempre incoraggiata e sostenuta e per aver seguito con disponibilità, entusiasmo e piena fiducia nelle mie capacità l'evoluzione della ricerca in questi tre anni.

Ringrazio, inoltre, i referenti delle università estere presso le quali ho svolto la sperimentazione delle unità didattiche: Margarita Borreguero Zuloaga (Università Complutense di Madrid), María Belén Hernández González e Pablo Zamora Muñoz (Università di Murcia), Esther Morillas García (Università di Malaga) e, in particolare, ringrazio Daniel Słapek (Università Jagiellonica di Cracovia) per avermi sempre incoraggiata e sostenuta con fiducia e per la completa disponibilità dimostrata in tutte le fasi di questo percorso. Ringrazio, inoltre, Piero Conficoni e mio fratello Domenico per il prezioso supporto tecnico e informatico.

Un pensiero particolare va ai miei genitori, che non mi hanno mai fatto mancare il loro supporto e che inconsapevolmente mi hanno trasmesso la loro passione per l'insegnamento e per l'evoluzione della lingua italiana.

Grazie, infine, alle studentesse e agli studenti di Lingua Italiana L2, che dal 2009, portando i loro dubbi in classe, mi spingono ad interrogarmi costantemente su come funziona la lingua italiana e come io possa trasmetterla al meglio a loro.

RIASSUNTO

La presente ricerca consiste in uno studio sui tratti fondamentali dell'italiano contemporaneo, avente come obiettivo la creazione di unità didattiche sul neostandard, destinate ad apprendenti stranieri. Come punto di partenza della ricerca è stata assunta una lista di fenomeni grammaticali, considerati tipici della nuova varietà standard dell'italiano, elaborata da Francesco Sabatini nel 1990 nella sua definizione dell'"italiano dell'uso medio".

La ricerca è finalizzata innanzitutto a determinare la diffusione di tali fenomeni grammaticali in un corpus rappresentativo dell'italiano contemporaneo scritto e, in secondo luogo, ad aprire una riflessione su quali di questi fenomeni siano da considerarsi rilevanti nell'insegnamento dell'italiano a stranieri.

Oltre al criterio della frequenza dei fenomeni all'interno del corpus, nella scelta dei fenomeni grammaticali su cui incentrare la progettazione delle unità didattiche si terrà conto anche dei contenuti grammaticali della manualistica per stranieri attualmente in commercio, al fine di individuare i fenomeni finora esclusi dalle pubblicazioni esistenti e per i quali possa quindi essere rilevante e utile la creazione di materiali didattici.

INDICE

Introduzione.....	3
1. L'ARCHITETTURA DELL'ITALIANO CONTEMPORANEO.....	11
1.1. Brevi cenni storici	11
1.2. Gli assi di variazione	13
1.2.1. La diamesia: da asse di variazione autonomo a sottocategoria della diafasia	13
1.3. Lo spazio linguistico italiano	18
1.3.1. Italiano standard	18
1.3.2. Italiano popolare (o “dei semicolti”)	20
1.3.3. Italiano regionale.....	22
1.3.4. Italiano neostandard	23
2. APPROFONDIMENTO TEORICO SULL'ITALIANO DELL'USO MEDIO	27
2.1. L'italiano dell'uso medio	27
2.2. I tratti linguistici dell'italiano dell'uso medio	31
2.3. La lingua ritrovata.....	54
3. I CORPORA NELLA DIDATTICA DELLE LINGUE	65
3.1. Alcune nozioni di linguistica dei corpora.....	65
3.2. Metodo per l'identificazione dei fenomeni tramite corpus	70
3.3. Tipologie di corpora esistenti	71
3.4. Il web come corpus	74
3.5. Corpora nella linguistica cognitiva e nella construction grammar	76
3.6. Corpora e apprendimento delle lingue	77
3.7. Uso dei corpora per la creazione di materiali didattici.....	79
3.8. L'uso di materiali autentici.....	86
4. OBIETTIVI E METODI DELLA RICERCA	90
4.1. Domande di ricerca	90
4.2. Scelta del corpus	92
4.3. Corpora disponibili dell'italiano contemporaneo scritto.....	92
4.4. Corpus selezionato per l'indagine: PAISÀ.....	111
4.5. Confronto tra i corpora e individuazione del corpus da utilizzare	114
4.6. Progettazione delle stringhe di ricerca	116
4.6.1. Il linguaggio CQL	116
4.6.2 Elaborazione delle stringhe di ricerca	121
4.7 Metodo per il salvataggio e il trattamento dei risultati dell'interrogazione del corpus ..	139
4.8 Metodo per la progettazione e sperimentazione delle unità didattiche	141

4.8.1. Selezione dei testi.....	141
4.8.2. Valutazione di alcune piattaforme di e-learning	142
4.8.3. Progettazione delle unità didattiche	147
4.9. Questionario sulle opinioni degli studenti	153
FREQUENZA DEI FENOMENI GRAMMATICALI INDAGATI	157
5.1. Frequenza dei fenomeni indagati	157
5.2. Fattori di variazione	164
5.3. Conclusioni sulla frequenza dei fenomeni indagati	170
5.4. I fenomeni grammaticali selezionati	176
PROGETTAZIONE E SPERIMENTAZIONE DELLE UNITÀ DIDATTICHE	180
6.1. I tratti linguistici selezionati	180
6.2. Descrizione delle unità	182
6.2.1. Unità didattica sui verbi riflessivi “affettivi”	186
6.2.2. Unità didattica sul congiuntivo	195
6.3. Esiti della sperimentazione	200
6.3.1. Unità didattica sui verbi riflessivi	203
6.3.2. Congiuntivo	209
6.3.3. Esiti delle attività finali sui verbi riflessivi e sull’uso del congiuntivo vs indicativo. 213	
6.4. Esiti del questionario sulle opinioni degli studenti	224
6.4.1. Unità didattica sui verbi riflessivi “affettivi”: indagine conoscitiva dei partecipanti .224	
6.4.2. Unità didattica sui verbi riflessivi “affettivi”: opinioni sulle attività	231
6.4.3. Unità didattica sui verbi riflessivi “affettivi”: opinioni sulla piattaforma Wizer.me..238	
6.4.4. Unità didattica sul congiuntivo: indagine conoscitiva dei partecipanti.....239	
6.4.5. Unità didattica sul congiuntivo: opinioni sulle attività	243
6.4.6. Unità didattica sul congiuntivo: opinioni sulla piattaforma Wizer.me	248
6.5. Riflessioni finali sugli esiti delle sperimentazioni	249
CONCLUSIONI	254
Riferimenti bibliografici.....	265
Grammatiche	285
Sitografia di riferimento	285
Indice dei nomi	288
Appendice 1	292
Appendice 2.....	295
Appendice 3.....	305
Appendice 4.....	309

INTRODUZIONE

Questa ricerca si pone come obiettivo principale la progettazione di unità didattiche per l'insegnamento dell'italiano a stranieri che favoriscano l'approccio con i fenomeni grammaticali dell'italiano neostandard e che diano una percezione realistica della lingua utilizzata al giorno d'oggi. Il punto di partenza per la progettazione delle unità didattiche è la lista dei 14 fenomeni grammaticali identificati da Francesco Sabatini nel 1990 come caratteristici del nuovo standard linguistico emerso nel corso del Novecento e, più decisamente, nella seconda metà del secolo.

Negli ultimi due secoli la lingua italiana ha subito, infatti, un processo di trasformazione a livello morfologico, morfosintattico e lessicale. Prima dell'Unità d'Italia (1861) l'italiano era la lingua usata da un numero ristretto di persone e coincideva grosso modo con la varietà parlata in Toscana, mentre nel resto del paese era utilizzata solo nella burocrazia e dai ceti sociali più elevati.

Dopo aver assunto il ruolo di lingua nazionale, questa lingua, fino ad allora utilizzata principalmente nella scrittura di registro formale e con uno standard tendente all'aulico, dovette estendere il proprio raggio d'azione alle esigenze pratiche della comunicazione nella vita quotidiana. Fu così che, sotto l'influenza delle varietà linguistiche parlate nelle diverse regioni, l'italiano subì una graduale trasformazione che fece emergere vari tratti linguistici già esistenti da secoli e ampiamente utilizzati nel parlato, ma considerati scorretti secondo la grammatica tradizionale che codificava l'italiano scritto.

Censurati per secoli dai grammatici, ma maggiormente rispondenti alle esigenze di una comunicazione reale, parlata e scritta (cfr. D'Achille 1990), questi tratti sono giunti fino ai giorni nostri, coesistendo con le relative forme standard, e ispirando nel 1985 uno studio di Francesco Sabatini, intitolato *L'italiano dell'uso medio: una realtà tra le varietà linguistiche italiane*. Questo studio aveva l'obiettivo di elencare e descrivere i tratti linguistici caratteristici di una nuova varietà standard della lingua nazionale, che erano ormai entrati anche nello scritto di una certa formalità, sebbene alcuni di essi storicamente fossero percepiti come una deviazione dalla norma codificata dalle grammatiche. La conclusione di questo studio fu che nella scelta della varietà linguistica da trasmettere nell'insegnamento dell'italiano bisognasse tenere conto dei mutamenti subiti dalla lingua, riproponendo, così, la necessità di definire uno standard linguistico.

Pertanto, alcuni degli interrogativi che si pongono al docente di italiano per stranieri sono i seguenti: *quale* italiano insegnare? Quale varietà viene effettivamente proposta dai manuali attualmente in commercio? Come fornire un supporto a chi insegna e a chi apprende l'italiano che possa favorire una percezione realistica delle stratificazioni sociolinguistiche della lingua usata al giorno d'oggi?

Questa ricerca nasce dalla necessità di dare una risposta a queste domande, proponendo un metodo per la progettazione di unità didattiche sui tratti del neostandard, e facendo fronte in tal modo anche alla attuale carenza di materiali sul tema (cfr. Ślapek 2017). Tali necessità, che si inseriscono nel dibattito sullo standard della lingua italiana, sono emerse nella didattica della lingua italiana a stranieri presso il Dipartimento di Interpretazione e Traduzione dell'*Alma Mater Studiorum* - Università di Bologna (Campus di Forlì), dando ulteriori spunti a questo studio.

Quando una lingua viene appresa, come nel contesto appena descritto, ossia nel paese in cui è parlata dai suoi parlanti nativi, è definita “lingua seconda”, o L2, per differenziarla dalla lingua materna (L1)¹. Ciò che caratterizza il contesto di apprendimento di tipo L2 è che il contatto con la lingua oggetto di studio avviene sia in maniera guidata, tramite le attività in aula, sia in maniera spontanea, attraverso le interazioni con gli italofoni nativi e l'approccio con i testi scritti, visivi, sonori e audiovisivi eccetera prodotti nel mondo circostante. Di conseguenza, i contenuti del corso di lingua e gli stimoli provenienti dal contesto italofono si integrano a vicenda:

Ciò che caratterizza la situazione di insegnamento della lingua seconda è il fatto che le ore di didattica diretta svolta in classe dall'insegnante sono in interrelazione con la vita extrascolastica dello studente, in cui egli è esposto alla lingua viva, non strutturata e graduata, non spiegata e commentata. (Balboni 1994: 13)

Da questa citazione si possono cogliere due aspetti fondamentali: se da un lato il contatto diretto con la lingua viva ha il pregio di dare una percezione realistica delle varietà usate dai parlanti nativi, dall'altro questa situazione porta l'apprendente ad imbattersi anche in strutture linguistiche che possono risultare incomprensibili. Ciò succede perché, mentre in aula l'approccio con le strutture linguistiche sconosciute è sempre mediato e pianificato dal

¹ Quando invece una lingua non materna viene appresa in un paese diverso da quello in cui è parlata dai suoi parlanti nativi, si parla di “lingua straniera” (LS). Ad esempio, quando un italiano si trova in un paese anglofono per studiare l'inglese, lo starà studiando come “lingua seconda” (L2), mentre se studia l'inglese a scuola in Italia, lo starà studiando come “lingua straniera” (LS).

docente, che stabilisce tempi e modalità del percorso formativo, nella vita quotidiana la comunicazione avviene invece in maniera spontanea e non programmata.

Di conseguenza, l'esposizione alla lingua viva può anche causare incomprensibilità e suscitare dubbi sull'uso linguistico ("Ho sentito gli italiani dire [...]: è corretto?"), che spesso vengono portati in aula, sperando che il docente possa scioglierli. I dubbi possono derivare o dal fatto che la competenza linguistica degli apprendenti non è ancora abbastanza elevata da includere anche quella specifica struttura linguistica che crea loro difficoltà, oppure dal fatto che gli apprendenti si imbattono in strutture linguistiche per le quali non trovano riferimenti teorici nei manuali per stranieri (né, di fatto, li troverebbero nelle grammatiche destinate agli italofoeni nativi) o che in qualche modo contrastano con le norme linguistiche in essi contenute.

Quest'ultimo scenario si verifica ad esempio nel caso di strutture linguistiche che, pur essendo ampiamente utilizzate dai parlanti nativi, si collocano al di fuori della varietà standard dell'italiano e che pertanto o non compaiono nei manuali di lingua italiana L2/LS attualmente in commercio, oppure sono oggetto di una trattazione molto limitata. Il divario tra norma linguistica e uso effettivo della lingua è un tema che accompagna costantemente la lingua italiana e, data la specificità della sua storia, persiste ancora a trent'anni dagli studi di Sabatini (1985, 1990) e si riflette anche nei contenuti della manualistica attualmente in commercio per l'insegnamento della lingua italiana.

Sebbene i tratti linguistici elencati da Sabatini come costitutivi del nuovo standard linguistico siano ampiamente utilizzati dalla comunità dei parlanti (Sabatini 1990, 2016), la loro trattazione nei manuali per l'insegnamento della lingua italiana a stranieri è molto limitata, come verrà in seguito illustrato in questo lavoro. Poiché "per il docente di italiano L2 il problema fondamentale è quello di evitare l'eccessivo scollamento tra lingua utilizzata in aula e quella effettivamente presente nelle situazioni comunicative quotidiane" (Diadori, Palermo, Troncarelli 2015: 227), la presente ricerca ha come punto di partenza il divario tra l'uso linguistico effettivo e l'uso prescritto dalle grammatiche. Dato il difficile reperimento di materiali didattici di linguistica destinati (o che perlomeno risultino accessibili) agli stranieri che apprendono l'italiano, si è fatta strada l'idea di progettare da zero alcune unità didattiche sui fenomeni dell'italiano neostandard partendo dalla lista di 14 fenomeni linguistici classificati da Sabatini (1990). Come emergerà con maggiore chiarezza nell'esposizione, si ritiene dunque che uno dei tratti innovativi della presente ricerca risieda proprio nella necessità di far fronte alla carenza di materiali didattici destinati ad apprendenti stranieri sul neostandard.

Questo aspetto rappresenta il punto di intersezione, in questo studio, tra la linguistica italiana e la linguistica dei corpora, un ambito di ricerca che si è sviluppato negli ultimi anni del ventesimo secolo all'interno degli studi sull'uso della lingua e sulla variazione (Biber, Conrad, Reppen, 1998; Biber 2012). I corpora, sono delle raccolte di grandi dimensioni, di testi orali o scritti prodotti in contesti comunicativi reali (ad esempio, raccolte di articoli di giornale, libri, registrazioni di discorsi, ecc.). Tali testi sono conservati in formato elettronico e possono essere consultati tramite l'uso di strumenti informatici (cfr. Baroni 2010).

Quanto più un corpus sarà esteso e composto da fonti variegate, tanto più potrà essere considerato rappresentativo di una certa lingua. Di conseguenza, la sua consultazione permetterà di ricavare un modello del funzionamento della lingua, verificandone le tendenze generali su base quantitativa, osservando, ad esempio, quali siano le parole più utilizzate, come le parole si combinino le une alle altre e quali combinazioni di parole siano più comuni.

Uno dei metodi di ricerca principalmente utilizzati nella linguistica dei corpora è definito *corpus-based* (Tognini-Bonelli 2001), ossia basato su corpus, in cui i dati estratti da un corpus vengono utilizzati per verificare una teoria, con l'obiettivo di convalidarla, confutarla o perfezionarla. Ciò è in opposizione all'approccio di tipo induttivo definito *corpus-driven*, ossia guidato dal corpus, che verrà illustrato successivamente, nel corso dell'esposizione.

Nel contesto di questo lavoro, l'approccio *corpus-based* è quello che risulta più utile e conforme al raggiungimento degli obiettivi della ricerca. Esso consente innanzitutto di verificare la diffusione dei tratti linguistici dell'italiano dell'uso medio nei testi scritti di media formalità degli ultimi vent'anni e, di conseguenza, consente di avere a disposizione un numero elevato di testi che contengono i tratti grammaticali oggetto di studio, da poter poi utilizzare per la progettazione da zero di unità didattiche. Questa scelta si inserisce nel filone di ricerca che prevede l'applicazione dei corpora alla didattica delle lingue.

Dopo aver accennato agli ambiti da cui ha tratto spunto questa ricerca, ossia la storia della lingua italiana e le esigenze didattiche emerse nell'insegnamento, e alla metodologia adottata, di seguito verrà illustrata l'articolazione di questa trattazione in due parti.

La prima parte (capitoli 1-3) fornisce un inquadramento teorico di linguistica italiana e linguistica dei corpora. Nel primo capitolo, dopo aver dato alcuni brevi cenni sulla storia della lingua italiana (D'Achille 2003, De Mauro 2011, Dardano 2020 a c. di, Frosini 2020 a c. di, Palermo 2020), verrà tracciata una panoramica delle varietà linguistiche, dando

particolar rilievo all'italiano dell'uso medio / neostandard (Sabatini 1985, 1990, Berruto 1987).

Quest'ultimo tema sarà approfondito nel secondo capitolo, in cui verranno esaminati nel dettaglio i fenomeni grammaticali che caratterizzano questa varietà. La descrizione sarà integrata sia con gli studi sui singoli fenomeni (si citano a titolo di esempio Simone 1993, Cortelazzo 2000a, Tavoni 2002, D'Achille, Proietti, Viviani 2005, D'Achille 2010, Renzi 2012), sia con gli studi che hanno fornito un contributo fondamentale nel dibattito sulla ristrutturazione dello standard linguistico italiano (ad esempio: Sabatini 1985, 1990, 1997, 2016, Berruto 2012 [1987], Nencioni 1987, Castellani 1991, Sobrero 1992, Dardano 1994, Lorenzetti 2002, D'Achille 2003, 2016, Tesi 2005, Renzi 2012, Ondelli, Romanini 2018, Prandi 2020). Come già affermato, la necessità di definire un modello linguistico di riferimento è un tema centrale in questa ricerca, poiché ha delle implicazioni nella scelta della varietà linguistica da trasmettere nell'insegnamento dell'italiano. A ciò si accompagna la necessità di reperire dei testi che permettano agli apprendenti di avere una percezione della lingua usata al giorno d'oggi.

Pertanto, nel capitolo 3 verranno illustrate le ragioni per le quali i corpora possono rappresentare delle valide risorse per l'apprendimento della lingua e, nel caso specifico, per il reperimento di testi autentici che contengano i fenomeni grammaticali dell'italiano dell'uso medio. Dopo aver illustrato alcuni concetti base della linguistica dei corpora (McEnery, Wilson 1996, Bieber *et al.* 1998, McEnery, Hardie 2001, Kilgarriff, Grefenstette 2003, Baroni 2010, McEnery, Hardie 2012), verranno presentati alcuni ambiti di applicazione dei corpora alla didattica delle lingue, dedicando particolare attenzione al loro utilizzo per la realizzazione di materiali didattici. Per quest'ultimo ambito si farà riferimento alle pubblicazioni contenute in Sinclair (2004), Granger, Meunier (2008), Meunier, Granger (2008), Aijmer (2009) e Hinkel (2011), agli atti dei convegni *Teaching and Language Corpora* (TaLC) ed alla bibliografia ivi contenuta, raggruppando gli studi per aree tematiche. In particolare, per l'italiano L2/LS si farà riferimento a Corino, Marellò (2009), Merlo (2009), Kennedy, Miceli (2010), Guidetti, Lenzi, Storchi (2012), Lyding *et al.* (2013) e Ivanovska-Naskova (2018).

La seconda parte di questo studio (capitoli 4-6) è dedicata all'impostazione della ricerca. Chiarite le ragioni per le quali i corpora possono rappresentare delle utili fonti per reperire testi da utilizzare come punto di partenza per la progettazione di attività didattiche, nel capitolo 4 verranno illustrati gli obiettivi e i metodi della ricerca. Dopo aver esposto le domande alla base di questa ricerca, verranno innanzitutto descritti i vari corpora

dell'italiano contemporaneo scritto attualmente disponibili. Successivamente verranno esposti i criteri di selezione utilizzati per individuare il corpus più idoneo per il reperimento di testi che contengano i fenomeni grammaticali in questione e che possano essere riutilizzati per la produzione di materiali didattici.

Una volta individuato il corpus, si illustrerà il metodo adottato per la sua interrogazione tramite la progettazione di alcune stringhe di ricerca in linguaggio Corpus Query Language (Evert 2009, Nesi, Thompson 2014), utilizzate per ricercare i fenomeni grammaticali elencati da Sabatini nel suo studio del 1990.

Infine, verrà presentato il metodo adottato per la progettazione e sperimentazione delle unità didattiche e per la realizzazione di un questionario per la raccolta delle opinioni degli studenti sulle attività svolte. Di quest'ultimo si descriveranno la struttura, il contenuto e le caratteristiche principali della piattaforma *online* su cui è stato collocato.

Per la progettazione delle unità didattiche si farà riferimento alla *Second Language Acquisition Theory* di Krashen del 1982, alla didattica acquisizionale (Vedovelli 2003, Vedovelli, Villarini 2003) e al modello dell'unità didattica ripreso e approfondito da Danesi (1988, 1998) da un punto di vista neurolinguistico. Verrà messa in luce l'importanza di strutturare le attività sulla base dell'ordine in cui si attivano i due emisferi cerebrali (Balboni 1994, Danesi 1998: 152 e sgg., Balboni 2002, 2008). Per l'elaborazione del questionario si farà riferimento agli studi di Nunan (1989), Coonan (2000) e Pozzo (2008).

Nel capitolo 5 si illustreranno i procedimenti adottati per il calcolo della frequenza all'interno del corpus dei fenomeni grammaticali oggetto di studio, al fine di determinare quali di essi siano i più ricorrenti. Questo aspetto è di particolare importanza per la presente ricerca, poiché, come verrà descritto, la frequenza dei fenomeni grammaticali nel corpus sarà uno dei criteri adottati per la selezione dei fenomeni grammaticali sui quali costruire dei materiali didattici per l'italiano per stranieri. Essa permette di individuare quali tra i fenomeni linguistici elencato da Sabatini (1990), siano ora maggiormente utilizzati nella produzione scritta di media formalità dagli italofoeni nativi e la cui conoscenza è necessaria agli apprendenti di lingua italiana L2/LS al fine di raggiungere una competenza linguistica vicina all'uso reale della lingua. L'analisi quantitativa sarà supportata dalla presentazione dei dati in tabelle e dalla loro rappresentazione grafica.

Oltre al criterio della frequenza, nella scelta dei fenomeni grammaticali su cui incentrare la progettazione delle unità didattiche si terrà conto anche dei contenuti grammaticali della manualistica per stranieri attualmente in commercio (cfr. Słapek 2017),

al fine di individuare i fenomeni finora esclusi dalle pubblicazioni esistenti e per i quali possa quindi essere rilevante e utile la creazione di materiali didattici.

Una volta identificati i fenomeni grammaticali su cui incentrare la progettazione, i testi che li contengono da includere nelle unità didattiche e la piattaforma di *e-learning* sulla quale collocarle, nel capitolo 6 si descriverà la struttura delle unità didattiche prodotte. Nella loro progettazione si terrà conto della collocazione tra i livelli del Quadro Comune Europeo di Riferimento per la conoscenza delle lingue (QCER) (Consiglio d'Europa 2018, 2002) dei fenomeni grammaticali selezionati e dei fattori neurolinguistici che regolano l'ordine in cui i contenuti linguistici vengono acquisiti dalla mente umana e che, pertanto, sono alla base dell'acquisizione di una lingua (cfr. capitolo 4). Verranno, inoltre, descritti gli esiti delle sperimentazioni delle unità didattiche presso i dipartimenti di italianistica di alcuni atenei europei, che saranno poi confrontati con i dati raccolti tramite un questionario sulle opinioni dei partecipanti. In particolare, nel questionario verranno presi in considerazione i seguenti fattori: la lingua materna degli apprendenti (L1), la conoscenza di altre lingue straniere (LS), la durata dello studio pregresso della lingua italiana, la familiarità dei partecipanti con gli argomenti trattati, sia dal punto di vista dell'uso standard che dell'uso neostandard e, infine, la presenza nella L1 e nelle altre LS conosciute di argomenti grammaticali simili.

Nelle Conclusioni verranno infine proposte alcune riflessioni sulla validità dell'approccio metodologico proposto in questa ricerca, mirata a promuovere una percezione realistica delle stratificazioni sociolinguistiche della lingua usata al giorno d'oggi e fornire degli strumenti di supporto a chi insegna e a chi apprende l'italiano. In particolare, si metteranno in luce sia gli aspetti innovativi sia i limiti della ricerca, ponendo le basi per dare vita in futuro ad uno studio più ampio.

CAPITOLO I

L'ARCHITETTURA DELL'ITALIANO CONTEMPORANEO

In questo capitolo si affronterà il tema dell'architettura dell'italiano contemporaneo. Dopo alcuni brevi cenni sulla storia della lingua italiana (cfr. D'Achille 2003, De Mauro 2011, Dardano 2020 a c. di, Frosini 2020 a c. di, Palermo 2020), verranno trattati gli aspetti sociolinguistici che permettono di identificare le varietà della lingua italiana e le differenze d'uso sulla base degli assi di variazione (Cosieriu 1973, Mioni 1983, Berruto 1993a e Berruto 1993b).

1.1. Brevi cenni storici

Come tutte le varietà romanze, l'italiano deriva dal latino volgare, in particolare dalle varietà della lingua latina parlate dalle diverse popolazioni italiche. Tali varietà si distaccavano dal latino classico-letterario, con forti variazioni di carattere geografico e legate ai gruppi sociali di appartenenza, attraverso gli influssi ricevuti dai substrati linguistici locali delle diverse regioni in cui il latino era stato importato. Dal fiorentino del Trecento, nella sua elaborazione letteraria realizzata da Dante, Petrarca e Boccaccio, derivarono alcune strutture linguistiche fondamentali, quali la fonologia, la morfologia, numerosi aspetti della sintassi e il vocabolario di base (Dardano 2020). Queste strutture vennero prese a modello dell'uso scritto della lingua da parte della linea classicistica dei grammatici, capeggiati da Pietro Bembo, che animarono la fase più acuta di una disputa in ambito letterario, definita *questione della lingua*, in corso sin dal Trecento, nella quale ci si interrogava su quale lingua utilizzare nella penisola italiana.

L'origine della disputa viene fatta risalire al *De vulgari eloquentia* (1303-04), nel quale Dante, alla ricerca di un *vulgare illustre* per la lingua poetica, individuò 14 dialetti, ma giunse alla conclusione che nessuno di essi fosse adeguato a divenire il canone letterario. Tale ideale linguistico poteva essere elaborato solo in una corte ideale, nella quale ciascun letterato italiano portasse i contributi migliori dal proprio dialetto.

Dal suo avvio nel Trecento, la disputa si è protratta per secoli, vivendo una delle sue fasi più acute nel Cinquecento anche in seguito all'avvento della stampa, che fece sorgere la necessità dell'uso di una varietà linguistica omogenea nella penisola italiana.

Tra le tappe importanti della *questione della lingua* e nella codificazione dello standard si ricordano la pubblicazione nel 1525 delle *Prose della volgar lingua* di Bembo, in cui come modelli per la redazione di opere letterarie vennero proposti Francesco Petrarca per la poesia e Giovanni Boccaccio per la prosa, poi la pubblicazione della prima edizione del *Vocabolario degli Accademici della Crusca* nel 1612, alla quale fecero seguito varie edizioni rivedute. Dunque, il modello di lingua codificato non fu il fiorentino parlato, ma “il toscano urbano della classe colta di Firenze” (Galli de’ Paratesi 1984: 60), ossia una varietà scritta letteraria con alcune influenze da altre varietà volgari e dal latino.

Se ne deducono due importanti conclusioni: innanzitutto, non tutte le caratteristiche del fiorentino parlato furono ammesse nella codificazione dello standard, caratteristica che influirà nel corso dei secoli nella percezione di distanza tra lo standard linguistico e la lingua comunemente in uso; inoltre, il fiorentino si impose sugli altri dialetti non grazie ad un potere politico in grado di unificare il territorio e di imporre una politica linguistica, bensì grazie all’affermazione del modello proposto da Bembo sulla base del valore letterario degli autori che lo utilizzarono, al prestigio economico e artistico di Firenze e anche ad alcune caratteristiche strutturali che lo collocavano in una posizione centrale tra gli altri dialetti italo-romanzi.

La varietà fiorentina, che solo a partire dal Settecento venne identificata con il termine *italiano*, fu usata fino all’Unità d’Italia nel 1861 nello scritto colto, il che favorì una certa stabilità e conservazione delle strutture linguistiche. Con l’unificazione del paese si crearono gradualmente le condizioni affinché l’italiano divenisse una lingua nazionale, non solo scritta, ma anche parlata, nonché la lingua materna della popolazione e non più una seconda lingua affiancata ai diversi dialetti: l’alfabetizzazione della popolazione dovuta all’insegnamento scolastico, i cambiamenti nelle condizioni sociali, economiche e culturali e, soprattutto, la diffusione soprattutto nel corso della seconda metà del Novecento della radio e della televisione, portarono l’italiano a contatto con fasce sempre più ampie della popolazione e ad allargare sempre più i suoi ambiti d’uso, raggiungendo anche la lingua orale. Tutto ciò ha innescato il processo da cui deriva l’attuale repertorio linguistico del territorio italiano, alla cui natura complessa, caratterizzata da diverse dimensioni, saranno dedicate le sezioni successive.

1.2. Gli assi di variazione

All'interno del repertorio linguistico italiano è possibile individuare una serie di varietà le cui differenze sono legate, come verrà chiarito di seguito, alla loro diffusione nello spazio e nel tempo, al canale di trasmissione del messaggio, al suo contenuto, alla situazione comunicativa, agli strati e gruppi sociali a cui appartengono gli interlocutori, eccetera.

Nello specifico, i parametri osservabili in una lingua in un dato momento storico, sistematizzati da Eugenio Coseriu (1973), sono tre: la diatopia, relativa allo spazio, in base alla quale una lingua ha caratteristiche diverse a seconda delle aree geografiche in cui è usata, la diastratia, legata alla condizione sociale dei parlanti e influenzata anche dal genere, dall'età, dalla classe sociale, dal livello economico e di istruzione, e la diafasia, legata alla situazione comunicativa, da cui deriva la scelta di un registro linguistico formale o informale (cfr. Coseriu 1973: 135-152). A questi parametri, che determinano la variazione linguistica sincronica, si aggiungono la diacronia, relativa ai cambiamenti nell'uso della lingua legati al passare del tempo e, infine, la diamesia, termine coniato da Alberto Mioni (1983: 508-510), che identifica la variazione di una lingua a seconda del mezzo o canale adottato per comunicare, esemplificabile con la differenza tra scritto e parlato.

Questi “assi di variazione”, combinati tra loro, definiscono lo “spazio linguistico italiano, cioè l'insieme delle possibilità di esprimersi a disposizione dei parlanti di italiano” (Lorenzetti 2002: 23), permettendo di identificare diverse varietà di italiano e di collocare le produzioni linguistiche all'interno dello spazio di variazione di una lingua.

Particolare rilievo verrà dato, ai fini di questa trattazione, alla dimensione diamesica, alla quale verrà dedicata la sezione successiva, con l'obiettivo di metterne in luce la natura complessa e trasversale rispetto agli altri assi di variazione.

1.2.1. La diamesia: da asse di variazione autonomo a sottocategoria della diafasia

Il termine *diamesia* fu introdotto da Mioni (1983), ispirandosi alla terminologia di Coseriu, caratterizzata dal prefisso *dia-* già in uso in linguistica, e fu riferito alle

[...] differenze del m e z z o via via usato per comunicare (per le quali si potrebbe usare il neologismo di ‘dimensione d i a m e s i c a’). Queste ultime non consistono in una pura e semplice opposizione polare tra scritto e orale, ma in un continuum di gradini intermedi: [...]

vi sono, ad es., testi scritti, testi scritti per la sola lettura e cioè per non essere letti ad alta voce, ecc. (Mioni 1983: 508)

L'elaborazione del concetto di diamesia come dimensione di variazione ha dato il via ad un dibattito articolato, avente la sua fase principale negli anni dal 1983 al 1987 e con strascichi successivi nel primo decennio del Duemila fino ai nostri giorni.

Tale dibattito fu incentrato principalmente sulla complessità del definire che cosa rientrasse all'interno della dimensione diamesica: come emerge già dalla definizione di Mioni, non si tratta di una semplice opposizione bipolare tra scritto e parlato, ma di un continuum scritto-parlato, quindi di una serie di diverse forme di realizzazione della comunicazione che vanno, attraverso diverse gradazioni, dal polo dello "scritto-scritto", caratterizzato dal rispetto delle norme linguistiche, al "parlato-parlato", nella terminologia di Nencioni (1983 [1976]), passando attraverso una serie di forme intermedie.

Con la diffusione dei nuovi mezzi di comunicazione di massa, l'opposizione tra parlato e scritto è stata, infatti, aggiornata per includere testi che presentano caratteristiche sia dello scritto che del parlato e che rientrano nell'italiano trasmesso (Sabatini 1982, 1997). Questo include le categorie del "parlato trasmesso", ossia la lingua parlata trasmessa attraverso i mezzi di comunicazione a distanza (radio, televisione, cinema), alla cui base può esservi un testo scritto realizzato per essere ascoltato da molti e che, quindi, ha caratteristiche sia dello scritto che del parlato, e lo "scritto trasmesso", con riferimento alla lingua scritta nelle comunicazioni che avvengono tramite Internet, quali la posta elettronica, i vari servizi di messaggistica istantanea ecc., che ha caratteristiche assimilabili a quelle del parlato (cfr. D'Achille 2003: 216). Rientrano in questo contesto anche forme ibride quali lo "scritto per essere detto come se non fosse scritto" (Lavinio 1986: 16), relative, ad esempio, ai copioni cinematografici.

La diamesia è una dimensione di variazione complessa, "particolarmente necessaria in una situazione linguistica come quella italiana" (Lorenzetti 2002: 22), nella quale la lingua scritta, codificata dalle grammatiche, è stata identificata per secoli con lo standard, in contrapposizione alla lingua d'uso comune, sia nello scritto che nel parlato. Ed è proprio nell'ambito del dibattito sull'italiano dell'uso medio (Sabatini 1985) o neostandard (Berruto 1987) e sull'italiano popolare, (si vedano rispettivamente le sezioni 1.3.4. e 1.3.2) che si sviluppa la proposta di Mioni di una nuova dimensione di variazione. Interrogandosi su quale posizione occupasse la dimensione diamesica nello studio dell'italiano popolare, Mioni propone che per lo studio di questa varietà linguistica le due dimensioni, solitamente

considerate un tutt'uno (ossia, formalità/scritto vs informalità/parlato), vengano separate, permettendo così di considerare separatamente la relazione dello scritto e dell'oralità con la diafasia, ossia con i diversi registri linguistici, poiché la loro sovrapposizione (scritto = formalità; parlato = informalità) avverrebbe solo nei casi più semplici di realizzazione della comunicazione.

La proposta di Mioni sull'esistenza di una vera e propria dimensione diamesica è accolta da Berruto nel 1985 nel contributo *Per una caratterizzazione del parlato: l'italiano parlato ha un'altra grammatica?*, come emerge dai riferimenti fatti a Mioni, dall'uso del termine *varietà* per definire la gamma di possibili realizzazioni scritte e orali e dalla domanda: "Come si interseca la variazione scritto/parlato con la variazione sociale e con quella diafasica?" (Berruto 1985a: 139), la cui risposta arriverà con maggiore definizione nel suo studio del 1987. In questo lavoro, infatti, si consoliderà la visione della diamesia come dimensione di variazione, una proposta adottata dalla comunità scientifica italiana, seppur con l'espressione di alcune cautele da parte dello stesso Berruto:

Il riconoscimento dell'autonomia della dimensione diamesica non è del tutto chiarito in sede teorica. Indubbiamente, uso scritto e uso parlato rappresentano due grandi classi di situazioni d'impiego della lingua: e questo è un buon argomento per ritenere la diamesia una sottocategoria della diafasia. D'altra parte, è anche vero che l'opposizione scritto-parlato taglia trasversalmente la diafasia e le altre dimensioni, e non è riconducibile completamente all'opposizione formale-informale. (Berruto 1987: 22)

A questo proposito nella relativa nota (1987: 22 n. 12) Berruto specifica che, sebbene di solito un registro molto formale possa coincidere con l'uso scritto e viceversa, ci sono anche situazioni in cui il parlato può essere più formale dello scritto: questo avviene perché la scelta dell'uso orale o grafico dipende dalla natura del mezzo utilizzato, mentre il codice parlato e il codice scritto dipendono dalla pianificazione, dalla formalità della situazione ecc.

Esemplificando: un messaggio in una chat è considerato un testo scritto a causa del mezzo utilizzato per la sua produzione (la tastiera del computer o del cellulare), ma la sua natura di testo scritto non lo rende automaticamente una produzione formale. Quel messaggio può avere, ad esempio, caratteristiche del parlato e risultare meno formale di una produzione orale vera e propria prodotta in situazioni di maggiore ufficialità. Questo è possibile poiché il grado di pianificazione dei contenuti e di formalità dipendono da una varietà di fattori, ad esempio il tipo di rapporto esistente tra mittente e destinatario, i loro

gruppi sociali di appartenenza, la loro provenienza geografica ecc.. Per comprendere questo aspetto si immagini, ad esempio, come una stessa notizia di sport, cronaca o politica possa essere comunicata in uno scambio in chat informale tra due utenti con un rapporto di familiarità e come la stessa notizia possa venire annunciata oralmente in un contesto formale da un giornalista televisivo.

Il tema della natura trasversale della diamesia viene ripreso da Berruto nel 1993, che pone da un lato l'accento sulla distinzione tra il canale scritto/orale e la situazione comunicativa e dall'altro sulla realizzazione concreta ricollegabile alla maggiore o minore formalità, ossia alla diafasia:

La differenziazione fra parlato e scritto, pur realizzandosi in concrete condizioni d'uso, è infatti preliminare e indipendente rispetto all'utente (e allo strato sociale di appartenenza) e alla stessa diafasia, in quanto almeno per una certa parte è determinata dalle caratteristiche generali del canale di comunicazione e dalle circostanze ambientali di attuazione della comunicazione. Però, nella realizzazione effettiva viene poi assorbita dalla variazione diafasica e dai tipi di testi relativi. Ciò rende particolarmente difficile sceverare, in molti casi, quello che è proprio del parlato e quello che è invece proprio delle varietà che si servono della modalità parlata e del canale fonico-acustico. (Berruto 1993b: 38)

Dunque, sebbene la proposta adottata dalla comunità scientifica italiana corrisponda a quella da lui elaborata nel 1987, che vede la diamesia come un asse di variazione autonomo, riprendendo il tema Berruto arriva ad elaborare una nuova articolazione dello spazio di variazione che vede rientrare la diamesia nell'area della diafasia, considerandole inscindibili¹:

All'interno della dimensione diafasica si situa tuttavia un'importante altra sottodimensione di variazione, che per certi aspetti trascende di fatto l'ambito della stessa diafasia, anche se fondamentalmente il fattore principale che correla con essa sta nella situazione comunicativa, e per la precisione nel mezzo o canale 'fisico' attraverso cui passa la comunicazione verbale. Si tratta della differenziazione fra uso parlato e uso scritto della lingua, dipendente per molti aspetti dalle caratteristiche strutturali e di realizzazione che il mezzo o canale impone alla codificazione del messaggio linguistico; spesso ci si riferisce a questo genere di variazione,

¹ Si vedano a questo proposito anche i successivi studi (Berruto 2011b; Berruto e Cerruti 2014: 147-148).

che a rigore si situerebbe all'interno della dimensione diafasica, come a una quarta dimensione fondamentale di variazione, la 'variazione diamesica'. (ivi: 93)

Da questo breve *excursus* è emerso come l'introduzione della dimensione diamesica nello spazio variazionale abbia avviato, fin dalla sua nascita, un lungo dibattito sulla sua definizione e sull'evoluzione di un dubbio relativo in particolare al suo rapporto con la diafasia. Le perplessità non riguardano tanto le differenze tra scritto e parlato, quanto come esse possano essere descritte su un asse di variazione, come è stato sottolineato prima da Holtus (1984) e poi da Radtke (1992). Una possibile proposta di definizione dell'asse diamesico è la seguente:

Poiché il repertorio [linguistico] nel suo complesso, ancor prima delle singole dimensioni di variazione, è definito come continuum pluridimensionale con addensamenti, anche la diamesia, una volta inserita in tale contesto, dovrà presentare caratteri adeguati. Il continuum diamesico non è orientato, perché le varietà non si dispongono da un estremo sociolinguisticamente 'alto' a uno 'basso'; né polarizzato, allo stesso modo in cui lo sono quello diafasico o diastatico. (Pistolesi 2015: 42)

Dunque, poiché l'uso scritto e l'uso dell'oralità non sono legati in maniera fissa alle variazioni diafasiche, ma possono essere impiegati in grande varietà di realizzazioni della comunicazione con diversi gradi di formalità (vale l'esempio sopra citato della notizia di cronaca scritta tra amici in una chat o comunicata ufficialmente da un giornalista radio/televisivo), le varietà diamesiche sono da considerarsi trasversali rispetto alle altre dimensioni, dalle quali non possono essere completamente isolate, se non per specifiche esigenze di studio (cfr. Berruto 1993a e 1993b).

Per quanto riguarda le dimensioni di variazione, la loro combinazione definisce, come già accennato in precedenza, lo spazio linguistico italiano:

All'interno di questo spazio si sono individuate aree corrispondenti a varietà di italiano relativamente ben individuabili e definibili. Tra esse l'italiano standard, l'italiano "neostandard" (o "comune" o "medio"), l'italiano popolare e l'italiano regionale sono, per così dire, varietà complete, dotate di peculiarità a tutti i livelli. (Lorenzetti 2002: 23)

A queste varietà si farà riferimento nelle sezioni successive.

1.3. Lo spazio linguistico italiano

Nelle seguenti sezioni si descriveranno le varietà dello spazio linguistico italiano, partendo dai poli estremi, ossia l'italiano standard e l'italiano popolare/dei semicolti, passando poi all'italiano regionale e concludendo con l'italiano neostandard, al quale verrà dato particolare rilievo coerentemente agli scopi di questa trattazione.

1.3.1. Italiano standard

Come già accennato nella sezione 1.1, il dibattito teorico sul concetto di standard linguistico è in corso almeno dal Trecento, cambiando nel corso dei secoli con il ciclico ritorno della cosiddetta *questione della lingua*. Ulrich Ammon (1986, 2003) identifica lo standard come:

1. codificato e descritto esplicitamente da un insieme di regole normative e da testi di riferimento (opere letterarie modello, grammatiche, dizionari);
2. sovraregionale, cioè non limitato ad una specifica area geografica;
3. elaborato, ossia possiede tutti i mezzi e le risorse linguistiche (vocabolario, strutture grammaticali e testuali, ecc.) per poter essere utilizzabile in tutti i domini, anche quelli specialistici;
4. in uso da parlanti appartenenti agli strati sociali superiori;
5. invariante, come conseguenza della codificazione normativa grammaticale;
6. utilizzabile in tutti gli usi della lingua scritta.

Tale definizione è successivamente ripresa da Lorenzetti (cfr. 2002: 25-28), che aggiunge le seguenti caratteristiche in base alle quali, inoltre, lo standard è:

7. la varietà sociolinguistica normale, ossia la meno marcata;
8. la varietà di riferimento, che rappresenta la norma (e in quanto tale garantisce l'invarianza a cui fa riferimento Ammon);
9. la varietà utilizzabile in forma orale in tutti i contesti, da qualunque parlante per qualsiasi argomento.

La moderna questione della lingua, che va riproponendosi sin dal Cinquecento, fino ad Alessandro Manzoni (Koch 2014), si interroga non solo su quale sia la varietà di riferimento, ma sulla possibilità che questa varietà esista. Il dibattito linguistico si è sviluppato secondo due correnti riassumibili come segue:

- a) l'italiano standard esiste e va identificato per la pronuncia con l'italiano fiorentino contemporaneo, privo di alcuni tratti prettamente locali, per la grammatica nell'italiano ottocentesco post-manzoniano;
- b) l'esistenza in sé di uno standard è dubbia². Nel caso specifico dell'italiano standard, la sua grammatica di riferimento, presentata nel dettaglio in Serianni (1988), si può definire nel complesso come un "fiorentino emendato" (Galli de' Paratesi 1984: 57), quindi una varietà che non coincide con nessuna delle varietà effettivamente usate dalla comunità linguistica all'interno del territorio nazionale. Come precisato da Berruto (2010):

L'italiano standard in effetti non ha mai, fin dalla codificazione cinquecentesca, coinciso esattamente con il fiorentino, e sin dal Seicento ha accolto, data anche la mancanza fra il tardo Cinquecento e l'avanzato Ottocento di un centro preminente che imponesse una norma, innovazioni di varia provenienza. La distanza dal fiorentino si è ancora accresciuta dopo l'Unità d'Italia, nonostante i tentativi puristici di imporre il fiorentino moderno come modello, in particolare per la pronuncia.³

Rispetto a quest'ultimo punto, come riportato da Lorenzetti (2002: 27-28), proprio nelle aree che identificano lo standard si rilevano tratti in regresso, rispetto ai quali si potrebbe affermare che lo standard non esiste dal punto di vista della pronuncia, poiché nessun parlante ha acquisito nativamente la pronuncia fiorentina emendata (es. la distinzione tra *e* e *o* aperte e chiuse e tra *s* e *z* sorde e sonore non si è mai imposta al di fuori della Toscana e del centro), che è di solito appresa tramite corsi di dizione da parte di parlanti professionisti (attori, presentatori radiofonici e televisivi). Un discorso simile vale per la grammatica, poiché la varietà codificata coincide per la maggior parte con l'italiano scritto formale, ma non riflette le varietà parlate, escludendo tratti morfosintattici effettivamente presenti nella lingua contemporanea.

Il prestigio della varietà standard normativa perde dunque terreno, riducendo la propria influenza allo scritto formale, mentre "lo spazio perduto dallo standard normativo viene

² In natura non esistono varietà di lingua standard, poiché esse sono il risultato di un'azione esplicita di codificazione di regole grammaticali e normazione. Lo standard si basa, dunque, inevitabilmente su un certo grado di artificialità (cfr. Berruto 2010). La formazione di una varietà standard è, infatti, un processo sociale e culturale che deriva dall'azione congiunta di più fattori (Ammon 2003): (a) parlanti e scriventi professionali, che producono testi modello; (b) autorità normative in ambito linguistico (per es., accademie) che sono la fonte di istruzioni e correzioni; (c) codici linguistici (i manuali descrittivo-normativi di riferimento); (d) esperti di lingua, che descrivono e giudicano le produzioni linguistiche.

³ Versione html, pertanto priva di pagine.

guadagnato dalle altre varietà, che diventano così più accessibili, meno censurate, più capaci di espansione” (Lorenzetti 2002: 28). A tale varietà di italiano, caratterizzata da tratti linguistici che, in passato erano esclusi dallo standard e che ora appaiono ampiamente diffusi e accettati, sarà dedicata la sezione 1.3.4, mentre la sezione successiva sarà dedicata alla varietà collocata sul polo opposto allo standard: l’italiano popolare o dei semicolti.

1.3.2. Italiano popolare (o “dei semicolti”)

L’italiano popolare è il “modo di esprimersi di un incolto che, sotto la spinta di comunicare e senza addestramento, maneggia quella che ottimisticamente si chiama la lingua ‘nazionale’, l’italiano” (De Mauro 1970: 49). In seguito, Manlio Cortelazzo descrisse l’italiano popolare come “il tipo di italiano imperfettamente acquisito da chi ha per madrelingua il dialetto” (Cortelazzo 1972: 11). Con l’espressione “italiano popolare” si indica, quindi, una varietà linguistica “di contatto”, poiché come l’italiano regionale (si veda la sezione seguente) deriva dall’incontro tra lingua e dialetto e quindi presenta fenomeni di interferenza linguistica, oltre che rappresentare una varietà marcata in basso lungo l’asse diastratico (Berruto 1993a e 2012 [1987]).

A questa definizione si affianca anche quella di “italiano dei semicolti” (Bruni 1984, D’Achille 1994), con la quale si mette in evidenza che questa varietà, sia parlata che scritta, è tipicamente utilizzata dalle fasce meno istruite della popolazione in alternativa al dialetto. È utilizzata da parlanti per i quali il dialetto è la lingua materna, e l’apprendimento dell’italiano è stato limitato ai pochi anni di istruzione scolastica ed è documentata in contesti quali la scrittura di diari, lettere indirizzate alle autorità pubbliche e lettere dal fronte durante la Prima e la Seconda guerra mondiale. Pur essendo una varietà sia scritta che parlata, prima della diffusione del telefono era una varietà per lo più scritta, il che non implica che gli utenti fossero in grado di dominare la correttezza formale della lingua, essendo questi, anzi, caratterizzati dal punto di vista sociolinguistico proprio dal basso grado di istruzione.

Dal punto di vista della sua diffusione, “Non ci sono dubbi che un it. pop. veramente unitario non esista. [...] La marcatezza diatopica in Italia è preliminare a tutte le altre dimensioni di variazione” (Berruto 2012 [1987]: 130). Si tratta, dunque, di una varietà fortemente caratterizzata in diatopia, poiché presenta un’elevata frequenza di caratteri regionali.

Tra le caratteristiche si cita qui a titolo esemplificativo (cfr. D'Achille 2003: 206-207, Berruto 2011a) una forte tendenza alla semplificazione che investe sintassi, morfologia, fonologia e lessico dello standard. A questo proposito, la nozione di semplificazione è intesa in senso lato:

Con il termine di “semplificazione” si designano, al di là del valore corrente intuitivo del termine [...]: a) una riduzione all'essenziale [...] eventualmente misurabile in termini di (un minor) numero di morfemi e di (una minore) articolazione strutturale interna; b) il richiedere meno regole, vale a dire un minor numero di regole ed eventualmente un'applicazione diversa, con ridotto numero di passaggi, nel processo di generazione e trasformazione, dalla struttura profonda o forma logica alla struttura superficiale [...]; c) l'abbandono della lingua alla “naturalità”, a tendenze evolutive latenti all'interno del sistema e miranti a una 'ottimizzazione' e a una minore marcatezza. (Berruto 1983: 40)

Ad esempio: la semplificazione del paradigma dell'articolo con la riduzione del numero delle sue forme e l'estensione dell'uso di alcune di esse: *il zio, dei amici, il spettatore, i altri, un sbaglio*; la regolarizzazione di paradigmi verbali (*dassi vs dessi, venghino vs vengano, persuadere vs persuadere, soddisfaceva vs soddisfaceva, vadi/vadino* al congiuntivo presente, *misimo* come prima plurale del passato remoto di *mettere, potiamo* come prima persona plurale del presente di *potere, bevavamo* come imperfetto di *bere* ecc.), la riduzione di un intero paradigma di pronomi di terza persona al solo *ci* (*dicci* al posto di *dì a noi, di loro, dille, digli*), uso del possessivo di terza persona *suo* anche per il plurale (gli amici con il *suo* [= loro] cane) e altri fenomeni (cfr. Berruto 2011a, Lorenzetti 2002 pp.33-34); alcune realizzazioni che mostrano lo scarso controllo ortografico, quali l'assenza della percezione dei confini tra le parole (*lamico vs l'amico*), mancanza o utilizzo improprio dell'acca, delle doppie e degli accenti eccetera. La semplificazione riguarda anche il lessico, caratterizzato da termini generici e da un vocabolario ridotto.

Rispetto al neostandard, a cui si farà riferimento successivamente, questa varietà è caratterizzata da una maggiore influenza dei dialetti, da una più elevata frequenza di elementi rifiutati dallo standard e da tratti linguistici idiosincratici. Sebbene si manifestino con frequenza diversa, condivide con il neostandard colloquiale alcune tendenze: ad esempio, la già citata tendenza a regolarizzare i paradigmi verbali è più frequente nell'italiano dei semicolti (le forme *venghi, venghino* sono esclusive dell'italiano dei

semicolti, mentre *persuàdere* e *soddisfava* si realizzano anche nell'italiano colloquiale). Al di là di queste lievi variazioni,

ciò che distingue più nettamente l'italiano popolare dall'italiano colloquiale e informale è lo status sociolinguistico: l'italiano colloquiale e italiano informale non sono vere e proprie varietà, ma piuttosto registri dell'italiano neostandard, che possono essere scelti dai parlanti per adattarsi a determinate situazioni. (Lorenzetti 2002: 34)

La sezione successiva sarà dedicata all'italiano regionale e alle caratteristiche che lo distinguono dall'italiano popolare o dei semicolti.

1.3.3. Italiano regionale

L'italiano regionale è stato identificato tra le varietà dell'italiano da Giovan Battista Pellegrini (1960), che ha articolato il repertorio linguistico italiano in quattro varietà: italiano letterario, italiano regionale, dialetto regionale (o *koiné* dialettale) e dialetto locale.

L'italiano regionale è dunque una varietà parlata in una determinata area geografica, nata dal contatto con la lingua di prestigio (analogamente alla derivazione dei dialetti italiani dal latino volgare), molto influenzata dal dialetto e che si differenzia sia dallo standard sia dalle altre varietà regionali per la presenza di elementi locali:

Storicamente è l'esito dell'adattarsi alle singole realtà dialettali preesistenti – e che normalmente erano l'unico modo di comunicare – di una lingua 'piovuta dall'alto': dalla scuola, dalle scritture, dall'amministrazione, poi in tempi più recenti anche dal cinema, dalla radio, dalla TV. (Poggi Salani 2010)⁴

Con la già citata diffusione dell'italiano come lingua nazionale, nelle diverse aree italiane si è verificato un rafforzamento delle varietà standard regionali che, pur essendo diatopicamente marcate (ossia, presentando caratteristiche che le differenziano dalle altre varietà geografiche), non sono percepite come scorrette e sono utilizzate anche dai parlanti più colti. Inoltre, l'italiano regionale include sia le varietà apprese direttamente come lingua materna, sia quelle sviluppate da dialettofoni nativi.

Per quanto riguarda il rapporto tra italiano popolare e italiano regionale, quest'ultimo è una varietà legata alla provenienza geografica dei parlanti, mentre l'italiano popolare è una

⁴ Versione html, pertanto priva di pagine.

manifestazione regionale dell'italiano particolarmente marcata in diastratia, influenzata quindi dall'estrazione socio-economica e culturale dei parlanti: “la marcatezza regionale e quella sociale si sommano nel dare luogo [...] a tanti italiani regionali popolari. Un it. popolare specifico sarà sempre anche un italiano regionale” (Berruto 2012[1987]: 132). Rispetto all'italiano standard, le diverse varietà di italiano regionale sono varianti geografiche, mentre le diverse varietà di italiano popolare sono varianti sia geografiche che sociali (Mengaldo 1994: 108-110). Riassumendo, si può affermare, quindi, che l'italiano popolare è un sottoinsieme più marcato dell'italiano regionale.

Un tratto essenziale nella regionalità linguistica è l'intonazione, ossia l'insieme dei tratti soprasegmentali dell'enunciato. Come nota Lorenzetti (2002: 30), è rarissimo trovare pronunce del tutto prive di tratti regionali, anche presso parlanti colti e in situazioni formali. Nella pronuncia “l'italiano regionale colto coincide, di fatto, con l'italiano neostandard parlato, mentre la regionalità nella grammatica, nella sintassi e nella scelta delle parole è spesso controllabile (e dunque, se il caso, evitabile) da parte dei parlanti che abbiano un certo grado di istruzione” (ivi: 30).

Un accento troppo marcatamente regionale è invece ritenuto sub-standard (Berruto 2011a). Inoltre, il grado di regionalità di un testo non è determinato solo dalla quantità di elementi regionali che presenta, ma anche dalla valutazione sociale attribuita ad essi (ossia, il fatto che alcuni tratti siano percepiti come un regionalismo più marcato). Il lessico rappresenta un fattore di forte diversificazione regionale: ad esempio, tutto il vocabolario della vita quotidiana è soggetto ad interferenze dei dialetti, poiché molti parlanti danno una morfologia italiana a lessemi dialettali, creando così dei geosinonimi che variano di area in area, alcuni dei quali entrano nel repertorio nazionale. Sul piano morfologico e sintattico vi sono, invece, meno tratti largamente condivisi e linee di tendenza comuni a tutti gli italiani regionali. In tal senso è possibile affermare che vi sia uno scambio reciproco tutt'ora in corso italiano standard e le varietà regionali.

Vi è, infine, un'altra varietà, derivata da un processo di ristandardizzazione dell'italiano contemporaneo, che sarà al centro della presente trattazione e a cui è dedicata, in particolare, la sezione successiva.

1.3.4. Italiano neostandard

Come già affermato precedentemente nella sezione 1.1, la diffusione dell'italiano in sempre maggiori ambiti d'uso, fino a diventare la lingua nazionale, è un processo iniziato

gradualmente con l'Unità d'Italia e che si è intensificato nella seconda metà del Novecento, grazie soprattutto all'istruzione obbligatoria, alla diffusione dei mezzi di comunicazione di massa e ai cambiamenti sociali, descritti da De Mauro (2011, 2014) e da Poggi Salani (2012), dovuti all'industrializzazione e alle migrazioni interne.

All'origine del bisogno comune di una lingua *passé-partout*, valida su tutto il territorio nazionale, vi sono vari elementi (De Mauro 2011: 201, Poggi Salani 2012:127-129). Innanzitutto, le migrazioni interne, dal Sud e dalle Isole verso il Centro-Nord più industrializzato, dagli anni '50 del Novecento in poi: la presenza di immigrati nelle nuove sedi ha creato la necessità di italianizzazione anche presso gli abitanti originari di quelle stesse sedi, dove il dialetto era risultato sufficiente solo finché la popolazione era stata geograficamente omogenea. Altri fattori sono i grandi mezzi di diffusione della lingua parlata, ossia la radio, il cinema sonoro dal 1930 e la televisione dal 1954, con spettacoli e trasmissioni di massa, ma anche i dischi e il telefono, l'obbligo scolastico a 14 anni (legge 31 dic. 1962), il servizio militare obbligatorio e, infine, la facilità di spostamenti e scambi sul territorio nazionale al di là della migrazione, che includono anche il servizio militare obbligatorio, che ha contribuito allo spostamento di una fetta della popolazione verso regioni distanti dal proprio luogo d'origine.

In questa fase di evoluzione, la varietà standard, fino ad allora una realizzazione aulica, utilizzata per lo più nello scritto formale, dovette adattarsi a tante esigenze pratiche della comunicazione nella vita quotidiana. Di conseguenza, si è verificata una pressione del parlato sulle strutture della lingua scritta, che ha causato una serie di ristrutturazioni del sistema linguistico: hanno incominciato, così, ad emergere tratti linguistici del tutto innovativi o già esistenti da secoli (ma sino ad allora sanzionati dalla grammatica tradizionale, perché ritenuti sub-standard), innescando così un processo di ristandardizzazione della lingua.

Il processo appena descritto è comune a tutte le tradizioni linguistiche, dove accanto alla norma tende a formarsi un uso caratterizzato da maggiore flessibilità, tipico dei registri informali del parlato, che accetta una serie di tratti un tempo sanzionati.

In ambito italiano a questa varietà Sabatini (1985) ha assegnato il nome di *italiano dell'uso medio*, denominazione che indica la media formalità diafasica sia del parlato che dello scritto in cui è accettato, mentre Berruto ha parlato di *neostandard* a partire dal 1987, attribuendo questa denominazione a quell'insieme di

costrutti, forme e realizzazioni che non erano presentati nel canone ammesso dalle grammatiche e dai manuali, o che, quando vi erano menzionati, lo venivano in quanto costrutti, forme o realizzazioni del linguaggio popolare o familiare o volgare, oppure regionali, e quindi da evitare nel bel parlare o scrivere. (Berruto 2012 [1987]: 73)

Il neostandard è dunque il risultato di un processo di graduale ampliamento di ciò che è compreso nella norma, come affermato da Alberto Sobrero, che nel 1992 descrive così questa varietà, per la quale utilizza il termine proposto da Berruto:

[Il neo-standard] è diffuso nelle classi medio-alte e nella parte più acculturata della popolazione, ed è realizzato nel parlato più che nello scritto. L'etichetta di neo-standard si riferisce al fatto che su questo livello, oggi in piena evoluzione, troviamo un gran numero di forme che via via "risalgono" dai livelli inferiori (sub-standard): prima relegate nell'area delle forme "colloquiali" (o, come dicevano i vocabolari, "triviali"), ora si diffondono e sono accettate nella lingua nazionale. Lo standard così, a sua volta, estende i propri confini. (Sobrero 1992a: 5)

Il sub-standard a cui fa riferimento Sobrero (1992a) è composto da quegli elementi che, pur essendo ampiamente usati dai parlanti, sono esclusi dalla norma riconosciuta e codificata, sono principalmente tipici dell'uso parlato informale spontaneo e rientrano nelle varietà diastratiche e diafasiche basse, nell'italiano regionale marcato e nell'italiano popolare.

Il sub-standard non ha prestigio nella comunità linguistica ed è in genere oggetto di giudizi negativi, quali "cattivo italiano", "non è corretto", "sbagliato" e "dialettale" (Galli de' Paratesi 1984: 143-205), che rispecchiano la sua sanzione come errore nella scuola. Non essendo soggetto alle norme che inquadrano la lingua standard, il sub-standard è molto variabile, poiché in esso trovano espressione la naturalezza, le strategie dei parlanti finalizzate alla semplificazione, ad esempio in quelle aree della lingua in cui lo standard presenta irregolarità, eccezioni e paradigmi complessi, soprattutto nella morfologia, e strategie di adattamento pragmatico e funzionale al contesto basate sull'utilizzo di strutture sintattiche influenzate dalla situazione comunicativa e dai partecipanti all'interazione, ad esempio un lessico più generico, collegato all'uso della deissi (si vedano le già citate sezioni dedicate all'italiano regionale e popolare; cfr. Berruto 1990, 2011b, Simone 1991).

Per tornare alla citazione di Sobrero (1992a), è per queste ragioni che spesso il sub-standard contiene in sé quelli che potrebbero essere dei potenziali mutamenti linguistici e

può svelare tendenze e caratteristiche del sistema linguistico che non emergono attraverso lo standard, ma che in alcuni casi “via via “risalgono” dai livelli inferiori (sub-standard)”, entrando nella varietà diffusa a livello nazionale che costituisce il neostandard.

A cavallo tra gli anni Ottanta e Novanta le caratteristiche innovative passate dallo status di “tratti o scelte sub-standard a tratti o scelte (quasi) standard, o comunque non più sub-standard, presenti o diventate comuni anche nel parlato (e spesso nello scritto) colto o mediamente colto” (Berruto 2012 [1987]: 11) sono state oggetto di numerosi studi, tra cui Sabatini (1985, 1990, 1997, 2016), Moretti, Petrini, Bianconi (1992, a c. di), Voghera (1992), Sobrero (1993, a c. di), D’Achille (1990), Lorenzetti (2002) e altri. Tali studi hanno dimostrato che tratti ritenuti nettamente sub-standard dal punto di vista diastratico e/o diafasico, tipici di “varietà basse, di solito cioè dell’italiano popolare o di un parlato molto informale e trascurato, emergono significativamente anche nell’italiano dei parlanti colti e/o di registro abbastanza controllato” (Berruto 2012 [1987]: 112). Secondo D’Achille (2003), l’individuazione di questa varietà tra gli assi di variazione ha permesso di descrivere il repertorio della lingua italiana in maniera più completa:

L’italiano dell’uso medio e il neostandard sono caratterizzati da fatti morfosintattici e lessicali che non sempre rappresentano delle effettive novità, spesso si tratta di fenomeni già documentati in testi del passato, ma censurati o ignorati dalle grammatiche [che] si sono progressivamente diffusi, tanto da apparire ormai del tutto normali non solo nel parlato, ma anche in certi tipi di testi scritti. (D’Achille 2003: 33)

A questo proposito, nel capitolo successivo verranno esaminati nel dettaglio i fenomeni grammaticali che Sabatini nei suoi studi del 1985 e del 1990 identifica come i tratti caratterizzanti dell’italiano dell’uso medio. Si tratta di un tema centrale rispetto a questa trattazione, poiché sulla base di quanto è stato esposto, si ritiene che questa varietà, da tempo oggetto di riflessione linguistica, meriti di essere ulteriormente indagata dal punto di vista delle implicazioni che la ridefinizione dello standard linguistico ha nell’ambito dell’apprendimento dell’italiano come lingua straniera.

CAPITOLO II

APPROFONDIMENTO TEORICO SULL'ITALIANO DELL'USO MEDIO

Dopo aver tracciato nel capitolo precedente una panoramica delle varietà dello spazio linguistico italiano, dando particolar rilievo all'italiano neostandard, in questo capitolo verranno approfonditi i tratti linguistici che caratterizzano quest'ultima varietà sulla base degli studi di Sabatini del 1985 e del 1990, che hanno fornito un contributo fondamentale nel dibattito sulla ristrutturazione dello standard linguistico italiano.

2.1. L'italiano dell'uso medio

Come anticipato precedentemente, Sabatini aveva in un primo tempo (1985) descritto l'*italiano dell'uso medio*, definendolo come una varietà linguistica diffusa su tutto il territorio nazionale, tipica dell'uso parlato e scritto di media formalità. Le sue caratteristiche linguistiche sono identificabili in una lista di 35 tratti fonologici, morfologici e sintattici diffusi su tutto il territorio italiano, “usati da persone di ogni ceto e ogni livello di istruzione; [...] [che] non sono limitati al discorso “orale-non pianificato”, ma risultano pienamente funzionali anche per un discorso “scritto-pianificato”, purché non decisamente formale” (Sabatini 1985: 171) e che rappresentano la diffusione e “l'accettazione, nell'uso parlato e scritto di media formalità, di un tipo di lingua che si differenzia dallo “standard” ufficiale” (ivi: 155).

Tali tratti sono ritenuti sufficienti da Sabatini per identificare una varietà della lingua nazionale “a cui si addice – meglio di altre denominazioni più riduttive o più vaghe, come “italiano colloquiale” o “colloquiale-informale” – il nome di “italiano dell'uso medio parlato e scritto”, e, per brevità, ITALIANO DELL'USO MEDIO” (ivi: 171), che “si candida ad occupare, dopo secoli di ostracismo, il baricentro dell'intero sistema linguistico italiano o per lo meno a condividere con lo standard ufficiale il crisma della norma” (ivi: 175).

Successivamente, in un suo studio del 1990, Sabatini rielaborò la lista, riducendola ai seguenti 14 tratti linguistici¹:

1. *lui, lei, loro* in funzione di soggetti;

¹ Le denominazioni di ciascun fenomeno grammaticale sono tratte dalla lista del 1990. Allo scopo di facilitarne la comprensione, in alcuni casi alla denominazione originale è stata aggiunta una formulazione semplificata, che sarà poi utilizzata nel prosieguo di questo lavoro per ragioni di brevità e maggiore chiarezza.

2. *gli* come forma dativale onnivale, ossia l'impiego nella forma dativale di *gli* al posto di *le* (*a lei*) e *loro* (*a loro*);
3. la combinazione di una preposizione e di un articolo partitivo, ossia l'impiego del partitivo preceduto da preposizione (*con degli amici*);
4. la dislocazione dell'elemento tematico a sinistra o a destra della parte rematica della frase, con la ripresa pronominale atona, ossia l'impiego di dislocazioni a sinistra (*quel libro l'ho già letto*), a destra (*l'ho già letto, quel libro*) della parte rematica della frase con la ripresa pronominale atona, che può giungere fino a costruzioni con uso pleonastico della particella pronominale, tipo *a me mi piace leggere, di pane non ne ho più* (ossia, il doppio partitivo), e all'anacoluto;
5. il *che* polivalente, sia con valore temporale, sia con altri valori (spaziali, personali);
6. il connettivo *per cui* (derivato per ellissi da *motivo, fatto, ragione per cui*, oppure per generalizzazione di casi con normali antecedenti nominali) in funzione di congiunzione causale-consecutiva;
7. il semplice *cosa?* interrogativo diretto e indiretto, ossia il suo uso al posto di *che cosa?*;
8. forme d'ingresso o connettivi testuali come *allora, comunque* e anche *ma, e, o* iniziali di discorso o di enunciato;
9. l'indicativo invece del congiuntivo nei seguenti cinque casi: dopo i *verba putandi*, ossia i verbi di opinione (*credo che hai ragione*), dopo una dichiarativa negativa (*non credo che hai torto*); in un'interrogativa indiretta (*gli chiesi se poteva aiutarmi*); nel periodo ipotetico dell'irrealtà (*se me lo dicevi, ci pensavo io*); nelle relative restrittive (*sei l'unico che parla arabo*);
10. la concordanza a senso tra il soggetto collettivo singolare seguito da un'espressione partitiva e il predicato plurale (*un milione di elettori non hanno votato*);
11. la posposizione del soggetto rematico al predicato (ossia il soggetto in posizione rematica rispetto al predicato), talvolta con discordanza di numero con questo (*non ci sono giornali e non c'è giornali*, da cui si passa anche a *niente giornali*)²;

² La definizione corretta di questo fenomeno, a cui si fa riferimento in questa trattazione, è contenuta nello studio di Sabatini del 1990, p.225. A questo proposito si segnala che, invece, nella riedizione del saggio nel 2011 all'interno della raccolta *L'Italiano nel mondo moderno* il fenomeno è riportato erroneamente come "preposizione del soggetto rematico al predicato" e non come "posposizione del soggetto rematico al predicato", cfr. Sabatini (2011: 98). Dunque si poteva ipotizzare o che la riedizione del 2011 contenesse un refuso o che l'attenzione dell'autore si fosse effettivamente spostata sul tema del soggetto rematico *preposto* al verbo (esempio: *ANNA parla*, ove il tono di voce mette in rilievo il soggetto in risposta a una domanda del tipo: *Parla Marco?*). Il dubbio è stato sciolto dal confronto con lo stesso

12. la diatesi media dei verbi (non solo i più comuni *mi ricordo, mi sbaglio, me ne vado*, ma *mi vedo, mi godo, mi bevo, mi fumo, mi mangio* ecc.), ossia, secondo la versione del 1985, la costruzione dei verbi con forma pronominale per indicare una più forte partecipazione affettiva o di interesse (*mi bevo un caffè*);
13. la frase scissa (*sei tu che lo vuoi*);
14. *ci* attualizzante con *avere* e con alcuni altri verbi (*non c'ho tempo; non ci capisco niente*).

Come è evidente, si tratta di forme ormai parte dell'italiano contemporaneo che fino a non molto tempo fa erano accettate solo a livello di varietà diatopiche e “basse” (cfr. Sabatini 1985): “tratti sub-standard vengono attratti nella sfera dello standard, dando origine a quell'italiano neo-standard che abbiamo considerato come una delle varietà cardine nell'architettura dell'italiano contemporaneo” (Berruto 2012² [1987]: 67). Come già anticipato nel capitolo 1, la sintassi dell'italiano contemporaneo presenta, dunque, alcuni mutamenti dovuti spesso a quella che è stata definita come la “pressione del parlato” sulla lingua scritta, che ha portato ad una “forte spinta verso la semplificazione” (Ramat 1993: 35) con una riduzione delle possibilità offerte dal sistema e un generale alleggerimento delle strutture sintattiche (cfr. D'Achille 1993: 162-163). A questo proposito Ondelli e Romanini (2018: 187) affermano che l'italiano dell'uso medio può essere definito come

il risultato dell'accoglienza o della riammissione nei testi scritti di tratti linguistici tipici dell'italiano parlato (sulla persistenza del parlato nella scrittura nella storia della lingua italiana cfr. D'Achille 2006). Questa permeabilità si acutizza ulteriormente all'inizio del nuovo millennio, quando Antonelli (2011) registra una nuova fase di oralizzazione e abbassamento dei registri in tutti i tipi di testo proprio nel momento in cui, grazie all'avvento dei nuovi media, la scrittura torna a svolgere un ruolo centrale per la lingua nazionale anche nella comunicazione quotidiana.

È necessario precisare che tali tratti non sono innovazioni assolute dell'italiano contemporaneo, dove “il vecchio resiste o persiste nel nuovo”, come affermato da Nencioni (1987: 10). Essi risultano, infatti, ampiamente presenti nei volgari del territorio italiano anche in epoche molto antiche e se ne trova spesso un riscontro nelle altre lingue

Francesco Sabatini, che in una sua e-mail del 05/10/2020 ha confermato la correttezza della versione originaria del 1990.

romanze, sebbene nella maggior parte dei casi non è possibile risalire alle “prime attestazioni”, come si fa nel lessico (Renzi 2012: 40).

Quasi tutti i tratti morfosintattici che caratterizzano il nostro “italiano dell’uso medio” non sono innovazioni recenti: molti di essi sono attestati da vari secoli (in alcuni casi si può risalire fino al latino tardo) e dall’uso orale erano già passati anche nell’uso scritto, perfino in opere altamente letterarie. [...] [A]bbiamo fin da ora elementi sufficienti per affermare che la “novità” dell’italiano dell’uso medio riguarda sostanzialmente la validità della *norma*, non le caratteristiche profonde del *sistema*. Per dirla con i termini di E. Coseriu, vari tratti dell’italiano dell’uso medio erano già presenti da tempo nel *sistema* (o insieme di sistemi) che è alla base della lingua italiana ed erano stati accolti anche nelle varie *norme* scritte regionali (a volte di diffusione anche nazionale), ma non furono accolti in quella particolare norma, definibile come *supernorma*, che dal secolo XVI in poi ha dominato l’uso standard della lingua italiana: la norma letteraria di tipo bembesco, alla quale, in ultima analisi, si sono attenute le codificazioni grammaticali. (Sabatini 1985: 178, dove a seguire si trova il dettaglio dei tratti più antichi)

Ciò trova conferma nel titolo del saggio con cui Sabatini nel 1990 ritorna sulla questione, ossia, *Una lingua ritrovata: l’italiano parlato*, ove l’aggettivo *ritrovata* fa riferimento al fatto che l’uso parlato dell’italiano, gradualmente confluito nello scritto informale o di media formalità, non si basa sull’inserimento di nuovi elementi, bensì sul recupero nella norma di tratti linguistici esistenti da tempo nell’uso vivo, ma sanzionati a partire dal Cinquecento dalla norma stabilita dai grammatici.

Il concetto dell’*italiano dell’uso medio* di Sabatini fu non esente da critiche e precisazioni, tra cui si citano in particolare le obiezioni di Castellani, secondo il quale nell’insieme dei tratti elencati non ci sia “nulla che possa servire alla definizione d’una varietà nazionale d’italiano diversa dall’italiano normale o italiano senza aggettivi” (Castellani 1991: 256), riprese anche da Dardano (1994: 370-374) e Tesi (2005: 229). Oltre a questi, numerosi studiosi si sono occupati di precisare le caratteristiche di questa varietà, tra cui si citano Tavoni (2002), Simone (1993) e D’Achille (2010) per quanto riguarda la ristrutturazione del sistema verbale, Cortelazzo (2000a: 16-21) sui pronomi, Renzi (2012: 37-76) in particolare sulle dislocazioni e D’Achille (2010) e D’Achille, Proietti, Viviani (2005) sulla frase scissa. Tuttavia, a prescindere dalle diverse denominazioni date a questa varietà e dalle successive elaborazioni delle sue caratteristiche, è ormai accettata la convinzione che l’italiano abbia vissuto un processo di riconfigurazione da cui è derivato

un registro medio della lingua distinto dalla norma, ossia un *italiano dell'uso medio* o *italiano neo-standard* e distinto dallo standard tradizionale e che tale cambiamento non possa essere ignorato. La riflessione sulla definizione dell'*italiano dell'uso medio* è, poi, proseguita nei decenni successivi: come già affermato, lo stesso Sabatini nel 1990 è ritornato sul tema, riducendo i tratti da 35 a 14, per poi riprendere il tema nel 2016, definendo questa varietà come la vera spina dorsale dell'intero sistema linguistico. Nella sezione successiva si entrerà nel dettaglio dei singoli fenomeni grammaticali che la caratterizzano.

2.2. I tratti linguistici dell'*italiano dell'uso medio*

Dopo aver tracciato le linee generali dell'*italiano dell'uso medio*, verrà data una descrizione dei singoli tratti linguistici elencati nella lista del 1990, fornendo anche un confronto con la lista del 1985.

1) *lui, lei, loro* in funzione di soggetti

Sabatini (1985: 159) a questo proposito scrive che “l'uso di *egli, ella, essa, essi, esse* è ristretto al parlato che possiamo definire “celebrativo” e alle scritture di tipo argomentativo e situazionale”, mentre *lui, lei, loro* in funzione di soggetti “sono ormai la norma in ogni tipo di parlato, anche formale, e nelle scritture che rispecchiano atti comunicativi reali”. Tali forme in funzione di soggetto hanno dei precedenti rintracciabili nel fiorentino del XIII secolo, si attestano con una certa frequenza nel XIV secolo, affermandosi maggiormente nel secolo successivo. “La sostituzione di *egli* con *lui* è avvenuta attraverso una lunga fase di convivenza, cioè di uso alternativo, anche se non indifferente, delle due forme: tra queste, *egli* teneva la posizione stilistica alta, *lui* quella bassa” (Renzi 2012: 28)³. A questo proposito sempre Renzi:

³ A questo proposito Renzi (2012: 26-27) rimarca quanto segue: “Il cambiamento linguistico avviene attraverso alcune fasi. Poniamo che una lingua possieda due registri, uno alto, accurato e uno basso, spontaneo [...]. Le fasi di cambiamento linguistico possono essere rappresentate così: 1) nel contesto in cui si trova A, comincia a penetrare anche la nuova forma B, originariamente usata solo in un contesto diverso; 2) come conseguenza di 1), adesso A e B coesistono e si fanno concorrenza tra di loro, ponendosi generalmente a due livelli di registro diversi: la vecchia forma A si situa nel registro alto, la nuova forma B in quello più corrente, basso; 3a) A si indebolisce, e dopo un certo periodo di esistenza precaria, speso nello strato iperletterario della lingua, scompare; 3b) nella concorrenza tra A e B, A resiste allo sfidante B: quest'ultima si indebolisce e scompare.”

La concorrenza con *egli* dura da circa sei secoli: al tempo di Dante un *lui* soggetto al posto di *egli* sarebbe stato inconcepibile, nel Quattrocento la *Grammatica* di Leon Battista Alberti (1440 circa) lo propone invece come forma grammaticale, il che vuol dire che si usava correntemente a Firenze almeno già da qualche tempo. Nel 1525 Pietro Bembo nelle *Prose della volgar lingua* (III, 16) riesce a farlo regredire, almeno nell'uso colto, per un bel po'. (ivi: 114)

L'uso di *lui*, *lei*, *loro* come pronomi soggetto fu, infatti, condannato da Bembo, che raccomandava l'uso delle forme antiche dei pronomi, e ciò influenzò notevolmente la produzione di gran parte degli scrittori dell'epoca, ma non di tutti. Essendosi ormai affermate nel parlato, queste forme non scomparvero mai del tutto dal sistema linguistico, mostrando, invece, dal XVII secolo una graduale ascesa nella narrativa, che raggiunse il suo culmine con la revisione de *I Promessi Sposi*. In quest'opera Manzoni scelse o il soggetto sottinteso, conformemente alla tendenza moderna ad uno scarso utilizzo di pronomi soggetto, o la sostituzione delle forme soggetto con le forme oblique (cfr. Serianni 1989a: 190-192). I grammatici mantennero in proposito comunque un atteggiamento cauto, che cambiò nel corso del tempo.

L'avanzata di *lui* è avvenuta attraverso tappe successive, cioè occupando diverse posizioni sintattiche, da quella predicativa (per es. *Io non sono lui*), nella quale già il Bembo ammetteva la legittimità della forma obliqua, al pronome in isolamento (per es. *Chi è stato? Lui*), alla combinazione con un altro pronome o nome (per es. *Io e lui*, *Gianni e lui*) fino a raggiungere la posizione detta canonica, quella in cui il pronome precede direttamente il verbo (per es. *Lui dice che...*), che è l'ultima posizione a essere conquistata. (Renzi 2012: 28)

Nella *Grande Grammatica Italiana di Consultazione* (Renzi, Salvi, Cardinaletti 2001-2) si presentano come forme principali *lui*, *lei*, *loro* e le altre come forme alternative, definendole “pronomi rari nella lingua parlata, impiegati prevalentemente in alcune varietà della lingua scritta (scientifica, burocratica, letteraria) o in varietà della lingua parlata stilisticamente alte” (Cordin 2001, I: 549). Inoltre, si mette in evidenza anche come l'uso di *lui*, *lei*, *loro* non risponda solo a criteri stilistici, di registro o normativi, ma in molti casi si tratti di una scelta obbligata sintatticamente: “I pronomi *egli*, *ella*, *esso*, *essa*, *essi*, *esse* pur non essendo clitici, hanno proprietà particolari che li distinguono dagli altri pronomi liberi” (ivi: 550), poiché non possono essere usati nei seguenti casi: in coordinazione con altri sintagmi nominali: **Egli e Mario/ Mario e *egli vanno a Roma domani*; in isolamento, senza l'esplicitazione del verbo di cui sono soggetto: *Chi viene? *Essi*; se tra

il pronome personale e verbo si frappongono altri elementi lessicali: * *Essi, che sono arrivati per ultimi, non vogliono partire subito*; in posizione postverbale: *È arrivato *egli* e in funzione enfatica: **Egli è partito, non Piero; Proprio/solo *egli mi ha tradito*. Ciò trova conferma nello studio di Renzi (2012) e in alcune grammatiche per l'università e per la scuola secondaria superiore, citate a seguire a titolo di esempio, quali Salvi e Vanelli (2004), Fogliato (2015), Antonelli e Picchiorri (2016) e Serianni, Della Valle e Patota (2016).

In aggiunta a quanto esposto in precedenza (cfr. Cordin 2001), Fogliato (2015) precisa l'obbligatorietà dell'uso di *lui/lei/loro* quando si vuole mettere in evidenza il soggetto (*questo lo sostiene lui*), con gli avverbi e le congiunzioni *anche, proprio* ecc. (*neanche lui sa qualcosa, nemmeno lei si vestirà elegante*), con il verbo al gerundio o al participio (*arrivato lui, ci saremo tutti; parlando lei, tutti pendevano dalle sue labbra*). Antonelli e Picchiorri (2016), invece, oltre a ribadire la caduta in disuso delle forme soggetto, sottolineano come però molto spesso sia preferibile omettere del tutto il pronome (Ø) o sostituirlo con un nome (Esempi: *Manzoni lavorò alla seconda edizione dei Promessi sposi per molti anni. In questo periodo, egli si dedicò al romanzo quotidianamente* → *Manzoni lavorò alla seconda edizione dei Promessi sposi per molti anni. In questo periodo, (Ø) si dedicò al romanzo quotidianamente.* → *Manzoni lavorò alla seconda edizione dei Promessi sposi per molti anni. In questo periodo, lo scrittore..*). Ciò lascia intendere come l'eredità delle forme soggetto originarie sia stata raccolta non solo dalle forme oblique, ma anche dal pronome sottinteso.

Sulla base di quanto esposto, poiché l'uso dei pronomi *lui/lei/loro* in funzione di soggetto è ormai indiscusso nell'italiano contemporaneo e, dunque, non sussiste più alcun dubbio sulla reperibilità di tali forme nello scritto di media formalità/informalità, il fenomeno è stato escluso dalla ricerca basata su corpus dei tratti elencati da Sabatini (1990) (cfr. capitolo 5).

2) *gli* come forma dativale onnivale, ossia l'impiego di *gli* nella forma dativale al posto di *le (a lei)* e *loro (a loro)*

L'uso generalizzato del pronome *gli*, per entrambi i generi al posto delle altre forme pronominali di terza persona singolare e plurale è definito da Sabatini "accettabile nell'uso medio, soprattutto parlato, ma anche scritto. S'incontra spesso nei giornali e nelle riviste; è quasi normale nella narrativa" (1985: 158). A giustificazione di quest'uso generalizzato,

Sabatini ne richiama le origini storiche. Questa forma pronominale, originariamente *li*, deriva dalle forme del dativo latino singolare e plurale (*ILLI* e *ILLIS*), valide per tutti e tre i generi e questo uso indifferenziato per genere e numero rimase attivo nell'intero sistema linguistico italoromanzo (ossia, dialetti e lingua letteraria) come pronome atono e, dunque, utilizzabile come clitico nelle varie posizioni. “*Li* era originariamente anche la forma del f., essendo *ILLI* l'unica forma sg. Ma accanto a *li* si è sviluppato il clitico *le* da una forma latina analogica, *ILLAE*” (Vanelli 1999: 115). Oltre alla restrizione dell'ambito d'uso con l'emergere di *le*, a partire dal XII secolo si affiancò l'uso di *loro*, ritenuto un gallicismo, avente originariamente un uso sia tonico (come nell'italiano moderno, in cui il suo uso senza preposizione è quasi esclusivamente postverbale, esempio: *dico loro*) che atono proclitico. Con il passare del tempo questo uso proclitico è stato abbandonato, per cui l'unica forma clitica rimase *gli*, nel suo uso generalizzato per il maschile e il femminile fino al XVI secolo. Successivamente Manzoni ripristinò *gli* per il plurale maschile al posto di *loro*, ma non per il femminile singolare e plurale. Da ciò Sabatini suggerisce che l'uso di *gli* nel parlato come forma unica corrisponda al riemergere del suo antico uso generalizzato. Serianni nella sua grammatica (1989b: 249) afferma che

un punto molto delicato è costituito dall'uso della forma dativale di 6a persona: *gli* o *loro*? L'atono *gli* [...] – largamente attestato in tutti i secoli della nostra lingua – appartiene al registro familiare; il parlato formale e la massima parte dello scritto (tecnico-scientifico, letterario e in una certa misura anche giornalistico) preferiscono la forma *loro*.

Per quanto riguarda l'uso di *gli* per *le*, Serianni lo segnala invece come “decisamente da evitare anche nel parlato colloquiale” (*ibidem*) a prescindere dai precedenti letterari illustri da lui stesso menzionati (Boccaccio, Machiavelli, Carducci, Verga e Moravia). Tornando sul tema anni dopo, Serianni ha rilevato che *loro* è una forma “ormai minoritaria anche nello scritto” (2006: 12), ma che “la distinzione *gli* maschile / *le* femminile si è addirittura consolidata nel corso degli ultimi decenni” (ivi: 24), costituendo, quindi, un’ “evoluzione apparentemente bloccata nonostante la prevedibilità del percorso” (*ibidem*), ossia nonostante la possibilità che *gli* potesse arrivare a sostituire anche la forma *le*, come avvenuto per *loro*.

Sulla base di ciò, si può giungere alla conclusione che l'affermazione dell'uso di *gli* al posto di *loro* come forma unica per il plurale sia avvenuta anche sulla base di motivazioni sintattiche: contrariamente a *loro*, *gli* presenta il vantaggio di fungere sia da proclitico (*gli*

dico) che da enclitico (*digli*), caratteristiche condivise con forma *le* (esempi: *le dico; dille*). Dunque, è possibile ipotizzare che l'uso di *gli* al posto di *le* sia ancora percepito come tratto substandard, come confermato, ad esempio, dallo studio di Ondelli e Romanini (2018)⁴, proprio perché non vi è alcun vantaggio sintattico nella sostituzione di *le* con un'altra forma pronominale avente le stesse caratteristiche e dunque si tratterebbe di una deviazione del tutto arbitraria rispetto alla norma.

3) la combinazione di una preposizione e di un articolo partitivo, ossia l'impiego del partitivo preceduto da preposizione (*con degli amici*)

Circa questo tratto grammaticale, si riporta la descrizione fatta da Sabatini nella prima versione della lista del 1985:

Sconsigliata dalle grammatiche, è in realtà di uso frequentissimo proprio nella lingua media (non in quella molto formale, né in quella regionale o popolare): Condiscilo con dell'olio crudo; È andato in gita con degli amici che non conosco. [...] Nella varietà più alta e più bassa s'incontrano altre soluzioni, come l'eliminazione del partitivo o la sua sostituzione con altri tipi (alcuni, taluni, certi, espressioni composte): con olio o con un po' d'olio; con amici o con alcuni (o certi) amici; su un gruppo di passanti." (Sabatini 1985: 159-160)

4) la dislocazione dell'elemento tematico a sinistra o a destra della parte rematica della frase, con la ripresa funzionale atona, ossia l'impiego di dislocazioni a sinistra (*quel libro l'ho già letto*), a destra (*l'ho già letto, quel libro*) della parte rematica della frase con la ripresa pronominale atona

Queste strutture rientrano nella classe che i linguisti definiscono "costruzioni sintatticamente marcate", cioè quelle frasi in cui l'ordine delle parole non segue la sequenza prevista dal sistema linguistico di riferimento, che nel caso dell'italiano è la sequenza SVO, ossia soggetto – predicato verbale – oggetto diretto – altri sintagmi, detta anche "ordine non marcato" (cfr. Gossen 1954, Berruto 1985b, Benincà, Salvi, Frison

⁴ Nel loro studio hanno sottoposto ad un campione di 150 persone, provenienti da tutto il territorio nazionale e prive di competenze specifiche nelle questioni linguistiche, dei brevi estratti di articoli tratti da *Corriere della Sera* e da *Repubblica* (2015-2016), contenenti alcuni dei tratti grammaticali elencati da Sabatini (1985) e Berruto (1987) nella definizione dell'italiano dell'uso medio/ neostandard. L'obiettivo di tale ricerca era di indagare il grado di accettabilità percepita dai rispondenti rispetto a tali tratti. A questo proposito, l'uso del pronome dativale maschile *gli* in sostituzione del femminile *le* è generalmente rifiutato dai partecipanti.

1988, Berretta 1995 e Ferrari 2003) e che sono un tratto caratteristico dell'oralità (Renzi 2012: 41). Nella versione del 1985 della lista Sabatini definisce questo costrutto come “frase segmentata, cioè con tematizzazione (a sinistra o a destra) del dato “noto” assunto come “tema”, e ripresa di esso mediante un pronome nella frase che predica l’ “informazione nuova”, cioè il “rema”” (ivi: 162). La sua funzione è l'enfasi, ossia la focalizzazione dell'informazione che rappresenta il dato “nuovo” nell'atto comunicativo (ivi: 161).

Quando un costituente della frase viene spostato a sinistra rispetto alla sua posizione non marcata, si parla di “dislocazione a sinistra”⁵. Il costituente dislocato presenta un legame con il resto della frase tramite una ripresa pronominale (*Le donne, io proprio non le capisco*, in Cignetti 2006: 208 vs *io proprio non capisco le donne*, secondo l'ordine non marcato) e/o tramite la presenza di una preposizione, nel caso dell'oggetto indiretto (esempio: *A Elisabetta, non parlo da secoli, ibidem*).

Lambrecht (1994) individua gli ambiti d'uso principali della dislocazione a sinistra: presentare un costituente della frase come l'elemento su cui verte il contenuto dell'enunciato, imporre un nuovo tema, mettendo in risalto un costituente e attirando l'attenzione dell'interlocutore su di esso e, infine, a risolvere casi di ambiguità, quando vi è più di un costituente che potrebbe fungere da punto focale del discorso (si veda anche Berruto 1985b). In generale, si evince l'intento di creare nel testo una gerarchia delle informazioni, realizzata attraverso l'ordine marcato dei costituenti della frase. In aggiunta, la dislocazione a sinistra può essere utilizzata come strumento di coesione testuale, in particolare quando è abbinato ad espressioni anaforiche, come i dimostrativi: “Segue e cura – fin dai primi giorni – l'educazione della primogenita [...]. E di questa esperienza stende un resoconto [...]” (Soletti, in Cignetti 2006: 212).

La frequenza d'uso nella lingua scritta varia a seconda del genere testuale: rare nei testi specialistici (Mortara Garavelli 2001: 86 e sgg.), sono molto diffuse nella prosa

⁵ La dislocazione a sinistra rientra tra le “costruzioni sintatticamente marcate a sinistra” (cfr. Ferrari 2003: 145sgg). ossia quelle costruzioni in cui un costituente della frase, che normalmente avrebbe una posizione postverbale, occupa invece una posizione preverbale. Tra queste vi sono, oltre alla già citata dislocazione, anche le costruzioni con tema sospeso come l'anacoluto (esempio: *Elisabetta, non penso più a lei da tempo*), ove tra l'elemento marcato a sinistra e il resto dell'enunciato non vi è un legame sintatticamente esplicito, e la topicalizzazione (esempio: *MIA CUGINA ho visto [non mia sorella]*), che pone l'enfasi (intonativa, nel parlato) sull'elemento dislocato (gli esempi sono tratti da Cignetti 2006: 207 e sgg.). Benincà, Salvi e Frisson (1988: 141-143) aggiungono anche “l'anteposizione anaforica”, che si verifica quando il costituente dislocato non è né ripreso da un pronome, né enfatizzato (esempio: *la stessa proposta fece poi il partito di maggioranza*). Tra le costruzioni marcate vanno inserite anche le frasi scisse (cfr. il punto 13 della lista di Sabatini 1990 e Roggia 2006) e le frasi presentative (De Cesare 2006), ad esempio: *c'è una ragazza che ha scritto un annuncio*.

giornalistica (Ferrari 2003: 148), nei testi pubblicitari (*Oil of Olaz. L'ho scelto da sempre*, in Cignetti 2006: 212). In aggiunta, oltre a svolgere le funzioni di tematizzazione già menzionate, la dislocazione a sinistra sta ormai sostituendo la forma passiva, che è percepita dai parlanti come di registro troppo elevato (Renzi 2012: 42).

Nel caso, invece, in cui un costituente della frase è dislocato a destra rispetto alla sua posizione nell'ordine non marcato della frase si parla di “frasi dislocate a destra” (della vasta bibliografia sul tema si citano: Sornicola 1981, Benincà, Salvi, Frisson 1988, Cresti 2000, Ferrari 2003). Se l'elemento dislocato è un oggetto diretto, è presente normalmente un pronome clitico che anticipa l'elemento dislocato (esempio: *io proprio non le capisco, le donne*). Questa struttura serve ad introdurre informazioni aggiuntive con un tono, per così dire, “confidenziale” rispetto al contenuto dell'elemento dislocato, che è qualcosa su cui verte già il discorso.

Per quanto riguarda gli ambiti d'uso, la dislocazione a destra ha “funzioni meno informativo-coesive, per dir così, e più emotivo-espressive, quali l'isolamento di un tema in funzione di appendice (*detachment* o *extra-position*: Ferrari 1999); il ripensamento (*afterthought* o *self-repair*: Tomlin 1986, Berretta 1986, 1992); ammiccamento o confidenzialità (*camaraderie*) tra interlocutori (Berruto 1986, Ferrari 2008)” (Rossi 2016: 847).

Nel suo studio sulle dislocazioni Rossi prende in esame sia la prosa giornalistica, esaminando gli articoli presenti nel sito www.repubblica.it dal 10/1/2014 all'indietro, sia la narrativa, per la quale sono stati considerati i 100 romanzi vincitori e finalisti del Premio Strega dal 1947 al 2006. Sebbene la dislocazione a destra sia “più diffusamente avvertita come struttura ridondante, e dunque sanzionata negativamente dal sistema scolastico” (Ondelli, Romanini 2018: 193), dallo studio di Rossi emerge che la sua frequenza d'uso, contrariamente ai dati raccolti da Berruto 1985b, 1986 e negli altri studi già precedentemente citati sulle strutture marcate, non risulta sempre inferiore alla dislocazione a sinistra: la narrativa è molto più incline alla dislocazione a destra, mentre invece la prosa giornalistica “preferisce la [dislocazione a sinistra] e soprattutto la frase scissa, in quanto tecniche predilette per il dinamismo comunicativo e di introduzione o riproposizione di nuovi temi nel discorso (Bonomi 2002: 212-215 e 321-323, Gatta 2014: 339). Fatte queste precisazioni, sul computo totale la [dislocazione a destra] è oltre 2/3 della [dislocazione a sinistra], dunque la percentuale risulta invertita rispetto ai dati raccolti finora da Berruto e altri.” (Rossi 2016: 850).

Da una valutazione a campione emerge un'elevatissima occorrenza di dislocazioni a destra in brani di discorso diretto (con, invece, presenze molto inferiori della dislocazione a

sinistra), da cui si evidenzia che la presenza di dislocazioni a destra si innalza proprio nei contesti che simulano il dialogo (ivi: 851), ossia nel teatro, nel cinema e anche nei testi poetici in cui il poeta instaura un finto dialogo col lettore, mentre si riduce nella prosa giornalistica (che generalmente predilige la dislocazione a sinistra), dov'è utilizzata solo in presenza di discorso riportato e interrogativo. Il suo uso sembra rispondere, dunque, all'intento di coinvolgimento dell'interlocutore (ivi: 851-852).

Per quanto riguarda gli ambiti d'uso della dislocazione a destra, sempre dallo studio di Rossi emerge che sono molto più variegati di quelli finora attribuiti a questa struttura negli studi precedenti:

- a) usi grammaticalizzati o in via di grammaticalizzazione (ossia, usi ormai integrati o in via di integrazione nella norma e, dunque, non più percepiti come forme marcate. Ad esempio, i costrutti completivi espliciti dislocati come *lo so/sai/sa che* e *lo vedi che* sono le forme marcate dei costrutti completivi *so/sai/sa che* e *vedi che*; il costrutto *chi lo sa se* vs *chi sa se*, ecc.);
- b) autocorrezione e glossa esplicativa (Rossi 1999, 2005);
- c) ancoramento deittico (ossia, accentuazione della forza deittica degli enunciati. Esempio: *Ma la vedi lí, la porta? aveva aggiunto indicando con un gesto veloce l'uscita verso il ballatoio. La vedi? Apri, esci, chiudi e non tornar piú.* (Testori in Rossi 2016: 853);
- d) ancoramento emozionale (ripetizione) (si veda l'esempio precedente);
- e) modalità interrogativa (coinvolgimento dell'interlocutore, ammiccamento) (si vedano l'esempio precedente e il seguente: *Ecco, lo vedi a cosa serve la tua superbia? Lo vedi?*, Testori, in ivi: 854);
- f) messa in evidenza del centro di interesse comunicativo ed emotivo: a questo proposito vengono citati vari esempi tratti dalla pubblicità, tra cui *Noi, lo facciamo così / da cinquant'anni* pubblicità del 2013 di un panificio di Messina, ove l'elemento dislocato a destra (*noi, lo facciamo così, [il pane]*) è di fatto dislocato iconicamente nell'immagine sul cartellone;
- g) coesione, una funzione, ritenuta negli studi precedenti a quello di Rossi esclusiva della dislocazione a sinistra (esempio: *Io non lo vedo quasi mai il mio papà, perché egli lavora tutta la settimana a Torino e viene a trovarmi solo la domenica. Mio papà è un uomo molto alto e molto ordinato; e quando arriva, la domenica, mi deve sgridare perché se fossi da solo non mi laverei mai* (Romano, in ivi: 855).

Come già anticipato in apertura di capitolo, Sabatini (1990: 97) inserisce all'interno della tipologia delle dislocazioni anche le strutture come *a me mi piace leggere, di pane non ne ho più* (ossia, il doppio partitivo) e l'anacoluto⁶.

Le strutture come *a me mi* si basano sulla ripetizione della forma pronominale in forma tonica (*a me*) e atona (*mi*). In queste il primo elemento (*a me*) equivale a *per quanto riguarda me* e rappresenta il tema sul quale poi si svolge il discorso o rema (*mi piace*). Nella presente trattazione saranno classificate all'interno del fenomeno delle dislocazioni, nello specifico come dislocazioni a sinistra dell'elemento pronominale e saranno oggetto di ricerca tutte le forme pronominali toniche del sistema pronominale, precedute dalla preposizione *a* e seguite dalla loro forma atona (*a me mi, a te ti, a lui gli...* ecc.) (cfr. capitolo 4).

Il fenomeno del doppio partitivo rientra anch'esso nella categoria della dislocazione a sinistra e consiste nell'anticipazione di un sintagma preposizionale, la cui testa è costituita dalla preposizione *di* (esempio: *di pane*), successivamente ripreso attraverso la particella pronominale *ne*. Esempio: *di pane non ne ho più*. “Per il *ne* si nota la frequente ricorrenza, senza più valore pronominale pieno e soltanto come rafforzativo, con alcuni verbi, soprattutto *fare, parlare, sapere, pensare, occuparsi, preoccuparsi, interessarsi, essere convinto*. Sono frequentissime frasi come *Di lui non me ne parlare; Di questa faccenda non ne voglio più sapere niente*” (Sabatini 1985: 163).

Un estremo della categoria delle strutture con tematizzazione è l'anacoluto (esempio: *Giorgio, non gli ho detto nulla*), dove il tema è solo un'enunciazione, un *nominativus pendens* che non ha un raccordo sintattico ma solo semantico con il rema, di cui, come riportato da Sabatini (*ibidem*), sono rintracciabili esempi anche in Dante.

⁶ Entrando nel dettaglio, nella versione del 1985 Sabatini riunisce sotto un'unica categoria quattro procedimenti di natura sintattica che contribuiscono al “fenomeno dell’“enfasi” tramite la tematizzazione dell'elemento noto a destra o a sinistra (p.161): la frase scissa (cfr. punto 13 dell'elenco del 1990), la posposizione del soggetto al predicato (punto 11), la frase segmentata, ossia la dislocazione a destra e a sinistra, includendo anche le costruzioni come *a me mi* e il doppio partitivo (punto 4), e infine l'anacoluto (punto 4). Il motivo di questo accorpamento è dato dalla condanna generalizzata da parte delle grammatiche tradizionali delle costruzioni con ripresa pronominale del tema, definite “pleonastiche”, e delle strutture con inizio “è a lui che”, definite come “francesismi”, una posizione che, di fatto, distoglie l'attenzione di chi insegna la lingua dal fenomeno della messa in prospettiva comunicativa (*ibidem*). Sulla condanna generalizzata delle grammatiche rispetto a questi costrutti, si veda anche Prandi (2020: 174-175).

5) il *che* polivalente, sia con valore temporale, sia con altri valori (spaziali, personali).

Sabatini (1985: 164) nota come il *che*, originariamente pronome relativo con il significato di *di cui, in cui, a cui*, si sia poi trasformato in un connettivo generico con molte funzioni, il cui uso “viene accolto con maggiore larghezza” (*ibidem*, cfr. Fiorentino 2010):

[Mentre] nell’italiano standard la congiunzione *che* è tipica di alcune frasi subordinate, quali le oggettive, le soggettive, le dichiarative e le relative (nei casi in cui l’antecedente del pronome relativo è un soggetto o un complemento oggetto), al di fuori dell’uso standard il *che* viene utilizzato anche come forma pronominale generica e come introduttore generico di subordinate che nell’italiano standard prevedono, invece, congiunzioni subordinanti semanticamente più precise. (Zingaro 2018)

Riallacciandosi agli studi di Alisova (1972: 254-268), Cortelazzo (1976: 93-98), Sornicola (1981: 61-74), Sabatini (1985: 164) articola la polivalenza della congiunzione *che* in quattro categorie principali:

- a) il *che* con valore temporale, ossia come equivalente dei più formali *in cui, dal momento in cui, nel momento in cui*. Tra gli esempi citati da Sabatini: *La sera che ti ho incontrato; è un’ora che ti aspetto;*
- b) il *che* che congiunge le due parti di una frase scissa (cfr. punto 13 della lista del 1990): *è qui che ci siamo incontrati;*
- c) il *che* con apparente funzione di soggetto o oggetto, contraddetta da una successiva forma pronominale che ha la funzione di complemento indiretto: *Quel mio amico che gli hanno rubato la macchina;*
- d) il *che* sostitutivo di una congiunzione finale, consecutiva o causale: *aspetta che te lo spiego; vieni, che ti pettino.*

Le grammatiche, come riportato da Sabatini (ivi: 180), definiscono “abnorme” questo tratto grammaticale. Successivamente D’Achille (1990: 212) ne ammette nella norma il suo uso temporale, condannandone, invece, gli usi come relativo obliquo.

6) il connettivo *per cui* (derivato per ellissi da *motivo, fatto, ragione per cui*, oppure per generalizzazione di casi con normali antecedenti nominali) in funzione di congiunzione causale-consecutiva.

L'uso di questo nesso relativo come congiunzione consecutiva risulta attestato sin dal Trecento (Sabatini 1985: 165).

7) il semplice *cosa?* interrogativo diretto e indiretto, ossia il suo uso al posto di *che cosa?*

Rispetto a questo fenomeno, Sabatini spiega che l'interrogativo *che cosa* ha perso terreno rispetto all'affermazione a livello nazionale del solo *cosa*, di provenienza settentrionale, e del solo *che*, di provenienza meridionale e predominante da Roma in giù (ivi: 165).

8) forme d'ingresso o connettivi testuali come *allora, comunque* e anche *ma, e, o* iniziali di discorso o di enunciato

A questo proposito si vedano gli studi di Sabatini (1997) e Mandelli (2006).

9) l'indicativo invece del congiuntivo nei seguenti cinque casi: dopo i verba putandi, ossia i verbi di opinione (*credo che hai ragione*), dopo una dichiarativa negativa (*non credo che hai torto*); in un'interrogativa indiretta (*gli chiesi se poteva aiutarmi*); nel periodo ipotetico dell'irrealtà (*se me lo dicevi, ci pensavo io*)⁷; nelle relative restrittive (*sei l'unico che parla arabo*)

Nella versione della lista del 1985 di Sabatini si precisa che già all'epoca si registrava una notevole tendenza alla sostituzione del congiuntivo con l'indicativo nei costrutti sopra indicati (ivi: 166-167), giudicati normali nel parlato di media formalità e aventi tutti alle spalle una secolare tradizione letteraria: in particolare il caso delle ipotetiche ha dei saldi precedenti storici in Dante, Machiavelli e Manzoni (ivi: 167). Questi precedenti storici, uniti alle numerose norme d'uso del congiuntivo, che non chiariscono del tutto l'alternanza con l'indicativo, hanno dato vita ad un'annosa questione che “ricorre più frequente non solo nelle discussioni di linguisti e insegnanti ma anche nelle quotidiane riflessioni metalinguistiche” (Marani 2016: 28) e che ha trasformato l'uso del congiuntivo in un argomento largamente studiato. Già nel 1987 Berruto affermava che “la ‘morte’ o ‘sparizione’ del congiuntivo” era diventata “un *topos*, a proposito delle vicende odierne del

⁷ Si veda a questo proposito lo studio di Mazzoleni 1992.

sistema verbale” (p. 70), proseguendo nel corso dei decenni successivi con vari studi che, invece, ne hanno dimostrato la resistenza⁸. In risposta a una lettera pubblica, scritta dagli alunni di una scuola di Treviso, in cui si chiedeva ai giornalisti di rispettare la grammatica, lo stesso Sabatini invitò alunni e insegnanti ad approfondire la storia della lingua, invece di chiudersi in “gabbie di norme rigide”, così da scoprire che l’alternanza del congiuntivo con l’indicativo è “una tendenza praticata da secoli”⁹, sulla quale sono stati prodotti numerosi studi¹⁰, con l’intento di fare chiarezza nell’insegnamento dell’italiano sia ad italofoeni nativi e sia a stranieri. Per quanto riguarda l’insegnamento della lingua italiana a stranieri si vedano, ad esempio, gli studi di Moneti 1988, Malinowska 1996, Maiden 2005, Manili 2015 e Marani 2016. Questi ricercatori hanno posto in particolare la loro attenzione sulle questioni relative a norma e uso e all’impiego del congiuntivo da parte degli studenti, nel tentativo di trovare una soluzione alla difficoltà messa in evidenza dal già citato studio di Manili, che afferma che “l’uso del congiuntivo nell’italiano contemporaneo non offre agli apprendenti non nativi indizi chiari e univoci delle sue funzioni” (2015: 154), in particolare in riferimento al tema dell’alternanza congiuntivo/indicativo.

A questo proposito, come riportato da Ondelli e Romanini (2018: 197), la scelta di utilizzare nelle subordinate l’indicativo al posto del congiuntivo può basarsi su un’interpretazione semantica (indicativo come “modo della certezza” e congiuntivo come “modo dell’incertezza”: cfr. Schneider 1999) oppure sintattica (congiuntivo come espressione di dipendenza, oppure usato per innalzare il registro. Si vedano a questo proposito Sgroi 2013, Lepschy e Lepschy 1988: 204 e Renzi 2012: 52-53, secondo i quali la scelta congiuntivo-indicativo è di tipo stilistico).

Nella riflessione sull’alternanza dell’uso dell’indicativo e del congiuntivo, questa ricerca si riallaccia a quanto espresso da Prandi (2002). Sebbene la posizione tradizionale circa il valore del congiuntivo sia univoca, ossia che l’opposizione tra indicativo e congiuntivo sia da ricondursi all’opposizione tra realtà e non realtà (cfr. le grammatiche Battaglia, Pernicone 1951, Sorrento 1951, Dardano, Trifone 1985, Fogarasi 1969 [1983],

⁸ Già Serianni (1986: 59-61) aveva constatato la resistenza del congiuntivo, in particolare nelle frasi complete, nella narrativa rosa e in fotoromanzi e fumetti. Restando nell’ambito della narrativa, la scarsa sostituzione del congiuntivo con l’indicativo è documentata anche in Bonomi (1996; 1997), Berisso (2000), Bonomi e Mauroni (2003), De Santis (2003) e nella letteratura per l’infanzia (Pesce 2003). Circa i quotidiani, cartacei e *online*, la buona tenuta del congiuntivo è verificata da Bonomi (2002), mentre Antonelli (2007: 130) ha notato un ritorno significativo all’uso del congiuntivo nei testi delle canzoni dei 15 anni precedenti.

⁹ Dall’articolo *Un appello dalle elementari per la difesa del congiuntivo* (La Stampa, 20/02/2002): www.repubblica.it/online/societa/congiuntivo/congiuntivo/congiuntivo.html (consultato il 23/09/2020).

¹⁰ Si vedano gli atti del convegno *Intorno al congiuntivo* a c. di L. Schena, M. Prandi, M. Mazzoleni (2002) e la bibliografia ivi contenuta.

Sensini 1990 e gli studi di Wandruszka 1991, Giorgi, Pianesi, 1997), nella sua argomentazione Prandi (2002), di cui a seguire si citano gli esempi utilizzati, mette in evidenza come, in realtà, la sola presenza del congiuntivo non è condizione sufficiente per esprimere la non realtà di una proposizione (esempio: *Ho ottenuto che Luca si assentasse dal lavoro per un'ora*, ove l'oggetto di *ottenere* è qualcosa che è realmente accaduto) e, viceversa, l'uso dell'indicativo non determina necessariamente la fattualità della proposizione (esempio: *Ho sognato che mia suocera aveva vinto un miliardo al lotto*, ove il contenuto della proposizione oggettiva – il contenuto del sogno – non è reale). Il fatto che sia possibile costruire frasi come quelle presentate negli esempi secondo Prandi dipende da quanto segue:

Il congiuntivo non ha un valore – o una costellazione coerente di valori – che accompagnerebbe tutti i suoi usi. Ci sono casi in cui il congiuntivo ha un suo valore, che rientra in linea di massima nel ventaglio di valori che gli sono attribuiti da sempre, e dei casi in cui è inutile cercare di identificare un valore perché questo valore non c'è. (ivi: 32-33).

Per spiegare questo punto, Prandi fa una riflessione sulla struttura delle frasi complesse e, in particolare, sul loro regime di codifica.

Si parla di codifica di tipo relazionale quando all'interno della frase complessa sono presenti relazioni grammaticali, o categorie funzionali, in particolare, il soggetto e l'oggetto diretto. Un esempio è la relazione che lega una frase principale e una subordinata soggettiva o oggettiva. Ad esempio l'oggettiva: *Spero che Giovanni venga*.

Si parla invece, di codifica puntuale, nel caso in cui la codifica non si basa su relazioni grammaticali, ma su una congiunzione e coinvolge la complementazione del nome, la connessione transfrastica e i ruoli periferici del processo: *poiché* codifica una relazione causale, *dopo che* e *prima che* codificano relazioni temporali, *benché* e *anche se*, una relazione concessiva ecc.

Nel caso della codifica relazionale, in presenza di relazioni soggettive e oggettive, la realtà o l'irrealtà – nei suoi diversi gradi – della frase subordinata è determinata dal verbo principale. Ad esempio in *Suppongo che tu sia al corrente delle nostre decisioni* il contenuto incerto della subordinata è dato dalla semantica del verbo principale *supporre* e non dalla presenza del congiuntivo. La funzione del congiuntivo, dunque, è quella di codificare una relazione grammaticale in sé vuota di contenuto: la relazione tra il verbo

principale e il suo contenuto in posizione di oggetto. Per frasi come quella citata nell'esempio si può affermare quanto segue:

In caso di alternanza tra indicativo e congiuntivo in presenza di una stessa accezione di uno stesso verbo, cambia il registro stilistico dell'espressione, ma non cambia il valore modale della completiva, dato che non cambiano né il contenuto del verbo principale né la relazione grammaticale, che sono gli unici fattori pertinenti¹¹ (ivi: 38).

Ciò si evince anche da alcuni esempi citati dall'autore nel suo stesso studio (ivi: 38-39): confrontando l'esempio precedente, *Suppongo che tu sia al corrente delle nostre decisioni*, con *Gli inquirenti suppongono che il suo compito era quello di mantenere [...]*, emerge che la scelta dell'indicativo è giustificata da una differenza di registro, ma non influisce sul contenuto della subordinata. In aggiunta, si prendano in esame i seguenti esempi: 1) *Spero che Giovanni venga*; 2) *Spero che Giovanni verrà*; 3) *Spero che Giovanni viene*. Dall'esempio 1 all'esempio 3 si assiste ad un graduale abbassamento del registro, tuttavia l'assenza del congiuntivo non cancella il contenuto del verbo *sperare*: dato come incerto nel caso 1, non diviene certo nel caso 3.

Una riflessione simile può essere fatta circa i verbi principali che esprimono emozioni: *Temo che Giovanni sia partito* vs *Mi dispiace che Giovanni sia partito*: “Posto che *temere* e *dispiacere* reggono entrambi il congiuntivo, ad esempio, il contenuto della proposizione oggettiva di *temere* è dato come non reale, mentre il contenuto della proposizione soggettiva di *dispiacere* è presupposto come vero” (*ibidem*). Ciò ovviamente non significa che il congiuntivo presente nei due esempi abbia valori opposti, bensì che la presenza del congiuntivo non influisce sul valore reale o non reale del contenuto della completiva. I due esempi appaiono incoerenti solo se si pretende di giustificare i due valori opposti della subordinata non a partire dal contenuto di verbi principali distinti, ma dalla presenza del modo congiuntivo, secondo quanto prescritto tradizionalmente dalla norma:

¹¹ Ciò si contrappone ai casi, elencati da Serianni (1989: 556) e ripresi da Schmitt-Jensen (1970), di verbi per i quali l'alternanza della reggenza dell'indicativo e del congiuntivo determina un cambiamento di accezione di significato. A questo proposito Prandi (2002: 39, nota 3): “*Ammettere*, ad esempio, regge l'indicativo con il valore constativo di “rendersi conto, riconoscere” e il congiuntivo con il valore di “fare un'ipotesi”: *Ammetto che hai fatto il tuo dovere*; *Ammettiamo che non piova*. Un'alternanza analoga si ha con l'uso constativo e l'uso direttivo dei verbi di comunicazione: *Gli ho telefonato che era ora di partire*. *Gli ho telefonato che partisse quanto prima*. In alcuni casi l'alternanza è coerente con una differente disposizione del soggetto enunciante: *Ti ho detto che Luigi è partito*; *Dicono che Luigi sia partito*. Tutte le volte che la reggenza ammette un'alternanza coerente con una distinzione modale, la distribuzione del modo non è arbitraria ma motivata dall'affinità elettiva di cui si è detto”, ossia l'indicativo come modo della realtà e il congiuntivo come modo dell'irrealtà.

Se ammettiamo che il valore modale della proposizione completiva è controllato dal verbo indipendentemente dalla presenza del congiuntivo, il problema rappresentato dai predicati emozionali e valutativi, e da predicati come *ottenere* semplicemente si dissolve. [...] Di fronte a esempi di questo genere, la domanda pertinente non è: “Che differenza di valore c’è tra congiuntivo e indicativo?”, ma “Fino a che punto le nostre orecchie sono disposte a accogliere costruzioni che usano l’indicativo invece del congiuntivo in dipendenza da certi verbi, come in *Spero che Giovanni viene?*”. Le mie stridono, ma non è questo il punto. Resta il fatto che le costruzioni più avventurose dell’italiano popolare non influiscono sullo statuto modale della completiva, alla sola condizione che la relazione grammaticale tra verbo e completiva rimanga trasparente. (Prandi 2002: 39)

Per quanto riguarda la codifica puntuale, in alcuni casi la scelta tra indicativo e congiuntivo è motivata: *poiché* e *dopo che* reggono l’indicativo in quanto la verità della proposizione è implicitata; *prima che* regge il congiuntivo dato che al momento della realizzazione dell’evento principale, la verità della subordinata è sospesa. Tuttavia, questa osservazione non vale per la relazione concessiva, in cui sia l’azione della frase principale, sia quella della frase subordinata sono entrambe vere e questa caratteristica non viene invalidata dal fatto che per la sua costruzione possono essere utilizzate sia congiunzioni che reggono l’indicativo, come *anche se* (esempio: *Anche se pioveva sono uscito*), sia congiunzioni che reggono il congiuntivo, come *sebbene*, *quantunque* o *benché* (esempio: *Sebbene piovesse sono uscito*). In entrambi i casi, la realtà della pioggia non è messa in dubbio. Una riflessione analoga può essere fatta per il periodo ipotetico:

In italiano standard, l’alternanza tra indicativo e congiuntivo nel periodo ipotetico è certamente significativa. La forma (10) esprime irrealtà o, in subordine – in presenza di un contesto incoerente con l’inferenza sollecitata di irrealtà – possibilità di grado basso nel passato. La forma (11) esprime possibilità di grado basso nel presente-futuro o, in subordine, irrealtà nel presente. La forma (12) esprime possibilità di grado alto:

- (10) Se Anna avesse preso il treno delle tre sarebbe già arrivata a casa
- (11) Se Anna fosse in casa potremmo passare a salutarla
- (12) Se Anna è in casa passiamo a salutarla

Come appare evidente, la frontiera tra i modi non coincide con la frontiera critica tra non realtà e realtà, ma separa, all’interno della non realtà, gradi diversi di possibilità. Anche

quando partecipa attivamente alla codifica del contenuto, l'opposizione tra congiuntivo e indicativo non si lascia dunque ridurre necessariamente alla dicotomia tra realtà e non realtà. (Prandi 2002: 41)

Per concludere, l'ipotesi di Prandi è la seguente: nei casi in cui l'uso del congiuntivo è imposto, il contenuto di realtà o di gradi diversi di possibilità della frase dipende in effetti non dalla sua presenza nella frase, ma da fattori diversi, quali il contenuto del verbo principale e la presenza di una relazione grammaticale. Il congiuntivo ha invece un suo valore solo nei casi in cui il suo uso si basa su una scelta libera del parlante, che può alternarlo all'indicativo per cambiare il registro stilistico della frase. Dunque, l'opposizione indicativo-congiuntivo non andrebbe più ricondotta sempre alla sola opposizione tra codifica della realtà e dell'irrealtà, secondo la posizione tradizionale, bensì bisognerebbe tenere conto anche dei casi in cui la sua scelta è un'opzione del parlante, e di conseguenza, il suo valore si esprime nel contenuto del messaggio (cfr. i già citati Lepschy e Lepschy 1988: 204 e Renzi 2012: 52-53, secondo i quali la scelta tra congiuntivo e indicativo è di tipo stilistico).

10) la concordanza a senso tra il soggetto collettivo singolare seguito da un'espressione partitiva e il predicato plurale (*un milione di elettori non hanno votato*)

Quando il soggetto della frase è costituito da un nome collettivo (gruppo, serie, quantità, infinità, decina, dozzina, parte, classe, folla, gente, maggioranza, ecc.), il predicato spesso ha la forma del plurale, vale a dire è concordato "a senso". Questa concordanza è più frequente quando il collettivo è accompagnato dal partitivo al plurale: *Una quantità di uccelli si alzarono in volo. [...]* Ma anche senza la successiva specificazione partitiva al plurale, il nome collettivo a volte è seguito dal predicato al plurale: si sente dire, e talvolta si scrive, *Un centinaio partono domani; la terza C (sottinteso classe) vanno a Venezia; Venite, gente.* (Sabatini 1985: 167)

Si tratta di un fenomeno con antichissimi precedenti, nel latino, in altre lingue romanze e non, e in italiano, per il quale Sabatini (*ibidem*) fa riferimento ad esempi tratti da Dante, Boccaccio e Manzoni.

11) la posposizione del soggetto rematico al predicato (ossia il soggetto in posizione rematica rispetto al predicato), talvolta con discordanza di numero con questo (*non ci sono giornali e non c'è giornali*, da cui si passa anche a: *niente giornali*)

Come già anticipato per le dislocazioni (punto 4 della lista del 1990), in italiano l'ordine non marcato dei costituenti della frase è soggetto + verbo + oggetto (SVO). Tale ordine è parzialmente modificabile, dunque è possibile mettere in rilievo il soggetto, o altri argomenti del verbo, per attirare su di esso l'attenzione di chi ascolta, presentandolo come l'elemento nuovo o inatteso nel discorso. Questa evidenziazione può essere realizzata in vari modi¹², ad esempio sintatticamente invertendo l'ordine SV in VS: *ho preso io la decisione*. A questo proposito: “Se nella frase *Mario canta* l'informazione nuova è in *canta* (potenzialmente contrapposto a *balla, suona, ecc.*), nella frase *Canta Mario* l'informazione nuova è concentrata in *Mario* (potenzialmente contrapposto a *Fabrizio, Luigi, ecc.*)” (Sabatini 1985: 162).

La possibilità di spostamento del soggetto cambia a seconda delle tipologie di frasi, poiché per alcune di esse la costruzione ordinaria, non marcata, si realizza proprio tramite la posizione post-verbale del soggetto. Per citare alcuni esempi, in presenza di un participio assoluto¹³ il normale ordine non marcato della frase prevede che questo sia seguito dal soggetto (esempio: *Morto Napoleone* e non *Napoleone morto*). Vi sono poi le frasi che contengono verbi che rientrano nella classe dei verbi intransitivi detti *inaccusativi*. Tale classe di verbi può essere coinvolta in frasi in cui la posizione post-verbale del soggetto rappresenta la normale costruzione non marcata; seguono esempi tratti dal corpus PAISÀ: *occorre capacità, caparbietà e competenza; serve più umiltà; esiste un atto del Sindaco che nomina 5 membri del CSA*. In presenza di verbi inaccusativi il soggetto ha le caratteristiche dell'oggetto diretto, poiché si trova dopo il verbo, all'interno del sintagma verbale:

Si tratta di frasi presentative, costituite cioè dalla presentazione di un evento, dunque dal solo predicato. Esse, a differenza delle frasi predicative, non sono articolabili in soggetto e

¹² In alternativa all'inversione dell'ordine dei costituenti possono essere utilizzate strategie lessicali, facendo precedere il costituente da focalizzatori come *anche, proprio, perfino* (esempio: *proprio io ho preso la decisione*), e/o prosodiche. In quest'ultimo caso, il costituente della frase focalizzato viene pronunciato con enfasi, accentuandolo e separandolo dal resto della frase tramite una pausa o un innalzamento (esempio: *IO ho preso la decisione*; oppure: *io [pausa] ho preso la decisione*). Tali strategie sono, chiaramente, anche cumulabili.

¹³ Come indicato da De Roberto (2010) le subordinate participiali assolute sono quelle in cui il soggetto è esplicitato e non coincide con quello della proposizione principale. Esempio: *arrivato Gianni, la festa si animò* (cfr. Egerland 1996 e Loporcaro, Seiler 2009).

predicato. A questa particolarità logico-semantică se ne associa una di tipo sintattico: secondo alcune teorie in queste frasi il soggetto è un argomento interno al sintagma verbale (mentre nelle frasi predicative è argomento esterno). (Palermo 2011)

Un esempio è: [SV[VÈ arrivata] + [SNAnna]], ove SN è il sintagma nominale e SV è il sintagma verbale.

Le frasi passive e i costrutti introdotti da *si* impersonale (o *passivante*) possono avere un soggetto postverbale percepito come non marcato (Sansò 2003: 103). Alcuni esempi tratti da Palermo 2011: *furono presi provvedimenti eccezionali; si commettono molti errori per inesperienza; si verificano molti incidenti per la distrazione del guidatore*. Invece, la costruzione è più limitata nella frase interrogativa diretta, che tende ad avere un ordine fisso dei costituenti (Patota 1990, Palermo 1997, 2011).

Fatta questa necessaria premessa, il fenomeno inserito in elenco da Sabatini riguarda invece i casi in cui lo spostamento del soggetto in posizione post-verbale costituisce una strategia di prospettiva comunicativa e, dunque, viene percepito come marcato rispetto all'ordine SVO.

Anche questo fenomeno ha dei precedenti storici: l'inversione del soggetto come strategia di messa in rilievo si afferma stabilmente in italiano dal XVII secolo; nell'italiano antico e in quello letterario almeno fino alla metà del XIX secolo il soggetto era quasi sempre espresso e posposto al verbo nelle frasi interrogative (Patota 1990). Una traccia di ciò è ancora visibile nella lingua contemporanea in contesti formali e cerimoniali (*lo giurate voi?; vuoi tu prendere per legittima sposa la qui presente ...?*).

Per quanto riguarda, invece, l'uso di *niente* in funzione di aggettivo (esempio: *niente soldi*), questo fenomeno, assimilato al soggetto in posizione rematica nella presente lista del 1990, era invece trattato separatamente, in un punto a se stante della lista del 1985 (punto 31):

L'uso di *niente* in funzione di aggettivo (già attestato in epoca antica) permette di realizzare un tipo di espressione partitiva-negativa particolarmente efficace e di largo impiego: [...] *In questa marmellata niente coloranti, niente conservanti* (così spesso si legge sulle confezioni in commercio) (ivi: 168)

Per le caratteristiche generali della categoria grammaticale del soggetto, si veda anche Palermo 2011.

12) la diatesi media dei verbi (non solo i più comuni *mi ricordo, mi sbaglio, me ne vado*, ma *mi vedo, mi godo, mi bevo, mi fumo, mi mangio* ecc.), ossia, secondo la versione del 1985, la costruzione dei verbi con forma pronominale per indicare una più forte partecipazione affettiva o di interesse (*mi bevo un caffè*)

Quest'uso detto "costruzione riflessiva apparente o di affetto" è frequentissimo (con i verbi mangiare e bere (e loro sinonimi) e con altri che indicano azioni o atteggiamenti implicanti effetti sulla persona del soggetto: *Luca si è mangiato mezza torta; Verso le 11 mi bevo un caffè; Valerio si vede un film in televisione; Aldo si gode la vacanza; Stamattina mi sono fatto una splendida passeggiata.* (ivi: 167-168)

A questo proposito, si osserva come la definizione "verbo in forma pronominale per indicare partecipazione affettiva", coniata da Sabatini nel 1985, possa risultare lievemente equivoca, poiché in realtà il fenomeno si basa sull'uso di un clitico di oggetto indiretto che coincide, di fatto, con il dativo etico.

Questa struttura indica la partecipazione o il coinvolgimento emotivo di una persona rispetto a un'azione o a una circostanza indicata dal predicato; è sempre espresso da un pronome atono e non è necessario ai fini della compiutezza sintattico-grammaticale dell'enunciato (Salvi 1988: 65-66).

Questa denominazione deriva dalla grammatica latina, in cui era possibile l'uso del pronome personale al dativo per indicare la partecipazione o il coinvolgimento emotivo di una persona rispetto a un'azione (esempio: *quid mihi Tulliola agit?*, ossia "che cosa mi combina la piccola Tullia?", in Cicerone).

In italiano l'equivalente del dativo etico si realizza utilizzando i pronomi atoni. In particolare, quando sono utilizzati i pronomi di prima e quarta persona *mi* e *ci*, questa struttura "serve a enfatizzare la partecipazione emotiva del parlante, come effetto del riferimento deittico al soggetto enunciativo" (Cignetti 2010¹⁴). E ancora:

Un valore affettivo-intensivo, atto a segnalare una più attiva e sentita partecipazione del soggetto all'azione, affine a quello del dativo etico propriamente detto, può essere esteso agli impieghi di pronomi atoni pleonastici accompagnati a un verbo transitivo, in frasi come *bersi una birra, farsi una passeggiata, farsi quattro risate* e sim.; oppure in casi in cui sono presenti riferimenti a parti del corpo (*soffiarsi il naso*), ad attività biologiche (*asciugarsi le lacrime*) o a

¹⁴ Versione online, pertanto priva di numeri di pagina.

indumenti che riguardano la sua persona (*togliersi il cappello*) (Serianni 1988: 250). (Cignetti 2010)

13) la frase scissa (*sei tu che lo vuoi*)

La frase “scissa” (o “spezzata”) in due frasi, di cui la prima, col verbo *essere*, mette in forte rilievo il “nuovo”, mentre la seconda contiene il “noto”. [...] Questo tipo di frase non solo porta al massimo grado l’enfasi sul nuovo, ma amplia la durata dell’enunciato e quindi raggiunge l’effetto di quello “spezzettamento” dell’informazione che facilita la ricezione” (Sabatini 1985: 163).

Si tratta di una strategia sintattica che ha lo scopo di mettere in evidenza il punto di maggiore interesse del parlante e che contiene la maggiore quantità di informazione nuova. Tale segmento frasale su cui si focalizza l’attenzione si definisce *focus* o fuoco (si vedano gli studi di Benincà, Salvi, Frison 1988, Berretta 2002, D’Achille, Proietti, Viviani 2005, Roggia 2009).

La costruzione si impone a partire dal Settecento, sotto l’influenza del francese (Roggia 2009), mentre ci sono poche attestazioni per la lingua antica. Nell’uso contemporaneo, come fa notare Renzi (2012: 47), la frase scissa viene utilizzata nella lingua scritta colta anche senza la funzione di mettere l’elemento che regge il verbo *essere* a contrasto con un altro elemento (esempio: *sei tu che lo vuoi, non lui!*), ma semplicemente per metterlo in rilievo (esempio: *sei tu che lo vuoi!*). È utilizzata anche in costruzioni interrogative, in cui l’introduttore viene focalizzato ad inizio frase (esempi: *quand’è che torni? Con che cosa hai scritto? Dov’è che abiti?*, cfr. Renzi 2012: 47-48) e alcuni esempi di queste sono stati reperiti anche nell’italiano antico (Roggia 2009).

È interessante notare che tra gli esempi citati da Sabatini (1985; 1990: 98) c’è solo la tipologia della frase scissa esplicita, costituita dal verbo *essere* + [altro elemento grammaticale] + *che*; manca, invece, la tipologia implicita, basata su *a* + verbo all’infinito (esempio: *è lui a telefonare*).

14) ci attualizzante con *avere* e con alcuni altri verbi (*non c’ho tempo; non ci capisco niente*)

Nella versione del 1985 della lista si faceva riferimento all’uso attualizzante in presenza dei verbi *essere* e *avere* e di altri verbi:

originariamente con valore di avverbio di luogo ‘qui’ (dal lat. *ecce hic*), ha un uso larghissimo in unione con i verbi *essere* e *avere* (non con valore di ausiliari) e con altri verbi. Essa ha in gran parte perduto il suo significato originario: la sua funzione è quella di rinforzo semantico e fonico alle forme verbali. Con il verbo *essere* il *ci* [...] ha un effetto più propriamente e solo “attualizzante”. (ivi: 160)

A questo proposito, Sabatini precisa che l’uso del *ci* è normale e obbligatorio quando si descrive un evento specifico del quale implicitamente sono richiamati aspetti materiali e localizzabili (*Oggi c’è sciopero dei giornali*) e quando il verbo *essere* è usato nel significato di “esistere”, anche senza un implicito riferimento concreto ad un luogo (*C’è della gente che ama perder tempo*).

Per quanto riguarda il suo uso con il verbo *avere*, “è più evidente ancora la funzione puramente attualizzante del *ci* originariamente avverbio di luogo. Nell’uso orale della lingua esistono casi in cui il *ci* è obbligatorio per evitare ambiguità: a una domanda come *hai l’ombrello?*, Si risponde *ce l’ho* e non *l’ho*” (*ibidem*). Ciò avviene per evitare che la risposta si confonda con *lo* (pronome o articolo), analogamente a *ce l’ha/ non ce l’ha*, per evitare ambiguità di suono con *la* o *non la*. La stessa domanda spesso si formula come: *c(i) hai l’ombrello? O ce l’hai l’ombrello?* “Sulla base di casi obbligatori come questi, si spiega l’uso (diffusissimo in tutte le regioni d’Italia e tra parlanti colti) delle espressioni *c(i) ho fame, c(i) ho freddo, non c(i) ho tempo, c(i) hai ragione e simili*” (ivi: 160-161). Sull’uso con il verbo *avere* si veda in particolare lo studio di Moro (1998), che traccia un parallelismo tra l’uso di *averci* ed *esserci*, arrivando a paragonare a livello funzionale la struttura *c’hanno molti libri* con *ci sono molti libri*.

Come precisato da Sabatini, nello scritto queste forme stentano ad entrare, perché percepite come troppo colloquiali e perché ci sono delle difficoltà nella resa con la grafia normale della pronuncia palatale della *c* isolata, conservando l’*h* grafica del verbo *avere*. A questo proposito Renzi (2012: 55-56) ricorda come *averci* è usato nella letteratura, in particolare da Sanguineti, per dare alla scrittura la connotazione del parlato spontaneo.

Nella versione del 1985 viene inoltre presentata anche la funzione attualizzante dei verbi *sentire* e *vedere* in espressioni come *ci vedo, ci sento* ecc., ove la funzione è puramente attualizzante, perché il significato è semplicemente quello di sentire e vedere bene, senza riferimenti al luogo, e altri verbi come *entrare, capire, credere* e *volere*, ove il *ci* conserva in parte il suo significato originario di “qui, in questa faccenda”.

Mentre il primo elenco del 1985 includeva fenomeni sia fonologici che morfosintattici e lessicali, nella seconda versione Sabatini si concentra esclusivamente sui fenomeni morfosintattici diffusi su tutto il territorio nazionale. Pertanto, volendo fare un confronto tra i due elenchi verranno qui riportati i soli fenomeni morfosintattici successivamente esclusi dalla seconda versione:

- l’uso di *codesto* come aggettivo e pronome dimostrativo e degli avverbi di luogo (*costì* e *costà*), che, al di fuori della Toscana, sono limitati al solo uso burocratico sull’intero territorio nazionale. Nel parlato per specificare la posizione degli oggetti rispetto all’interlocutore si ricorre ad altri elementi della frase o a gesti;
- le forme aferetiche *'sto*, *'sta* vs *questo* e *questa*, che connotano la lingua in senso colloquiale, ma di cui bisogna riconoscere la diffusione sull’intero territorio;
- l’uso di *questo*, *quello* e *lo* in funzione di neutro per richiamare un intero enunciato (esempi: *questo non è vero*; *vorrei dirti solo questo: sei ingenuo*; *ti dico quello che so*; *lo so*) o un complemento predicativo del soggetto o dell’oggetto (*si crede furbo, ma non lo è*);
- l’uso dei dimostrativi *questo* e *quello* rafforzate da *qui* e *lì*: *questo qui* e *quello lì*;
- la scelta nella lingua parlata della particella pronominale *ci* (o *ce*) con valor di avverbio di luogo, rispetto a *vi* (*ve*) (esempi: *ci resto*, *ci metto*, *metterci*, *mettercelo*);
- i vari usi della particella *ci* (*ce*) in unione con i verbi *essere* e *avere*, i verbi *sentire* e *vedere* (in espressioni come *ci vedo* e *ci sento* con il significato di vedere e sentire bene), con i verbi *entrare* *capire*, *credere*, *volere* (in espressioni come: *(non) c’entra*, *che c’entra?*, *non ci capisco niente*, *non ci credo* ecc., con il significato di *qui*, *in questa faccenda*, *a proposito di questi fatti*), dove il *ci* ha la funzione di rinforzo semantico e fonico alle forme verbali e conserva solo una sfumatura di avverbio di luogo (nella versione del 1990 viene conservata solo la funzione “attualizzante” (*c’è uno sciopero*, *c’è della gente che ama perdere tempo*, *ce l’ho*, ecc.));
- la tendenza del pronome clitico, in presenza di un verbo servile, a “risalire” nella sintassi, passando dalla posizione enclitica (*qui possono comprarsi*) alla posizione proclitica (*qui si possono comprare*). Oltre al *si* passivante, appena citato, si menzionano anche i casi di risalita in presenza dei verbi modali seguiti dall’infinito (*Luca ti deve parlare* vs *Luca deve parlarti*), con i verbi aspettuali *stare* + gerundio, *stare a*, *stare per*, *cominciare a*, *finire di* e con i verbi *andare* e *venire* (esempi: *lo stiamo aspettando*, *non mi sta a sentire*, *mi sta per finire la benzina*, *mi comincia a*

venire sonno, ti finisco di raccontare la storia, vi vengo ad aiutare, ce ne vogliamo andare);

- la stessa tendenza alla posizione proclitica si osserva con l'imperativo negativo (*non ti muovere vs non muoverti*);
- l'uso di *che* in funzione di aggettivo interrogativo al post di *quale* (*che vestito ti metti?*), assimilabile all'uso nelle esclamative (*che pazzia!*);
- nel parlato è avvenuta una selezione che ha portato alla prevalenza dell'uso di alcune congiunzioni rispetto ad altre: causali (*siccome, dato che vs poiché, giacché*); finali (*perché vs affinché*) e interrogative (*come mai vs perché*) sulle altre congiunzioni; in aggiunta, nella lingua d'uso medio, la relazione causale spesso viene espressa paratatticamente tramite la congiunzione *e* (esempio: *fa freddo e preferisco uscire*) e la finale può essere resa tramite un costrutto implicito che incorpora un verbo causativo (esempio: *te lo dico per fartici andare*). Nel caso delle interrogative, Sabatini (1985: 166) mette in evidenza che *come mai* denoti una maggiore disponibilità all'ascolto delle ragioni dell'interlocutore, rispetto a *perché*, che è più comunemente utilizzato con tono irato o che denota polemica o contrasto;
- la concordanza tra il participio passato e l'oggetto sotto forma di pronome relativo antecedente, raramente rispettata: con l'ausiliare *avere* il participio è accordato al maschile singolare (esempi: *le vacanze che ho passato; i libri che ho letto*); con il verbo *essere* il participio concorda col soggetto in genere e numero;
- la realizzazione delle forme impersonali con la terza persona plurale (esempio: *bussano alla porta*) o il passivo senza l'agente (esempio: *è aumentato il prezzo della benzina*);
- la giustapposizione di due sostantivi, come *treno lampo, marito modello* ecc., in cui il secondo sostantivo determina il primo. Nella tipologia come *treno merci, fine mese* ecc. è stata soppressa una preposizione;
- l'affermazione tra i pronomi allocutivi di cortesia della forma *Lei* per il singolare e *Voi* per il plurale rispetto ad *Ella* e *Loro*;
- ripetizione dello stesso sostantivo per rideterminare ed esaltare il significato di una parola (*Speriamo che questa sia una vacanza vacanza*, ossia pienamente goduta);
- una lista di vari elementi lessicali e locuzioni che svolgono varie funzioni a livello testuale nell'articolazione del discorso (esempi: *ci vuole/ci vogliono, mi sa, sennò, per forza* ecc.).

La lista del 1990, dunque, nasce dalla rielaborazione della versione precedente: i tratti linguistici appena menzionati sono stati del tutto eliminati; altri, inizialmente raggruppati sotto un'unica categoria sono, poi, stati trasformati in punti separati dell'elenco del 1990 (è il caso della posposizione del soggetto al predicato, la dislocazione, l'anacoluto e la frase scissa, inizialmente raggruppate nell'elenco del 1985 al punto 17, come fenomeni di enfasi e di tematizzazione), al contrario l'uso di *niente* in funzione di aggettivo, a cui era dedicato il punto 31 della lista del 1985, è stato poi assimilato nella lista del 1990 al soggetto in posizione rematica.

Dopo aver descritto i tratti della lista del 1990 e aver fatto un breve confronto con la versione precedente, nella sezione successiva verrà fatta una riflessione sulle ragioni che hanno determinato il passaggio dalla prima alla seconda versione. In particolare l'esposizione si concentrerà sulla ricerca di un modello di riferimento linguistico e sulle implicazioni che ciò ha sull'insegnamento dell'italiano.

2.3. La lingua ritrovata

Sebbene non vengano mai esplicitamente espresse le ragioni che hanno portato alla riduzione dei fenomeni da 35 a 14, osservando i fenomeni esclusi dalla versione del 1990, è ipotizzabile che in una prima fase sia stato fatto un elenco più ampio, che raggruppasse tutte le forme presenti nell'uso comune e che storicamente erano percepite come una deviazione dalla norma codificata dalle grammatiche. Tuttavia, già negli anni '80 del Novecento l'uso di alcune di queste forme poteva essere considerato ormai come un dato di fatto e non costituiva più un elemento di novità o di estraneità, dato che stavano entrando anche nello scritto di una certa formalità a livello nazionale, sebbene non fossero state codificate esplicitamente nella grammatica. A questo proposito Ondelli e Romanini (2018: 188) osservano che

alcuni tratti potevano essere percepiti come desueti già 30 anni fa (*in Svizzera e in Spagna*), fino a rasentare l'agrammaticalità (*le camicie che ho stirate*), mentre altri sfuggono alla percezione (la minor frequenza della *d* eufonica, ma anche le concordanze *ad sensum*, cfr. Ondelli & Viale 2010) o sono portatori di marche diamesiche o diafasiche (*gli per loro*).

Di conseguenza, la lista del 1990 può essere considerata come una seconda riflessione su ciò che può costituire davvero il cuore del nuovo standard linguistico, ossia con l'obiettivo

di focalizzare l'attenzione sui soli fenomeni identificativi della nuova varietà panitaliana. L'identificazione dei tratti grammaticali che costituiscono questa varietà ha delle conseguenze anche sull'insegnamento dell'italiano:

L'insegnamento dell'italiano, sia in Italia che all'estero, non può essere tenuto "al riparo" dai problemi fin qui considerati. Almeno la distinzione - ormai fondamentale per la situazione italiana- tra VARIETÀ STANDARD PER L'USO SCRITTO FORMALE e VARIETÀ DELL'USO MEDIO PARLATO E SCRITTO dovrebbe essere presa in seria considerazione, specialmente se nell'insegnamento della lingua si perseguono obiettivi differenziati e graduati: ciò si impone particolarmente per l'insegnamento dell'italiano come lingua seconda, a discenti che molto spesso puntano ad acquisire una competenza (dapprima passiva, poi attiva) innanzitutto sul piano della lingua dell'uso medio, parlato e scritto. (Sabatini 1985: 180)

La questione viene ripresa nello studio del 1990, ove si riflette sulla diffusione dell'apprendimento dell'italiano come lingua straniera all'estero, e in cui si afferma che la varietà di lingua che viene appresa non è

ricalcata sulle [...] norme enunciate nelle nostre grammatiche precettive, le quali propongono, come ormai sappiamo, un sistema linguistico incompleto e più che altro virtuale, in quanto quasi totalmente epurato di riferimenti alla componente pragmatica. [...] Si diffonde quell'italiano parlato, divergente sì, dalla norma per vari tratti, ma unificato sul piano nazionale, di cui sanno servirsi con scioltezza e sicurezza le nostre classi colte, e che rappresenta altresì il traguardo al quale tendono oggi anche le nostre classi meno istruite. (Sabatini 1990: 90)

Si tratta, dunque, di una lingua che corrisponde all'italiano parlato che si discosta dal sistema linguistico definito incompleto e virtuale proposto dalle grammatiche. All'origine di tale divario vi è l'unificazione linguistica italiana, che, come ricorda Poggi Salani (2012: 119), era avvenuta quasi esclusivamente nello scritto, lasciando scoperto il lessico della vita quotidiana tipico del parlato e, allo stesso tempo, causando la sopravvivenza di molteplici varianti lessico grammaticali:

Come insegnare dunque l'italiano nelle scuole? [...] Gli insegnanti, trovandosi per mestiere al centro dei problemi, si appigliarono, con inclinazioni diverse, alle due sole ancore che potevano avere a disposizione: LA VECCHIA NORMA LIBRESCA DELL'ITALIANO DELLA TRADIZIONE E LA NUOVA NORMA TOSCANA (più che strettamente fiorentina), e le due

norme trovarono per lo più punti di contatto e di compromesso. Sono questi i poli cui si ispirano anche libri di testo, grammatiche e vocabolari d'uso corrente. (*ibidem*)

Ciò portò alla nascita del cosiddetto “scolastichese” (Cortelazzo 2000b, 2012), ossia di una lingua ricalcata sulla norma toscana, ma che di fatto si rivelerà nel corso del tempo distante dalla varietà usata dalla comunità di parlanti.

Sabatini (1990: 90) a questo proposito coglie un'analogia: la varietà d'italiano a cui tendono gli apprendenti stranieri, più vicina al parlato che alla norma, può essere considerata lo specchio del cambiamento linguistico che la lingua vive – sebbene in maniera più inconsapevole – in Italia “e che consiste nell'affermarsi sempre più largo, nei più diversi contesti e nelle sfere sia del parlato che dello scritto, di un modello di lingua italiana fortemente comunicativa al tempo stesso unitaria” (*ibidem*).

Come precisato da Sabatini stesso, la scelta del verbo *affermarsi*, e non *nascere*, è dovuta al fatto che l'italiano parlato di carattere nazionale a cui fa riferimento non è una creazione recente, poiché dalle ricerche (si citano a titolo di esempio Nencioni 1987 e D'Achille 1990) è emersa la storia secolare di alcuni fenomeni morfosintattici, che erroneamente sono percepiti come mutamenti linguistici recenti, e l'evoluzione graduale avvenuta nella lingua italiana tra la fine dell'Ottocento e la fine del Novecento.

Ciò trova conferma nel titolo del suo studio del 1990, *Una lingua ritrovata: l'italiano parlato*, ove l'aggettivo *ritrovata* fa riferimento ai tratti grammaticali che, in realtà, esistono da secoli e coesistono, seppur in secondo piano, con la norma grammaticale imposta dalla tradizione che ha le sue radici in Bembo (cfr. capitolo 1). Nel corso del tempo tali tratti sono stati “smarriti”, poiché sottoposti a censura diretta o alla semplice omissione da parte delle grammatiche specialistiche e, soprattutto, scolastiche: “A volte, più che la censura, nuoce l'omissione: anche questa rivela la mancanza di spessore nella considerazione della lingua” (Sabatini 1985:180). Analogamente Cortelazzo (1984) ha affermato che la tradizione scolastica “prescrittivo-repressiva” rischiasse di condannare la lingua italiana all'ostracismo, rallentandone l'evoluzione spontanea.

Dal divario tra l'italiano codificato nelle grammatiche e l'italiano parlato nasce la necessità per l'insegnante di italiano di fare riferimento ad una norma e, allo stesso tempo, la difficoltà di come stabilirla, dati i continui cambiamenti nella lingua. Sul tema interviene anche Sobrero (1992b) che, richiamando il dibattito avvenuto tra la fine degli anni Ottanta e gli anni Novanta tra Sabatini e Castellani (cfr. sez. 2.1), fa suo il punto di vista di quest'ultimo. Castellani, prendendo in esame la lista del 1985 dei tratti dell'italiano

dell'uso medio, era giunto alla seguente conclusione: ad eccezione dei casi in cui questi tratti coincidano con quei tratti dell'italiano normale che nel corso del tempo sono stati omessi o non approfonditamente descritti dalla maggior parte delle grammatiche scolastiche, essi si riducono a delle semplici tendenze o a delle doppie possibilità all'interno della norma. E su questo Sobrero conclude che si tratterebbe, tutto sommato, di semplici variazioni di registro.

A mio avviso (dopo le ovvie considerazioni storiche o funzionali, che servono ad inquadrare il problema) si dovrebbero applicare, ad ognuno dei registri, criteri legati alla sincronia, all'attualità. Mi sembrano in particolare determinanti due parametri: della rilevanza statistica e della stigmatizzazione sociale. In base a questi criteri, per lo scritto formale si dovrebbero comunque accettare: (a) le forme che hanno una presenza e una frequenza statisticamente molto rilevanti nella produzione linguistica scritta e orale; (b) le forme che, già colpite da censura per la loro appartenenza al livello colloquiale o dialettale, ora sono generalmente accettate. E, naturalmente, si dovrebbero respingere le forme che non hanno questi requisiti. In questo modo, ad esempio, si accetterebbe senza tentennamenti il *lui* soggetto (criterio a) e la frase segmentata (tipo *i soldi te li ho dati*) (criterio b); ma si rifiuterebbe il *codesto* (a) e la costruzione del tipo *la valigia che ci ho messo i libri* (b). In modo un po' diverso ci si potrebbe regolare nell'orale di registro formale, e in modo ancor più permissivo nei registri colloquiali. (Sobrero 1992b: 153)

Dunque, in entrambi i casi, seppur declinata attraverso i due aspetti della rilevanza statistica e della accettazione sociale, la questione si basa sul determinare la preesistente affermazione di tali fenomeni, sia essa di carattere quantitativo o sociale. Tale punto di vista, tuttavia, conserva una certa vaghezza sulla questione della funzione linguistica (ossia, quali strutture diano un effettivo apporto alla comunicazione), che viene data come premessa all'inizio del brano sopra citato, ma non affrontata direttamente.

Il tema della funzione linguistica come criterio di valutazione dei tratti a distanza di trent'anni è messo, invece, in primo piano da Prandi (2020: 174-175), che nella sua grammatica traccia una riflessione sulla censura generalizzata delle grammatiche rispetto ai tratti dell'italiano dell'uso medio/neostandard, prendendo spunto dalla trattazione delle strutture marcate di frase. A questo proposito, viene messo in evidenza come, ad esempio, le strutture marcate di frase siano generalmente assenti nelle grammatiche, perché considerate tradizionalmente caratteristiche del parlato informale, o addirittura dell'improvvisazione, e causa di un abbassamento del registro.

Che si tratti di omissione nella trattazione o di sanzione dell'uso, Prandi afferma che dinanzi a questi atteggiamenti è necessario affermare con decisione che le frasi marcate non costituiscono una caduta di registro, poiché rappresentano degli strumenti indispensabili della prospettiva comunicativa, di cui la lingua non può fare a meno, pena l'impossibilità di costruire testi coerenti e, nel caso specifico, di esprimere le gerarchie di peso comunicativo

Ciò detto, se da un lato la critica all'omissione e alla censura nelle grammatiche fatta da Sabatini e la riflessione di Prandi pongono l'attenzione sulla tendenza errata a rifiutare in blocco tutto ciò che differisce dalla norma prescrittiva, dall'altro ciò non deve indurre, per contro, a trarre conclusioni affrettate circa l'accettabilità di tutti i tratti dell'italiano dell'uso medio/neostandard.

Questo punto di vista è condiviso anche da Renzi (2012: 39), che mette in guardia dal pericolo di ricadere sia nell'errore dei grammatici classici, che per secoli hanno inserito tra i loro scopi la tutela della buona lingua, la tutela dagli errori e, dunque, dal cambiamento, sia nell'errore opposto, ossia di credere che tutti gli errori di oggi saranno le regole di domani. Tuttavia, mentre la conclusione di Renzi sulla questione è solo di tipo evolucionistico, ossia alcune innovazioni si imporranno, mentre altre perderanno terreno e verranno eliminate, l'approccio di Prandi, invece, poggia saldamente sul criterio della funzionalità o mancata funzionalità delle forme linguistiche, da cui è possibile trarre dei criteri per l'insegnabilità dei tratti.

Come messo in evidenza da Prandi (2020: 174-175), nell'elenco dei tratti distintivi dell'italiano dell'uso medio (Sabatini 1985; 1990) e del neostandard di Berruto (1987) sono presenti sia strutture grammaticali funzionalmente indispensabili, come la dislocazione a sinistra e a destra e la frase scissa, sia forme substandard che dal punto di vista funzionale non sarebbero necessarie, poiché sono delle mere alternative a forme standard perfettamente funzionanti. Per spiegare meglio questo punto, Prandi cita i seguenti esempi di strutture che abbassano effettivamente il registro senza apportare di fatto alcun vantaggio sul piano funzionale: 1) l'uso del *che* polivalente in sostituzione di strutture perfettamente funzionanti, come le congiunzioni subordinanti (esempio: *questa corsa la lascio fare a voi, perché io sto bene qui* vs *questa corsa la lascio fare a voi, che io sto bene qui*) e i pronomi relativi (secondo la struttura (*preposizione*) + *articolo* + *quale/cui*, esempio: *la persona con cui/ con la quale ho parlato ieri* vs *la persona che ho parlato ieri*). Per tutte le realizzazioni possibili del *che* polivalente in sostituzione del pronome relativo o di congiunzioni subordinanti si rinvia rispettivamente a Berretta (1993: 230-231)

e Berruto (2012 [1987]: 78); 2) l'uso combinato di *ne* e del pronome relativo (esempio: *il libro del quale ne abbiamo appena discusso*); 3) l'uso ridondante di *ci* (esempio: *Non c'ho tempo*); 4) l'uso dell'indicativo al posto del congiuntivo nelle frasi argomentali: *Mi spiace che hai perso il treno* (cfr. la riflessione sull'alternanza dell'uso dell'indicativo e del congiuntivo di Prandi 2002 al punto 9 dell'elenco dei tratti del 1990); 5) l'uso del pronome atono *gli* al posto di *le* per il femminile (esempio: *Gli hanno regalato una bambola (a Maria)*). Quest'ultimo fenomeno, ad esempio, rappresenta un importante punto di riflessione sulla pertinenza funzionale: come già affermato in precedenza nella sezione 2.2, non vi è alcun vantaggio sintattico nella sostituzione di *le* con un'altra forma pronominale avente le stesse caratteristiche, dunque l'uso di *gli* al posto di *le* è una deviazione del tutto arbitraria rispetto alla norma. Al contrario l'uso di *gli* per la terza persona plurale atona riempie un effettivo vuoto formale e funzionale nel paradigma (per i dettagli si rinvia alla descrizione dei due fenomeni al punto 2 dell'elenco della sezione 2.2).

Riesaminando, quindi, l'elenco dei tratti del 1990, Prandi (2020: 174-175) fa notare come al suo interno si trovino sia fenomeni che si distaccano dalle norme della tradizione prescrittiva, ma che sarebbero degne di fare parte dello standard linguistico perché svolgono funzioni linguistiche indispensabili, sia fenomeni che infrangono la norma senza apportare di fatto alcun vantaggio funzionale.

Questa riflessione non nega che l'elaborazione di concetti come *italiano neostandard* o *italiano dell'uso medio* abbia dato risposta ad un'esigenza esistente da secoli e più volte affiorata nel corso della storia della lingua italiana: la revisione dello standard ereditato dai secoli passati, quando il pieno accesso alla lingua italiana era limitato a poche persone colte, in ambiti d'uso molto ristretti.

Tuttavia, come già anticipato, l'inclusione nelle liste di tutto ciò che si distaccava dalle norme grammaticali ha portato, di fatto, all'accostamento di risorse funzionali essenziali a forme che invece non apportano alcun vantaggio funzionale e che producono una mera caduta di registro. Se da un lato, dunque, gli studi di Sabatini e di Berruto hanno avuto il pregio di documentare, anche da un punto di vista storico, le derive linguistiche, dall'altro permangono delle ombre dal punto di vista della messa a punto di uno standard dell'uso comune che dia sicurezza linguistica a chi lo usa e un metro sicuro a chi ne valuta gli usi. Infatti, al criterio funzionale sopra menzionato per valutare i fenomeni grammaticali si aggiunge la funzione sociale alla quale una norma (standard o neostandard che sia) non può sottrarsi: dare un punto di riferimento sicuro sia a chi usa la lingua, sia a chi ne valuta gli usi:

Se facciamo confluire nel neostandard non solo risorse funzionali essenziali tradizionalmente bollate come errori sulla base di un pregiudizio, ma anche deviazioni dalla norma prive di giustificazione funzionale, indici sicuri di una caduta di livello soggetti a censura in contesti anche solo mediamente formali, la nozione perde ogni coerenza e valore funzionale. Per intenderci, il candidato che a un colloquio di lavoro usa la dislocazione e la frase scissa si colloca a tutti gli effetti nella norma standard e non corre alcun rischio. Il candidato che usa il cumulo di *ne* e pronomi relativo o la forma *gli* per il femminile singolare, al contrario, si espone certamente a forme di censura, ed è normale che sia così. (Prandi 2020: 175)

Riflessioni analoghe sono presenti anche nello studio di Ondelli e Romanini (2018), in cui il problema del modello di riferimento linguistico, che persiste ancora a trent'anni dallo studio di Sabatini, è inquadrato dal punto di vista dell'insegnamento: se da un lato, secondo i due studiosi, il canone linguistico di riferimento per la "buona scrittura" non è più da ricercarsi nella letteratura, ma nei testi giornalistici di qualità (cfr. Dardano 1994: 34, Antonelli 2011: 51) che sono i depositari della varietà di italiano dell'uso medio, dall'altro la scuola non sembra recepire tale cambiamento, contribuendo al persistere del cosiddetto "scolastichese" (Cortelazzo 2000b, 2012), basato sulle prescrizioni contenute nelle grammatiche.

Allo stesso tempo, come osservato da Ondelli e Romanini (2018), l'esigenza di tenere conto sia della ricchezza e variabilità degli usi della lingua contemporanea, sia di regole di riferimento che possano essere considerate stabili, ha delle ripercussioni anche nella linguistica. Ne sono un esempio le difficoltà che si pongono nella definizione di parlante nativo dell'italiano (cfr. Acquaviva 2000, Berruto 2012) e nella scelta di una varietà linguistica di riferimento standard in qualsiasi tipo di analisi linguistica, soprattutto se di tipo quantitativo (cfr. Ondelli, Pontrandolfo 2014):

In effetti, se l'[italiano dell'uso medio] può essere oggi considerato lo standard nel senso della frequenza statistica, il giudizio si fa più difficile quando si passa a considerare i concetti di correttezza grammaticale e neutralità di registro. Il problema sembra almeno in parte risiedere anche nella diversa natura dei 35 tratti a suo tempo individuati da Sabatini (anche se poi ridotti a 14 nel 1990) e nella conseguente diversa accettabilità da parte dei parlanti. (Ondelli, Romanini 2018: 187)

A questo proposito, come già anticipato, sempre Ondelli e Romanini (2018) nel loro studio hanno indagato il grado di accettabilità di alcuni tratti dell'italiano dell'uso medio presso un campione di 150 persone provenienti da tutto il territorio nazionale, prive di competenze specifiche nelle questioni linguistiche, a cui sono stati sottoposti brevi estratti di articoli pubblicati tra il 2015 e il 2016 da *Corriere della Sera* e *Repubblica*, contenenti alcuni dei tratti elencati da Sabatini (1985) e Berruto (1987)¹⁵. Oltre alla valutazione del grado di accettabilità su una scala di cinque descrittori che spaziano dalla totale inaccettabilità alla piena accettabilità, con tre gradi intermedi (accettabile nel parlato formale, nel parlato informale e nel parlato o nello scritto di media formalità) è stato chiesto loro un commento libero sulla loro qualità linguistica.

Sulla base dei risultati è possibile suddividere i tratti dell'italiano dell'uso medio presi in considerazione considerati nel sondaggio in tre macrocategorie: quelli che non sono oggetto di alcuna censura da parte dei partecipanti o sono oggetto di segnalazioni minoritarie; quelli che invece vengono segnalati con costanza come non accettabili e, infine, quelli per i quali le valutazioni sull'accettabilità sono molto variabili, sulla base del co-testo in cui sono inseriti o delle caratteristiche diastratiche dei partecipanti. In generale, si può notare come le risposte dei partecipanti rispecchino quanto precedentemente affermato da Sobrero (1992b) e Prandi (2020).

Non sono oggetto di alcuna censura da parte dei partecipanti o sono oggetto solo di segnalazioni minoritarie i fenomeni di segmentazione (frase scissa e pseudoscissa, *c'è* presentativo) e la dislocazione a sinistra (meno quella a destra), le concordanze a senso e la sostituzione del pronome dimostrativo neutro *ciò* con *questo* o *quello*: in base alle risposte dei partecipanti sembra che questi fenomeni non vengano più percepiti come forme caratteristiche del parlato o di registri informali.

Vengono, invece, percepiti come errati l'uso del pronome dativale maschile *gli* in sostituzione del femminile *le*, che è generalmente rifiutato dai partecipanti, mentre il *che* polivalente è percepito come errato non solo nei suoi usi prettamente sub-standard (ossia, come pronome relativo generico), ma anche nel suo valore temporale, benché considerato

¹⁵ Una ricerca simile, avente tra i vari obiettivi quello di testare la percezione di accettabilità rispetto ad alcuni tratti tipicamente associati al neostandard tramite delle liste di frasi, è stata svolta da Grandi (2018, 2019). A differenza dello studio di Ondelli e Romanini (2018), condotto solo su persone prive di competenze specifiche nelle questioni linguistiche, l'indagine di Grandi ha coinvolto invece un gruppo di circa 200 studenti dell'Università di Bologna. Tra i tratti che in assoluto mostrano i valori di accettabilità più alti vi sono il futuro epistemico, l'uso di *loro* e *lui* come soggetto, il passato prossimo in luogo del passato remoto e il *ne* ridondante. Il gruppo di informanti che accetta questi tratti tende anche a rifiutare in modo omogeneo *a me mi* e le varie manifestazioni del *che* polivalente, ad eccezione parziale di quello con funzione temporale.

ormai accettabile nello standard (D'Achille 1990: 212). Rispetto a quest'ultimo punto è necessario, però, tenere in considerazione il fatto che la maggioranza dei soggetti che lo sostituiscono con *in cui* rientra nella fascia di età compresa tra i 50 e i 65 anni. Invece, l'uso in funzione consecutivo-esplicativa è segnalato solo da un numero limitato di partecipanti al sondaggio.

Tra i fenomeni per i quali le valutazioni sull'accettabilità sono molto variabili c'è l'uso di *gli* al posto del plurale *loro*, che risulta più accettabile soprattutto per i giovani tra i 18 e i 25 anni, ma il loro numero è pari alla metà degli adulti che lo segnalano come inaccettabile. Un altro tratto che rientra in questa categoria è l'alternanza indicativo/congiuntivo, sul quale i giudizi sono molto variabili e che viene percepito come una variazione di registro e nella cui valutazione si richiede di tenere conto del contesto sintattico (in interrogative indirette, dopo verbi che esprimono incertezza, con verbi di opinione e nelle relative restrittive). Infine, anche per il *ci* attualizzante nella valutazione sull'accettabilità rientra il fattore del contesto: il *ci* è sempre rifiutato in combinazione con il verbo *avere* (esempio: *Io c'ho un problema grosso*), viene raramente segnalato nel caso della ridondanza pronominale (esempio: *Lui [...] alla necessità di una squadra sognante ci crede ancora*), mentre negli altri casi le risposte dipendono molto dal verbo a cui si accompagna, il che probabilmente è indice di un processo più o meno avanzato di integrazione nella norma.

Seppur con i dovuti limiti, segnalati dagli stessi autori, circa il numero ristretto di persone intervistate, i risultati del sondaggio evidenziano comunque che la norma linguistica interiorizzata dai partecipanti tramite la formazione scolastica ha una forte influenza sui loro giudizi: gran parte delle opposizioni dei partecipanti sono assimilabili ad alcune regole caratteristiche dell'italiano scolastico, come la sanzione di *e* e *ma* a inizio frase (Sabatini 1997), che hanno contribuito a formare il loro modello linguistico di riferimento.

Ciò trova conferma anche nelle risposte libere date dai partecipanti, che hanno censurato anche altri fenomeni non presenti tra gli esempi forniti, rispetto ai quali gran parte delle obiezioni si basano su "regole" che tradizionalmente contraddistinguono l'italiano scolastico, di cui seguono alcuni esempi: la critica verso l'uso espressivo della punteggiatura (cfr. Gualdo 2007: 81, Cardinale 2011: 145) molte frasi che iniziano con *e* e *ma*, da cui deriva il rifiuto dell'uso di queste congiunzioni a inizio frase (Sabatini 1997), l'intolleranza verso le ripetizioni lessicali, con la conseguente ricerca di alternative lessicali di registro elevato (Cortelazzo 2000b); a questi si aggiunge la critica nei confronti delle

frasi nominali e monoproporzionali (in contrasto con gli ampi periodi sintattici incoraggiati nei temi scolastici). Alcune di queste censure si rispecchiano anche nella frequenza con cui certi fenomeni si manifestano nella prosa giornalistica. Come affermano Ondelli e Romanini (2018: 190):

Già la maggiore o minore frequenza, o la completa assenza, di certi costrutti dal materiale estratto dagli archivi elettronici delle due testate è indicativa della loro maggiore o minore accettabilità nell'italiano dei giornali: se non c'è stata alcuna difficoltà a reperire articoli caratterizzati dall'oscillazione tra indicativo e congiuntivo o dalle frasi scisse, l'inserimento a posteriori¹⁶ è risultato imprescindibile per *gli* al posto di *le*, per il *che* polivalente introduttore di una frase relativa nei casi obliqui e per la concordanza *ad sensum*. Inoltre, anche l'accordo del participio passato con l'oggetto in posizione antecedente sotto forma di pronomi relativi (che non è considerato tratto caratterizzante dell'[italiano dell'uso medio], ma che abbiamo inserito come controprova) è risultato del tutto assente dagli articoli di giornale.

Pertanto, come affermato dai due studiosi, se da un lato l'italiano dell'uso medio può essere ormai considerato lo standard dal punto di vista della frequenza statistica, il giudizio diventa più difficile quando si prendono in considerazione i concetti di correttezza grammaticale e neutralità di registro. Di conseguenza, un possibile approccio nell'insegnamento della lingua potrebbe essere quello di affrontare i fenomeni grammaticali "ritrovati" – per usare la scelta lessicale di Sabatini – presentandoli come diffusi tra i parlanti nativi e come segnali di un cambiamento tendenziale dell'italiano (in accordo con quanto proposto da Sobrero 1992b), avendo cura, però, di distinguere i fenomeni che deviano dalla norma grammaticale tradizionale dando un effettivo apporto funzionale alla lingua e quelli che, ferma restando la loro diffusione, di fatto non apportano alcun vantaggio funzionale e rappresentano una deviazione del tutto arbitraria rispetto alla norma, il cui uso produce solo un calo di registro (in accordo con quanto proposto da Prandi 2020).

Infine, al termine di questa esposizione, è necessario anche definire la questione della varietà linguistica dal punto di vista terminologico. Avendo adottato la lista di Sabatini del 1990 come base della ricerca, per coerenza nel corso dell'esposizione si è fatto costantemente riferimento alla sua terminologia, ossia *italiano dell'uso medio*, citandola occasionalmente in parallelo con la definizione *neostandard* di Berruto. Tale scelta si basa

¹⁶ Ossia, nel caso di alcuni fenomeni grammaticali di cui sono state reperite poche occorrenze o nessuna, i due studiosi hanno inserito *ad hoc* degli esempi all'interno dei testi.

sul fatto che Sabatini e Berruto attribuiscono, di fatto, due termini diversi allo *stesso* fenomeno, ossia la nuova varietà linguistica standard e lo fanno semplicemente partendo da due prospettive diverse, rispettivamente la prospettiva dello storico della lingua e del sociolinguista, che non entrano in contrasto, bensì si integrano a vicenda. Pertanto, nella consapevolezza di ciò, d'ora in avanti in questo studio si adotterà il termine *neostandard* di Berruto riferimento alla nuova varietà dello standard linguistico, poiché si ritiene che trasmetta in maniera più immediata ed esplicita l'evoluzione avvenuta nella lingua.

Concludendo, dopo aver fornito una descrizione dei tratti linguistici della lista di Sabatini del 1990 e aver discusso la questione della necessità di un modello linguistico di riferimento e delle implicazioni che ciò ha per la didattica dell'italiano, nel prossimo capitolo verranno illustrate le ragioni per le quali i corpora possono rappresentare delle valide risorse per l'apprendimento della lingua.

CAPITOLO III

I CORPORA NELLA DIDATTICA DELLE LINGUE

Dopo aver descritto nei capitoli precedenti le varietà dello spazio linguistico italiano, dando particolare rilievo all'italiano dell'uso medio, e dopo aver fornito una descrizione dei tratti linguistici che lo caratterizzano, in questo capitolo verranno illustrati alcuni concetti base della linguistica dei corpora e le ragioni per le quali i corpora possono rappresentare delle valide risorse per l'apprendimento della lingua. A questo proposito verranno presentati alcuni ambiti di applicazione dei corpora alla didattica delle lingue, dedicando particolare attenzione all'uso dei corpora per la realizzazione di materiali didattici.

3.1. Alcune nozioni di linguistica dei corpora

Il punto di partenza sarà chiarire cosa siano i corpora, per i quali Baroni (2010)¹ elabora la seguente definizione:

I cosiddetti *corpora* (sing. *corpus*) linguistici sono collezioni, per lo più di grandi dimensioni, di testi orali o scritti prodotti in contesti comunicativi reali (per es., registrazioni di discorsi o articoli di giornale), conservati in formato elettronico e spesso corredati di strumenti di consultazione informatici.

Tale definizione si richiama a quella precedentemente elaborata da McEnery e Wilson (1996):

In principle, any collection of more than one text can be called a corpus... But the term "corpus" when used in the context of modern linguistics tends most frequently to have more specific connotations than this simple definition provides for. These may be considered under four main headings: sampling and representativeness, finite size, machine-readable form, a standard reference. (McEnery, Wilson 1996: 21 in Kilgarriff, Grefenstette 2003: 333)

Dunque, se in senso ampio qualunque raccolta di testi può essere considerata un corpus, nel contesto della linguistica moderna una raccolta di testi è definita un corpus se riveste il

¹ www.treccani.it/enciclopedia/corpora-di-italiano_%28Enciclopedia-dell%27Italiano%29

ruolo di oggetto di studio linguistico o letterario (“A corpus is a collection of texts when considered as an object of language or literary study”, Kilgarriff, Grefenstette, 2003: 334) e se è stata realizzata sulla base di alcuni criteri, quali: campionamento e rappresentatività, poiché i testi devono rappresentare un campione dell’uso di una determinata lingua o di una sua varietà; dimensioni definite, ossia la quantità indicativa di testi da includere nel corpus, che viene stabilita nelle fasi iniziali che precedono la realizzazione del corpus stesso. Un’eccezione in tal senso è rappresentata dai corpora di monitoraggio (cfr. sezione 3.3), i quali vengono periodicamente aggiornati con nuovi testi; formato elettronico, poiché tutti i corpora moderni sono in formato digitale e, infine, il concetto del riferimento standard: un corpus deve rappresentare non solo un punto di riferimento per lo studio della lingua o varietà che rappresenta, ma anche un modello rispetto al quale è possibile fare un confronto sia tra corpora diversi, sia con gli studi successivi su un determinato ambito: una volta definita la metodologia impiegata per la raccolta e l’analisi dei dati di un certo corpus, gli eventuali nuovi risultati ottenuti su argomenti correlati possono essere confrontati direttamente con i risultati già pubblicati negli studi precedenti, senza la necessità di un ricalcolo.

Tornando al legame esistente tra corpora, campionamento e rappresentatività, sempre Baroni (2010) chiarisce:

Da un punto di vista statistico, un *corpus* è un campione estratto dalla popolazione di tutti i testi prodotti in una certa lingua, in un certo periodo, in un certo registro, ecc., sulla base del quale possiamo trarre conclusioni che si applicano alla popolazione campionata nel suo insieme. Per es., analizzando un *corpus* di articoli di giornale pubblicati in Italia negli anni Novanta del Novecento possiamo formulare generalizzazioni che, se il *corpus* è sufficientemente grande e vario, in genere sono valide per i testi giornalistici italiani del periodo scelto.

Dunque, un corpus mira idealmente ad essere un campione rappresentativo di una certa lingua, intesa come produzione parlata e scritta. La conseguenza è che le proprietà linguistiche dei testi che contiene (ad esempio, la diffusione delle unità lessicali o delle strutture sintattiche, ecc.) possono essere generalizzate all’intera popolazione linguistica, ossia ad una certa lingua, o varietà linguistica, nel suo insieme. Nello specifico:

In corpus linguistics, the population is the entirety of a given language, or of a given variety of language. We attempt to make generalisations about a population, say of all English newspaper language or of all spoken British English, based on a corpus that contains only a sample of that population. (McEnery, Hardie 2012: 248)

Un campione è, invece, solo una sezione della popolazione:

The word sample may also be used in its statistical sense by corpus linguists. In this latter sense, it means a group of cases taken from a population that will, hopefully, represent that population such that findings from the sample can be generalised to the population. (ivi: 250)

Pertanto, quanto più un corpus sarà esteso e composto da fonti variegata, tanto più sarà rappresentativo di una certa lingua e, di conseguenza, la sua consultazione permetterà di osservare l'uso effettivo di quella lingua, verificandone tendenze generali su base statistica (Lüdeling, Kytö 2008-2009). La selezione di campioni di testi rappresentativi di una lingua dipende dal tipo di lingua che il corpus deve rappresentare:

A corpus seeks to represent a language or some part of a language. The appropriate design for a corpus therefore depends upon what is meant to represent. Representativeness of the corpus, in turn, determines the kind of research questions that can be addressed and the generalizability of the results of the research. (Bieber *et al.* 1998: 246)

La costruzione di un corpus dipende, dunque, da ciò che esso si prefigge di rappresentare (ad esempio, la varietà scritta/ parlata, un registro particolare ecc.) e la sua rappresentatività determina a sua volta il tipo di domande di ricerca per le quali il corpus può essere utilizzato (in un corpus rappresentativo della lingua parlata potrà essere ricercato, ad esempio, l'uso degli intercalari nel parlato). A questo proposito, per quanto riguarda gli scopi con cui vengono utilizzati i corpora, Kilgarriff e Grefenstette (2003: 343) citano i seguenti ambiti di applicazione:

One use of a corpus is to extract a language model: a list of weighted words, or combinations of words, that describe (1) how words are related, (2) how they are used with each other, and (3) how common they are in a given domain. Language models are used in speech processing to predict which word combinations are likely interpretations of a sound stream, in information retrieval to decide which words are useful indicators of a topic, and in machine translation to identify good translation candidates.

Questa definizione lascia intendere che attraverso l'uso dei corpora sia possibile ricavare un modello del funzionamento della lingua semplicemente osservando, ad esempio, quali siano le parole più utilizzate, come le parole si combinino le une alle altre e quali combinazioni di parole siano più comuni. Tali modelli linguistici trovano applicazione in vari ambiti: citandone a titolo di esempio alcuni che hanno delle analogie con la presente ricerca, l'estrazione di elementi linguistici dai corpora è applicata anche agli studi di lessicografia (cfr. Wunderlich 2018) e fraseologia (cfr. Cosme, Gilquin 2008, Coxhead 2008, Mudraya *et al.* 2008, Babych 2012, Orenha-Ottaiano 2012, Seracini 2018 e gli studi di Pecman 2008, Wible 2008 e Alhujaylan 2014, che si collocano in entrambi gli ambiti sopra nominati) e come supporto nell'apprendimento delle lingue (cfr. Curado-Fuentes 2012, Charkova, Charkova 2014, Vališová 2014, Brezina, Gablasova, Marín Cervantes 2018, Curry, Goodman 2018, Hadley, Hadley 2018, Picoral *et al.* 2018, Tono 2018; in particolare per l'italiano L2/LS: Corino, Marelli 2009, Merlo 2009, Kennedy, Miceli 2010, Guidetti, Lenzi, Storchi 2012, Lyding *et al.* 2013 e Ivanovska-Naskova 2018).

L'uso dei corpora come supporto nell'apprendimento delle lingue verrà indagato nella sezione 3.6, in cui verranno presentati degli studi che propongono delle strategie per favorire l'uso guidato o diretto dei corpora da parte degli apprendenti (dalle operazioni più semplici, come le liste di frequenza e la ricerca delle concordanze, alla padronanza di modalità di ricerca più avanzate) allo scopo di introdurre i corpora come strumento di studio per formulare e testare ipotesi sul funzionamento della lingua

La consultazione di una grande quantità di dati contenuta in un corpus, come descritto da McEnery e Hardie (2012: 2), è resa possibile solo attraverso alcuni strumenti informatici. A questo proposito, come verrà qui di seguito illustrato nella citazione di McEnery e Hardie (2012), sebbene nel corso della storia della linguistica dei corpora vi siano stati casi di risorse in formato cartaceo, o le cui dimensioni ridotte abbiano consentito una consultazione non elettronica, attualmente, come ad esempio nella definizione di Baroni (2010) sopra menzionata, si fa riferimento principalmente a testi in formato elettronico:

It should be noted that the term text here denotes a file of machine-readable data. [...] Although we have said that corpus linguistics always uses machine-readable text, in fact, historically, much work was undertaken on corpora held in paper form; for example Fries (1952) produced a grammar of English based upon such a corpus. Also, while it is true that

much research using corpus methods [...] uses corpora of millions of words, there are others studies such as those of Ghadessy and Gao (2001) and McEnery and Kifle (2001) which, appropriately, use smaller, specialised corpora that might conceivably have been analysed by hand and eye. Nonetheless, despite the exceptions, the generalisations above characterise much of the work that can reasonably be described as corpus linguistics. (McEnery, Hardie 2012: 2-3)

Vi sono vari strumenti informatici che mettono a disposizione diverse modalità di ricerca. Limitando gli esempi alle sole funzioni utilizzate nell'ambito di questa ricerca, si ricordano la *concordanza* e la *lista di frequenza*. Queste ultime rappresentano, come affermato da McEnery e Hardie (*ibidem*), due tipi di analisi che rivestono pari importanza nella linguistica dei corpora, ossia l'analisi qualitativa e quantitativa, alle quali si farà riferimento (cfr. capitolo 5) nell'ambito di questa ricerca.

Nello specifico, con la funzione concordanza la parola cercata può essere visualizzata nel contesto in cui è stata utilizzata (dalla singola frase, fino ad allargare il contesto ad un intero paragrafo o all'intero testo in cui è contenuta, a seconda delle impostazioni dello strumento di ricerca in uso), mentre con la *lista di frequenza* è possibile ottenere una lista di tutte le parole contenute in un corpus, per ciascuna delle quali viene indicata la frequenza, ossia il numero di volte in cui ricorre.

Ad esempio, i corpora possono essere utilizzati per indagare la frequenza d'uso di una parola (ad esempio, in un dato contesto oggetto di ricerca, quanto è frequente la parola *questione* rispetto a *problema*?), per individuare le collocazioni tipiche in cui questa ricorre (ad esempio, a quali verbi si accompagna la parola *questione*? *Sollevare una questione*, *sottoporre una questione a qualcuno*, ecc.) e per cogliere la prosodia semantica² sulla base del contesto d'uso. Un corpus elettronico è in grado di dare risposta a queste domande tramite alcune modalità di ricerca, tra cui quelle sopra citate. Dopo aver, dunque, definito cosa sia un corpus e avere descritto in linee generali alcuni possibili usi, nella sezione successiva si illustreranno brevemente i due principali metodi di applicazione dei corpora allo studio del linguaggio, al fine di identificare il metodo più adeguato alla realizzazione degli obiettivi della presente ricerca.

² Descrive il modo in cui alcune parole possono essere percepite con connotazioni positive o negative, poiché tendono a ricorrere assieme ad altre parole che appartengono ad una determinata area semantica. “[A] word may be said to have a particular semantic prosody if it can be shown to co-occur typically with other words that belong to a particular semantic set” (Hunston, Francis 2000: 137).

3.2. Metodo per l'identificazione dei fenomeni tramite corpus

Come illustrato nella sezione precedente, la linguistica dei corpora è un ambito di ricerca che si è sviluppato negli ultimi anni del ventesimo secolo all'interno degli studi sulla variazione e l'uso della lingua (Biber, Conrad, Reppen, 1998, Biber 2012) impiegando principalmente due metodi di ricerca.

Nel primo, definito *corpus-based* (Tognini-Bonelli 2001), ossia basato su corpus, i dati estratti da un corpus vengono utilizzati per verificare una teoria, con l'obiettivo di convalidarla, confutarla o perfezionarla. In questo approccio la linguistica dei corpora è intesa, dunque, come metodo per lo studio del linguaggio.

The term corpus-based is used to refer to a methodology that avails itself of the corpus mainly to expound, test or exemplify theories and descriptions that were formulated before large corpora become available to inform language study. [...] Corpus-based linguists adopt a 'confident' stand with respect to the relationship between theory and data in that they bring with them models of language and descriptions which they believe to be fundamentally adequate, they perceive and analyse the corpus through these categories and sieve the data accordingly. (ivi: 65-66)

Allo studio del linguaggio con approccio *corpus-based* si oppone l'approccio *corpus-driven*, ossia guidato dal corpus, di tipo induttivo. Non vi è una teoria a priori; i dati sono osservati e studiati nel loro contesto, in modo tale che “parlino da sé” e sia possibile derivarne una teoria sull'uso della lingua. Tale approccio rifiuta la concezione della linguistica dei corpora come metodo e afferma che la sola fonte per la teoria del linguaggio sia il corpus stesso:

[i]n a corpus-driven approach the commitment of the linguist is to the integrity of the data as a whole, and descriptions aim to be comprehensive with respect to corpus evidence. The corpus, therefore, is seen as more than a repository of examples to back pre-existing theories or a probabilistic extension to an already well-defined system. The theoretical statements are fully consistent with, and reflect directly, the evidence provided by the corpus. [...] The theory has no independent existence from the evidence and the general methodological path is clear: observation leads to hypothesis leads to generalisation leads to unification in theoretical statement. (Tognini-Bonelli 2001: 84-85)

La dicotomia *corpus-based* e *corpus-driven* divide la ricerca in due approcci distinti, sebbene McEnery e Hardie (2001:151) rifiutino di considerarla una vera e propria distinzione binaria, sostenendo che vi siano casi in cui l'utilizzo del corpus si collochi a metà strada tra entrambi gli approcci (“the corpus-based versus corpus-driven distinction implies a dichotomy where there is actually a sliding scale”) e che a separare tali approcci vi sia in realtà solo la diversa concezione del *ruolo* ricoperto dal corpus (metodo per verificare/confermare/confutare una teoria preesistente o fonte da cui indurre una teoria) piuttosto che il loro *uso* effettivo:

we would argue that the different schools of corpus linguistics are not reliably distinguished – or, alternatively, that their nature is not optimally communicated – by the corpus-based versus corpus-driven distinction, or by the ‘top-down’ versus ‘bottom-up’ distinction with which it is often equated. Rather, it is the contrasting stances on the conceptual status of the corpus and of corpus linguistics – as having theoretical status versus as a linguistic methodology – that truly separates the two schools. (*ibidem*)

Per tali ragioni McEnery e Hardie trovano più adeguata la distinzione *corpus-linguistics-as-method* vs *corpus-linguistics-as-theory* rispetto a *corpus-based* vs *corpus-driven* (ivi: 6).

Sulla base di tali riflessioni, l'approccio da utilizzare per la presente ricerca è di tipo *corpus-based* o *corpus-linguistics-as-method*, ove un corpus di lingua italiana è utilizzato come risorsa di verifica dell'ipotesi circa l'effettiva reperibilità – e in quale misura – dei tratti dell'italiano dell'uso medio nei testi scritti di media formalità. Dopo aver illustrato i due principali metodi di applicazione dei corpora allo studio del linguaggio, nella sezione successiva si illustreranno brevemente le principali tipologie di corpora esistenti.

3.3. Tipologie di corpora esistenti

Una prima distinzione tracciabile tra le diverse tipologie di corpora esistenti si basa sulla lingua dei testi raccolti. Nello specifico, esistono corpora monolingui e corpora bilingui (o multilingui), i quali sono ulteriormente distinguibili in corpora paralleli e corpora comparabili.

I corpora monolingui contengono testi redatti in un'unica lingua, il che non esclude che possano rappresentarne le diverse varietà esistenti, ad esempio nel caso di un corpus di lingua inglese che includa la varietà britannica, americana ecc.: “So the International

Corpus of English [...], for example, is a large monolingual corpus – it represents one language, English, though it allows linguists to compare and contrast a number of international varieties of that language.” (McEnery, Hardie 2012: 18)

Si definiscono, invece, corpora multilingui le risorse che coinvolgono tre o più lingue, mentre le risorse che coinvolgono solo due lingue vengono convenzionalmente definite bilingui (ivi:19).

Inoltre, i corpora che coinvolgono più di una lingua possono essere ulteriormente classificati come segue, sulla base delle teorie di Baker (1993: 248, 1995, 1999), McEnery e Wilson (2001: 70) e Hunston (2002: 15) riportate in McEnery e Hardie (2012: 19): si definiscono corpora paralleli quelle risorse che contengono dei testi sia nella loro lingua originaria, sia nelle relative traduzioni in una o più lingue. Nel caso in cui le unità linguistiche dei testi (ad esempio le frasi) nella loro lingua originaria vengano collegate alle unità linguistiche dei testi tradotti, le risorse vengono definite corpora paralleli allineati. Ad esempio, l’Opus Corpus³ è il risultato di un progetto di raccolta e allineamento di testi tradotti tratti dal web, con l’aggiunta di annotazioni linguistiche, allo scopo di fornire un corpus parallelo liberamente accessibile, l’EuroParl⁴, costituito dagli atti del Parlamento Europeo dal 1996 al 2011.

Si parla, invece, di corpora paragonabili nel caso in cui vi siano due o più corpora monolingui, (ad esempio: un corpus monolingue italiano e un corpus monolingue di qualunque altra lingua, tra i quali non vi sia un rapporto “testo di partenza – traduzione”) che siano tra loro paragonabili sulla base di variabili quali ad esempio la tipologia testuale, il dominio/argomento, la data di produzione, la varietà linguistica di riferimento, il registro o altro. I corpora paragonabili consentono di confrontare due o più lingue rispetto allo stesso genere testuale o tema.

Un’ulteriore distinzione tracciabile tra le diverse tipologie di corpora è quella tra corpora di riferimento e corpora specialistici:

Un corpus di riferimento si propone come campione rappresentativo di una lingua in tutti i suoi aspetti: scritto e parlato, registri e varietà d’uso, ecc. Questo obiettivo però non è mai pienamente raggiungibile, per ragioni sia pratiche (la raccolta dei dati richiederebbe costi enormi) che teoriche (quanti e quali registri e varietà andrebbero campionati? in che

³ <http://opus.nlpl.eu> (Consultato il 05/06/2020).

⁴ <http://statmt.org/europarl> (Consultato il 05/06/2020).

proporzione? ecc.). Tuttavia, più un corpus è composto da fonti variegata, più esso sarà rappresentativo di una lingua in generale. (Baroni 2010)

La denominazione dipende, dunque, dal fatto che tale tipologia di corpora si propone come risorsa di riferimento per la descrizione di una certa lingua ed è, dunque, trasversale rispetto alle diverse varietà che la compongono. Ciò non esclude che un corpus di riferimento possa essere organizzato in vari sottocorpora specialistici, rappresentativi delle diverse varietà della lingua.

A questo proposito, i corpora specialistici hanno, appunto, l'obiettivo di rappresentare particolari varietà ristrette di una lingua, come ad esempio i corpora per lo studio di domini linguistici settoriali (ad esempio, il gergo di determinati settori professionali) o i corpora di testi che rientrano esclusivamente in una specifica tipologia testuale, eccetera. Di conseguenza, poiché il loro scopo è di tracciare un quadro di una varietà linguistica specifica, i corpora specialistici sono caratterizzati da fonti meno variegata rispetto ai corpora di riferimento e da un'elevata omogeneità linguistica e sono queste due caratteristiche che ne determinano il grado di rappresentatività.

Ulteriori tipologie esistenti sono i corpora sincronici, che contengono testi appartenenti tutti ad una stessa finestra temporale e, pertanto, descrivono un particolare stadio del linguaggio; al contrario i corpora diacronici contengono testi che appartengono a diverse finestre temporali e, quindi, descrivono i cambiamenti di una lingua attraverso il tempo.

Infine, vi è la distinzione tra corpora di monitoraggio e corpora bilanciati.

The monitor corpus approach, proposed most notably by John Sinclair, seeks to develop a dataset which grows in size over time and which contains a variety of materials. The relative proportions of the different types of materials may vary over time. (McEnery, Hardie 2012: 6)

Dunque, un corpus di monitoraggio è una raccolta di testi che cresce nel tempo attraverso l'introduzione di nuovi testi selezionati secondo gli stessi criteri usati per costruire la raccolta iniziale. Di conseguenza, poiché vengono aggiornati nel corso del tempo, i corpora di monitoraggio sono adatti a seguire le evoluzioni che avvengono in una lingua.

In contrast to monitor corpora, balanced corpora, also known as sample corpora, try to represent a particular type of language over a specific span of time. In doing so, they seek to be balanced and representative within a particular sampling frame which defines the type of

language, the population that we would like to characterise. The population is the notional space within which language is being sampled. So, for example, if we want to look at the language of service interactions in shops in the UK in the late 1990s, the sampling frame is clear –we would only accept data into our corpus which represents service interactions in UK shops in the 1990s. (ivi: 8)

I corpora bilanciati (*balanced corpora* o *sample corpora*) includono, dunque, testi che rappresentano una varietà linguistica specifica “fotografata” in un preciso confine temporale. A tale scopo, per costruire un corpus bilanciato è necessario definire accuratamente la popolazione linguistica di interesse, tracciandone i confini spaziali e temporali (quali testi sono inclusi o esclusi dalla popolazione) e la tipologia dei testi (l'articolazione in strati della popolazione). La differenza tra corpora di monitoraggio e corpora bilanciati può essere sintetizzata come segue:

Two broad approaches to the issue of choosing what data to collect have emerged: the monitor corpus approach (see Sinclair 1991: 24–6), where the corpus continually expands to include more and more texts over time; and the balanced corpus or sample corpus approach (see Biber 1993 and Leech 2007), where a careful sample corpus, reflecting the language as it exists at a given point in time, is constructed according to a specific sampling frame. (ivi: 6)

A conclusione di questa panoramica sulle principali tipologie di corpora esistenti, nella sezione successiva verrà descritto l'uso del web come corpus.

3.4. Il web come corpus

Kilgarriff e Grefenstette (2003) hanno proposto l'approccio del *Web as a Corpus*, ossia dell'uso del web come un enorme corpus di testi. “The web is clearly a multilingual corpus” (Kilgarriff, Grefenstette 2003: 334): il web può, infatti, essere utilizzato sia per ottenere dati di tipo quantitativo sulla frequenza d'uso delle espressioni linguistiche tramite i motori di ricerca (ad esempio, quante pagine web si ottengono attraverso la ricerca di una data parola), sia come una fonte inesauribile di testi per costruire corpora. In quest'ottica, il web è un corpus che presenta numerosi vantaggi: è una risorsa enorme, facilmente accessibile, oltre che gratuita, e variegata negli stili, nelle tipologie testuali e nei contenuti.

Sebbene tanto per i corpora quanto per il web il punto di partenza sia il medesimo, ossia un'enorme raccolta di dati in continua crescita, in realtà l'approccio del *Web as a*

Corpus presenta alcuni limiti (Kilgarriff, Grefenstette 2003: 344-345), come evidenziato anche da McEnery e Hardie (2012: 7).

Innanzitutto, nel caso in cui per *web as a corpus* si intenda nello specifico l'interrogazione del web tramite l'uso di motori di ricerca commerciali⁵, vi è la mancanza di replicabilità delle ricerche: i risultati delle ricerche effettuate tramite motori di ricerca commerciali, come ad esempio *Google, Yahoo!* ecc., non sono sempre affidabili, poiché dipendono dalle specifiche caratteristiche di tali motori. Uno stesso tipo di ricerca può, infatti, condurre a risultati diversi sia tramite motori di ricerca differenti, sia tramite un unico motore di ricerca a causa degli algoritmi di indicizzazione e di recupero delle pagine e della presenza di duplicati, ragion per cui i dati di frequenza risultano inaffidabili e non paragonabili. Inoltre, i risultati corrispondono al numero di pagine web reperite ("pagina contenente *x*") e non all'effettiva frequenza dell'espressione linguistica ricercata.

Sebbene i motori di ricerca siano degli eccellenti strumenti dal punto di vista della loro funzione primaria, ossia la ricerca di informazioni, essi possono risultare inadeguati alle esigenze dei linguisti:

The default means of access to the Web is through a search engine such as Google. Although the Web search engines are dazzlingly efficient pieces of technology and excellent at the task they set for themselves, for the linguist they are frustrating. (Kilgarriff, Grefenstette 2003: 344)

Le ragioni citate dai due studiosi sono le seguenti (ivi: 345). Il web presenta l'impossibilità sia di fare ricerche in sottoinsiemi di testi, poiché il contenuto dei testi presenti nel web non è suddiviso per generi testuali, sia di fare ricerche tramite criteri prettamente linguistici, ad esempio fare ricerche per parti del discorso (sostantivi, verbi, aggettivi, ecc.) e lemmi ecc., senza contare che il contesto fornito per ciascuna occorrenza è molto ridotto (circa 10 parole su *Google*). Infine, il web include testi qualitativamente disomogenei dal punto di vista dell'accuratezza nella pianificazione testuale, molti dei quali non sono privi di errori di ogni genere.

Dunque, sebbene il web sia di fatto un immenso corpus ("The answer to the question "Is the web a corpus?" is yes." (ivi: 334)), le caratteristiche dei motori di ricerca non lo rendono adeguato ad un suo utilizzo come corpus linguistico, ma solo come fonte di testi a

⁵ In contrapposizione al caso in cui i testi vengono scaricati dal web e poi processati per la costruzione di un corpus da consultare tramite specifici programmi di interrogazione.

cui attingere per la costruzione di corpora da consultare tramite specifici programmi di interrogazione. I vantaggi principali di questo approccio sono, ad esempio, l'offerta illimitata di testi attuali e appartenenti a generi e domini diversi, in cui sono rintracciabili i vari registri linguistici, la disponibilità immediata dei testi e il fatto che siano reperibili in formato digitale.

3.5. Corpora nella linguistica cognitiva e nella *construction grammar*

Sulla base di quanto illustrato, i corpora rappresentano delle risorse importanti per la conoscenza della lingua, la cui consultazione tramite specifici programmi di interrogazione permette, ad esempio, di valutare la frequenza d'uso di elementi lessicali, costruzioni grammaticali e molto altro, fornendo preziose fonti di informazioni sull'evoluzione e lo stato corrente della lingua in uso. A questo proposito, si ricorda come l'utilizzo di corpora come base di dati linguistici reali, utili all'analisi e alla teorizzazione linguistica, venga applicato in ambiti quali la linguistica cognitiva e la *construction grammar*. Come evidenzia Bazzanella (2014), la linguistica cognitiva si affianca ad un ricorso sempre maggiore ai corpora, poiché le strutture linguistiche emergono dall'uso stesso della lingua, secondo i cosiddetti modelli basati sull'uso (*usage-based models*):

An important consequence of adopting the usage-based thesis is that there is no principled distinction between knowledge of language and use of language (competence and performance in Generative Grammar terms), since knowledge of language *is* knowledge of how language is used. (Evans 2007: 217)

Inoltre, come riportato da Newman (2011) in una riflessione sullo stato attuale della linguistica cognitiva, in particolare sulle sue metodologie, Geeraerts (2006: 29) fa riferimento alla tendenza crescente della linguistica cognitiva a sottolineare la sua essenza di linguistica basata sull'uso ("growing tendency of Cognitive Linguistics to stress its essential nature as a usage-based linguistics"), una tendenza che si inserisce chiaramente nella ricerca basata su corpora, come dimostrano alcune pubblicazioni che illustrano metodi basati su corpora in ambito cognitivo-linguistico, tra cui Gries, Stefanowitsch (2006), Stefanowitsch, Gries (2006), Lewandowska-Tomaszczyk, Dziwirek (2009).

Analogamente, per quanto riguarda la *construction grammar*, Hilpert (2013) afferma che i dati recuperati dai corpora possono rivelare dettagli precisi riguardo ai cambiamenti

linguistici di tipo costruttivo (ad esempio, i cambiamenti relativi alla frequenza di una costruzione linguistica, alla sua forma e alla sua funzione), facilitando l'analisi quantitativa e, nel contempo, mostrando come la linguistica dei corpora abbia influenzato in modo significativo la comprensione dei cambiamenti linguistici di tipo costruttivo (cfr. gli approcci *corpus-based* alla grammatica costruttiva in Gries, Yoon (2016); Hoffmann, Trousdale (2013) e Gries, Wulff (2005) sull'applicazione dei corpora allo studio delle costruzioni linguistiche degli apprendenti di una lingua straniera).

Dopo aver illustrato alcuni concetti base della linguistica dei corpora e le ragioni per le quali i corpora possano rappresentare delle valide risorse per la conoscenza della lingua, nella sezione successiva si illustreranno alcuni ambiti di applicazione dei corpora alla didattica delle lingue, dedicando particolare attenzione all'uso dei corpora per la realizzazione di materiali didattici.

3.6. Corpora e apprendimento delle lingue

In questa sezione verrà affrontato il tema dell'applicazione dell'uso dei corpora all'insegnamento delle lingue. Verranno esposti brevemente gli ambiti in cui tale applicazione ha luogo, dando particolare rilievo all'utilizzo dei corpora per il reperimento di testi da utilizzare per la progettazione di attività didattiche. A tale scopo verranno citati alcuni studi nei quali si propone l'uso dei corpora per la realizzazione di risorse lessicografiche e grammaticali a supporto dell'apprendimento linguistico; si escluderà invece dalla trattazione l'impiego dei corpora nella traduzione. Infine, si affronterà il tema dell'importanza dell'autenticità degli esempi linguistici impiegati nella didattica, in contrapposizione agli esempi creati *ad hoc*, e della varietà di lingua da trasmettere agli apprendenti.

L'applicazione dell'approccio *corpus-based* alla linguistica e all'insegnamento delle lingue ha assunto un'importanza sempre maggiore tra la fine degli anni '80 e i primi anni '90 del secolo scorso (McEnery, Xiao 2005: 364), con i primi studi sul tema di Sinclair (1991) e Johns (1991). Inoltre, a partire dal 1994 la serie internazionale di conferenze biennali *Teaching and Language Corpora* (TaLC), giunta al momento della scrittura alla sua quattordicesima edizione, testimonia il forte interesse per l'uso dei corpora nell'insegnamento delle lingue.

Nel corso degli anni, come affermato da Römer (2008), si è sviluppato dunque un rapporto bidirezionale tra la linguistica dei corpora e la glottodidattica: da un lato la

linguistica dei corpora ha permesso di ricavare dati sull'uso della lingua e ha fornito strumenti tecnologici, quali programmi per la creazione di corpora, liste di frequenza, concordanze e corpora di grandi dimensioni, che sono utili a chi studia una lingua; dall'altro lato la glottodidattica ha incentivato, ad esempio, la produzione dei *learner corpora*, che comprendono produzioni orali e/o scritte degli apprendenti di una lingua straniera.

Con l'obiettivo di indagare il rapporto tra corpora e apprendimento, gli atti del convegno TaLC del 2002 (Aston, Bernardini, Stewart 2004), furono divisi in tre possibili ambiti di applicazione dei corpora alla didattica: 1) *corpora by learners*: i corpora composti da materiali prodotti dagli apprendenti e finalizzati, ad esempio, allo studio dell'interlingua (Selinker 1972) e dell'"errore", con l'obiettivo di identificare le caratteristiche della lingua degli apprendenti e utilizzarle come punti di partenza per riprogrammare l'attività didattica e favorire l'acquisizione linguistica; 2) *corpora for learners*: corpora progettati per permettere ai docenti e, più generalmente, a coloro che creano materiali didattici, di trasmettere una descrizione più accurata della lingua oggetto di studio e, pertanto, di decidere cosa debba essere appreso; 3) *corpora with learners*: i corpora sono considerati degli strumenti utili ad acquisire conoscenze e competenze linguistiche tramite il loro uso diretto da parte degli studenti, con o senza l'aiuto dei loro insegnanti.

Tale tripartizione degli ambiti d'uso ricalca parzialmente quella elaborata da Leech (1997): a) uso indiretto dei corpora nell'insegnamento, che prevede la consultazione dei corpora al fine di ricavare esempi autentici e tipici dell'uso di elementi linguistici, nonché di informazioni sulla loro frequenza d'uso. Tutto ciò è finalizzato allo sviluppo di materiali e test linguistici e alla compilazione di pubblicazioni di riferimento, siano esse grammatiche o dizionari, poiché l'uso di corpora permette di tenere conto delle variazioni di registro e delle differenze tra lingua scritta e parlata e, pertanto, di ottenere descrizioni grammaticali più accurate rispetto a quelle comunemente fornite dalle grammatiche (McEnery, Xiao 2005); 2) uso diretto dei corpora nell'insegnamento, che si articola, secondo la tripartizione di Leech (1997) in *teaching about*, *teaching to exploit* e *exploiting to teach*:

"Teaching about" means teaching corpus linguistics as an academic subject like other sub-disciplines of linguistics such as syntax and pragmatics. [...] "Teaching to exploit" means providing students with "hands-on" know-how, as emphasized in McEnery, Xiao and Tono

(2006), so that they can exploit corpora for their own purposes. Once the student has acquired the necessary knowledge and techniques of corpus-based language study, the learning activity may become student centred. “Exploiting to teach” means using a corpus-based approach to teaching language and linguistics courses (e.g. sociolinguistics and discourse analysis), which would otherwise be taught using noncorpus-based methods. (McEnery, Xiao 2011: 370)

3) ulteriori sviluppi orientati all'insegnamento: corpora di linguaggi per scopi specifici (LSP), corpora incentrati sullo sviluppo della prima lingua (L1) e corpora degli apprendenti di seconda lingua (L2).

Sulla base di tali suddivisioni delle possibili applicazioni dei corpora alla didattica, è possibile affermare che la presente ricerca, avente tra gli obiettivi l'uso di corpora come fonte da cui ricavare testi autentici sui quali incentrare la progettazione di materiali didattici, si colloca nell'ambito dell'uso indiretto dei corpora nell'insegnamento, secondo la terminologia di Leech (1997). Partendo da tale presupposto, nella sezione seguente verranno citati alcuni studi nei quali si affronta l'uso dei corpora per la realizzazione di risorse a supporto dell'apprendimento linguistico.

3.7. Uso dei corpora per la creazione di materiali didattici

Come già illustrato nella sezione precedente, a partire dalla metà degli anni '90 sono stati svolti numerosi studi circa l'uso con finalità didattiche di testi o singoli elementi linguistici (lessico, verbi, fraseologie ecc.) estratti dai corpora. In particolare, in questa trattazione si farà riferimento alle pubblicazioni contenute in Sinclair (2004), Aijmer (2009), Granger, Meunier (2008), Meunier, Granger (2008), Hinkel (2011) e agli atti dei convegni TaLC a partire dalla pubblicazione del 2004 di Aston, Bernardini e Stewart fino al 2018, raggruppando gli studi per aree tematiche.

Tra gli studi che gettano le basi per la futura progettazione di materiali didattici si citano Cosme e Gilquin (2008), che sottolineano le ricadute positive che l'analisi contrastiva dei dati ricavati da un corpus bilingue inglese-francese sull'uso delle preposizioni *with* e *avec* può avere nell'insegnamento delle lingue, nella lessicografia bilingue e nella compilazione di dizionari. Alle stesse conclusioni giungono anche Mudraya *et al.* (2008) nel loro studio comparativo *corpus-based* dei *phrasal verbs* inglesi e delle loro espressioni equivalenti in russo, di cui evidenziano le possibili applicazioni all'insegnamento delle lingue, alla linguistica contrastiva e all'estrazione di lessico

multilingue. Babych (2012) descrive l'impiego di uno strumento di ricerca basato sul Corpus Workbench (Christ, 1994a) per reperire, analizzare e valutare strutture lessico-grammaticali aventi la funzione di connettori testuali che occupano la posizione iniziale di paragrafo o frase. Dalla consultazione di alcuni corpora in lingua inglese, russa e ucraina sono emerse delle liste di connettori che hanno permesso di ampliare dell'80% le tradizionali liste di connettori testuali presenti nelle grammatiche, dimostrandone la limitatezza originaria.

Vališová (2014) presenta un corpus specializzato costituito da manuali di lingua ceca come lingua straniera e illustra come il confronto della lingua usata nei manuali con la lingua usata nei corpora generalisti apra delle riflessioni sulla lingua insegnata agli apprendenti e sui possibili miglioramenti dei materiali didattici esistenti.

Particolarmente frequente risulta l'applicazione dei corpora alla descrizione grammaticale e all'insegnamento delle collocazioni e della fraseologia. Coxhead (2008:156) nel suo studio sulla preparazione dei materiali didattici dedicati all'inglese accademico (*English for Academic Purposes*, EAP) cita lo studio di Jones e Haywood (2004), nel quale, tra le varie attività proposte per l'apprendimento della fraseologia inglese, vi sono anche delle lezioni incentrate sulla produzione scritta, in cui agli studenti vengono fornite delle espressioni fisse ricorrenti (*target formulaic sequences*) estratte da un corpus composto da testi tratti da manuali di scrittura accademica. Successivamente, viene chiesto di analizzare quelle espressioni ricorrenti nel contesto più ampio del testo da cui sono tratte e, infine, di svolgere degli esercizi di completamento (*fill in the gap*). Lo studio di Pecman (2008) illustra, invece, la creazione di un corpus parallelo inglese-francese finalizzato a reperire la fraseologia tipica dell'EAP in ambito scientifico e tecnologico e a creare un dizionario elettronico bilingue della fraseologia. Wible (2008: 176) illustra lo sviluppo dello strumento *Collocator*, basato sui materiali del *British National Corpus* e finalizzato all'impiego in aula: durante la lettura di testi sul web, gli studenti possono ottenere delle collocazioni semplicemente evidenziando delle parole del testo e cliccando poi sull'icona del *Collocator*, che mostrerà le collocazioni che interessano le parole evidenziate. Orenha-Ottaiano (2012) presenta alcune attività, esercizi e materiali basati su corpus che si concentrano su collocazioni e destinati ad apprendenti di L1 portoghese. Alhujaylan (2014) propone attività basate su collocazioni estratte dal corpus Collins WordBanks, mentre Seracini (2018) propone attività sulla fraseologia estratta da un corpus specialistico composto da testi di area economica tradotti dall'italiano all'inglese da studenti universitari del settore dell'*international management*.

Circa il potenziamento dell'uso del lessico, Wunderlich (2018) illustra la creazione di elenchi di termini estratti da un corpus specialistico costituito da articoli accademici del settore dell'informatica aziendale a favore degli apprendenti della disciplina, mentre Tono (2018) propone lo sviluppo di risorse per l'insegnamento delle lingue basate su corpora e sul Quadro Comune Europeo di Riferimento per le Lingue, implementabili attraverso alcuni sistemi di e-learning, tra cui si citano a titolo di esempio un'app per smartphone per la creazione di flash-card lessicali e un'applicazione di assistenza alla traduzione che include il rilevamento di errori.

Per il potenziamento delle abilità di lettura e comprensione testuale si citano Curado-Fuentes (2012), che presenta la progettazione di test e attività di comprensione di testi scritti basati su parole chiave estratte da due corpora di testi giornalistici in lingua inglese e spagnola e Hadley e Hadley (2018), che propongono attività di lettura mirate al potenziamento della conoscenza lessicogrammaticale ricettiva e ad una maggiore velocità di lettura di apprendenti giapponesi utilizzando materiali estratti da un corpus costituito dai testi della collana Oxford Bookworms Graded Readers.

Nell'ambito della formazione dei docenti circa la realizzazione di materiali didattici derivati da corpora, Charkova e Charkova (2014) illustrano come effettuare su corpora ricerche di carattere morfologico, lessicale e grammaticale e come utilizzare i testi estratti dai corpora per produrre materiali didattici. Si citano, infine, i seminari svolti nell'ambito del TaLC13 da Curry e Goodman (2018), che si focalizzano sull'impatto che i dati estratti dal Cambridge Learner Corpus (CLC) di lingua scritta possono avere sullo sviluppo di materiali didattici, mentre Picoral, Staples, Shin e Swatek (2018) e Brezina, Gablasova e Marín Cervantes (2018) illustrano delle procedure di creazione di attività didattiche basate su dati estratti da corpora e riguardanti, rispettivamente, l'intertestualità negli scritti degli apprendenti di lingua inglese L2 e l'insegnamento della sociolinguistica.

Per quanto riguarda l'applicazione dei corpora alla didattica dell'italiano L2/LS, nel corso degli anni Duemila si è registrata una forte crescita di corpora di vario genere, come riportato negli studi di Ivanovska-Naskova (2018), Barbera (2015, a c. di), Panunzi, Cresti e Gregori (2013) e Baroni (2010). Di particolare importanza è lo sviluppo di svariati *learner corpora* costituiti da testi prodotti dagli apprendenti di italiano come L2 o LS (cfr. Ivanovska-Naskova 2018: 73, Giacalone Ramat, Chini e Adorno 2013: 179). Tuttavia, nonostante questa tendenza positiva, esiste un divario tra i corpora dell'italiano e il loro utilizzo nell'insegnamento della lingua e gli studi su questo tema sono ancora scarsi (cfr. Ivanovska-Naskova 2018: 73). Al di là degli studi di carattere generale sulle potenzialità

dei corpora di Corino (2014) e Viganò (2011), le ricerche che illustrano proposte ed esperienze didattiche per l'impiego in aula dei corpora dell'italiano L2/LS o per la formazione dei docenti sono poco numerose (cfr. Ivanovska-Naskova 2015).

Di seguito si citano a titolo di esempio alcuni studi, tutti caratterizzati dalla proposta dell'uso diretto dei corpora da parte degli apprendenti, dando particolare rilievo nell'esposizione a quelli che presentano alcune analogie con la presente ricerca, ossia che in aggiunta propongono l'uso dei corpora come fonte per la produzione di materiali didattici.

Lo studio di Guidetti, Lenzi e Storchi (2012) propone una serie di strategie per portare progressivamente gli apprendenti dell'italiano come LS ad un utilizzo sempre più autonomo dei corpora. Dopo aver tracciato una panoramica sui principali utilizzi dei corpora nella glottodidattica, viene proposto l'utilizzo di tre risorse giudicate dagli autori di facile utilizzo e caratterizzate da un'interfaccia *user-friendly*, ossia *La Repubblica Corpus*, il *Corpus LIP* e il programma di interrogazione di corpora *Sketch Engine* (per ciascuno di essi si rinvia al capitolo 4). Nello studio viene proposto un percorso graduato che parte con attività basate sull'uso dei corpora guidato dal docente, fino ad arrivare ad un uso autonomo come strumento di studio per ricercare la soluzione a quesiti sul funzionamento della lingua e che richiedono la padronanza di diversi stili di interrogazione dei corpora. Partendo con la produzione di liste di frequenza e concordanze, che permettono di identificare delle strutture linguistiche ricorrenti, gli apprendenti assumono gradualmente il ruolo di apprendenti-ricercatori che, sulla base dei dati raccolti dalla consultazione del corpus, traggono delle generalizzazioni sul funzionamento della lingua, consolidando e/o perfezionando così le loro conoscenze.

Kennedy e Miceli (2010) propongono l'uso del corpus CWIC (*Contemporary Written-Italian Corpus*) come risorsa utile al potenziamento della competenza scritta degli studenti universitari. Partendo dalle difficoltà riscontrate da studenti di livello intermedio di italiano LS nell'uso dei corpora come supporto all'apprendimento linguistico, le due studiose hanno progettato un percorso di apprendimento mirato allo sviluppo delle capacità di ricerca su corpus che sono necessarie per formulare e testare ipotesi sull'uso della lingua. Il percorso ha la durata di un semestre; nella fase iniziale l'uso dei corpora viene introdotto come risorsa di supporto alla produzione scritta da affiancare all'uso tradizionale del dizionario bilingue, per poi essere applicato gradualmente per la risoluzione di dubbi e problemi grammaticali sempre più specifici. Attingendo alle intuizioni dei tre partecipanti al *case study* in questione, viene definito il percorso di approccio alla consultazione dei

corpora da replicare in successive sperimentazioni e vengono individuati alcuni perfezionamenti da apportare rispetto al percorso iniziale.

A differenza degli studi precedenti, la ricerca di Corino e Marellò (2009) è incentrata sull'uso diretto dei *learner corpora*. Nello specifico si propone l'uso di VALICO⁶ (Varietà Apprendimento Lingua Italiana Corpus Online), un corpus di testi scritti da apprendenti stranieri di italiano, e di VINCA⁷ (Varietà di Italiano di Nativi Corpus Appaiato). Entrambe le risorse fungono da strumenti di supporto alla formazione di futuri insegnanti di italiano L2/LS e all'insegnamento della morfologia italiana a stranieri.

Una prima applicazione del corpus VALICO è la sua consultazione mirata all'individuazione degli errori più frequenti commessi dagli apprendenti stranieri. La selezione nella maschera di ricerca di alcuni dati sociolinguistici, quali la L1 degli apprendenti, l'anno di studio della lingua italiana, la conoscenza di altre lingue ecc. permette di creare dei sottocorpora con i testi prodotti dalla popolazione selezionata. In tal modo è possibile analizzare le caratteristiche dei testi ottenuti e ottenere dati funzionali alle esigenze della formazione di insegnanti e studenti di italiano L2/LS.

Oltre alla consultazione del corpus, anche la sua stessa costruzione è stata utilizzata con finalità di formazione degli studenti di Didattica delle Lingue Moderne e Didattica della Lingua Italiana come LS dell'Università di Torino: collaborando alla trascrizione dei testi, alla correzione manuale e all'attribuzione di PoS si sono imbattuti negli errori più frequenti e sono stati spinti a riflettere sulle cause di quegli errori.

Inoltre, è stata proposta la consultazione del corpus VALICO da parte di studenti che svolgono la formazione in didattica dell'italiano con l'obiettivo di individuare nei testi che compongono il corpus gli errori commessi nella produzione scritta in italiano e, in una seconda fase, basare su queste osservazioni la progettazione di alcuni test linguistici a scelta multipla con distrattori basati sugli errori più frequenti.

Infine, vengono fatti alcuni brevi cenni alla produzione di attività didattiche basate sui testi del corpus di italofoeni VINCA: la sperimentazione da parte di apprendenti sia italiani che stranieri di un'attività di riordino delle sequenze dei testi, finalizzata all'elaborazione di un'attività di verifica per apprendenti romeni (Merlo 2009) e delle attività di comprensione testuale destinate a studenti italofoeni nativi e statunitensi (Corino, Marellò 2009).

⁶ <http://www.valico.org>

⁷ <http://www.valico.org/vinca.html>

Questo aspetto relativo alla produzione di attività didattiche basate su testi tratti da corpora risulta, invece, maggiormente sviluppato dalle ricerche di Lyding *et alii* (2013) e Ivanovska-Naskova (2018), a cui verrà dato ora particolare rilievo, date le analogie con la presente ricerca: in entrambi i casi sono stati progettati dei materiali didattici per l'insegnamento dell'italiano a stranieri basate sull'uso del corpus PAISÀ.

Lo studio di Lyding *et alii* (2013) dopo aver illustrato le principali caratteristiche e funzioni del corpus PAISÀ, presenta un'unità di lavoro dedicata ad apprendenti adolescenti di lingua italiana, la cui competenza linguistica si colloca tra i livelli A2-B1, incentrata sull'uso diretto da parte degli apprendenti dell'interfaccia di PAISÀ e delle sue funzionalità. I contenuti tematici dell'unità includono il saper parlare e scrivere delle proprie preferenze musicali, mentre dal punto di vista grammaticale l'unità è incentrata sull'uso dei comparativi e superlativi e su alcuni tratti di fonetica.

Dopo aver dato delle nozioni generali su cosa sia un corpus, viene chiesto agli apprendenti di ricercare con la modalità semplice e avanzata dell'interfaccia le diverse forme del comparativo e del superlativo. Successivamente viene proposta la lettura di un testo tratto dal corpus e viene chiesto agli apprendenti di ricercare nel corpus delle occorrenze di alcuni tratti tipici della scrittura di registro informale sul web presenti nel testo (ad esempio, acronimi, abbreviazioni, troncamenti ecc.) Infine, si richiede di creare un sottocorpus con le occorrenze reperite e di riscrivere le frasi trovate con un registro più formale. L'unità propone anche un'attività sul lessico: attraverso la funzione Word Cloud è possibile visualizzare quali parole co-occorrono con la parola ricercata sulla base della loro frequenza e, partendo dai risultati ottenuti, si può richiedere agli studenti di raggruppare le parole sulla base di relazioni semantiche.

Lo studio di Ivanovska-Naskova (2018), invece, è incentrato sull'uso dapprima guidato e infine diretto di corpora (tra cui PAISÀ) applicato allo studio dei connettivi concessivi italiani nell'insegnamento dell'italiano a stranieri. Viene presentata una proposta didattica sperimentata nell'ambito del corso di Sintassi dell'italiano, destinato a studenti di livello B2 nel corso di laurea quadriennale in Lingua e Letteratura italiana presso l'Università di Skopje. Tra gli obiettivi delle attività proposte vi è quello di abituare gli studenti all'uso incrociato di più risorse elettroniche, ossia corpora monolingui, corpora paralleli e dizionari elettronici; nello specifico le risorse utilizzate sono: il corpus PAISÀ (cfr. capitolo 4), un corpus parallelo italiano-macedone e, per quanto riguarda i dizionari, il Nuovo De Mauro e il Sabatini Coletti.

Le attività proposte, di cui di seguito si darà una breve descrizione, prevedono una prima fase in cui si richiede agli studenti di riflettere su alcuni testi tratti dal corpus PAISÀ (che è la parte dello studio che presenta analogie per la presente ricerca, poiché il corpus viene utilizzato come fonte di testi su cui costruire le attività didattiche) e una seconda fase che prevede l'uso diretto del corpus PAISÀ da parte degli studenti.

La parte introduttiva è mirata a far familiarizzare gli studenti con la costruzione della subordinata concessiva, le possibili traduzioni in macedone e il legame esistente tra la relazione concessiva e quella avversativa. Nelle attività iniziali si richiede di individuare le strutture concessive in alcuni brevi testi tratti dal corpus PAISÀ e, in particolare, di focalizzare l'attenzione sulla varietà di connettivi utilizzati per esprimere la concessione (congiunzioni e locuzioni concessive), sulla posizione della frase concessiva nella sintassi del periodo e sul tipo di frase concessiva (concessiva verbale – esplicita o implicita- o nominale).

In una seconda fase vengono proposte attività che prevedono l'uso diretto di PAISÀ da parte degli studenti: data una lista di connettivi concessivi, si richiede di ordinarli in base alla frequenza d'uso nel corpus; in un'altra attività viene chiesto di ricercare nel corpus degli esempi contenenti alcuni connettivi, in modo da verificare le intuizioni circa l'uso di questi connettivi.

Successivamente viene proposto un uso combinato di dizionari monolingui e del corpus PAISÀ: in alcune attività in una prima fase si richiede di consultare il dizionario De Mauro per individuare le funzioni principali dei connettivi *per quanto* e *malgrado* e, successivamente, si richiede di ritrovare tali funzioni in alcune frasi tratte da PAISÀ che contengono i connettivi sopra menzionati. La stessa tipologia di attività viene svolta con *anche se*, consultando però il dizionario Sabatini Coletti. In quest'ultimo caso, oltre a fornire degli esempi tratti da PAISÀ, viene chiesto agli studenti di tradurne alcuni, allo scopo di stimolare la riflessione su analogie e differenze tra il sistema dei connettivi italiani e quelli macedoni.

Nella parte conclusiva viene proposta un'attività basata interamente sull'uso diretto del corpus PAISÀ da parte degli studenti, a cui si richiede di ricercare del costrutto *pure + gerundio*.

Infine, vengono forniti degli esempi di frasi concessive italiane con le corrispettive traduzioni in macedone tratte da un corpus parallelo (Ivanovska-Naskova 2014) e si richiede di individuare le strutture concessive nelle due lingue e riflettere sulla loro forma.

Come è evidente, la maggior parte degli studi precedentemente citati riguardano principalmente l'estrazione di elementi linguistici dai corpora al fine di facilitare l'approccio da parte degli apprendenti con il lessico e la fraseologia, gettare le basi per la compilazione di dizionari e grammatiche o per fornire un supporto nell'apprendimento delle lingue. Solo in alcuni degli studi precedentemente citati, che includono anche l'insegnamento dell'italiano a stranieri, i corpora sono utilizzati come fonte per la produzione di materiali didattici o come strumenti tecnologici a supporto della didattica. Ciò che, però, emerge in generale è la necessità di far riferimento ai contenuti dei corpora al fine di sensibilizzare gli apprendenti circa alcune strutture linguistiche e fornire loro strumenti metodologici e/o tecnologici per l'analisi dell'italiano, oltre che la necessità di ottenere descrizioni grammaticali più accurate rispetto a quelle generalmente fornite dalle grammatiche e dai manuali, poiché gli esempi sono tratti da produzioni autentiche della lingua e non da esempi creati *ad hoc*.

A questo proposito, Wible (2008:179) mette in evidenza due modalità di utilizzo del web a supporto della didattica delle lingue, il secondo dei quali è quello che coincide con uno degli obiettivi della presente ricerca, ossia il reperimento di testi autentici da reimpiegare nella progettazione di attività didattiche:

to language learning, there are roughly two ways of constructing the Web as a language learning environment. First, the Web can be used as a location for storing and publishing content designed explicitly for language learning. Content of this sort would include online language lessons or activities. [...] A second way of viewing the Web as a source for language learning is to see it as a massive archive of pre-existing content that has not been designed for the purpose of language learning and teaching but which learners and teachers can exploit for that purpose.

Il riferimento a testi preesistenti che non sono stati progettati per gli scopi dell'apprendimento linguistico rinvia alla questione dell'autenticità dei testi, che è uno dei temi teorici e metodologici fondamentali nell'uso dei corpora con finalità didattiche, di cui si tratterà nella sezione successiva.

3.8. L'uso di materiali autentici

Nella definizione di materiale autentico, Bonvino (2004: 26); afferma che “un materiale può considerarsi autentico se si tratta di un testo che non è stato concepito per studenti

stranieri, ma al contrario se nasce per un pubblico di parlanti nativi”. Si tratta, dunque, della pluralità di materiali prodotti dalla comunità dei parlanti della L1, “creati per i madrelingua, quindi non a scopo didattico” (Begotti 2006: 14), e classificabili, sempre secondo Begotti, come cartacei, audio, video e multimediali (ivi: 15-18). L’uso di materiali autentici nella didattica delle lingue seconde e straniere è oggi ampiamente diffuso e consolidato, “in quanto testimonianze vive e spontanee della cultura e della civiltà [...] e delle relative manifestazioni” (Comodi 1995: 10). Si vedano a questo proposito gli studi di Breen (1985), Taylor (1994), Widdowson (2000), Stubbs (2001) Amor (2002) e, nell’ambito dell’italianistica, Harrich (2011), Santipolo (2014) e i già citati Comodi (1995) Begotti (2006 e 2011). L’autenticità dei materiali fu, inoltre, anche il tema principale della *keynote session* di Sinclair e Widdowson al TaLC 2002 “corpora and language teaching tomorrow”⁸, sulla base di due articoli di Sinclair (1991) e Widdowson (1991) che avevano sollevato l’interesse sul tema.

Il concetto di materiale autentico, in contrapposizione a quello di materiale didattico, fu introdotto da Wilkins (1976: 79), sebbene la spinta ad utilizzare esempi linguistici autentici abbia iniziato a farsi strada sin dagli anni ‘50, con lo sviluppo nella glottodidattica dell’approccio comunicativo. Quest’ultimo ha posto l’accento in tutte le sue declinazioni teoriche (metodo situazionale, nozionale-funzionale, metodo naturale di Krashen ecc.) sull’esigenza di esporre gli apprendenti della lingua a situazioni comunicative reali e, dunque, ad un contatto diretto con la lingua in uso.

Sempre a partire dagli anni ‘50 nell’ambito della linguistica Firth si espresse negativamente circa i manuali di lingua, sostenendo che molti degli esempi linguistici impiegati nella didattica fossero senza senso (“just nonsense”) dal punto di vista semantico (1957: 24) Questa opinione fu ripresa in seguito da Sinclair (1991: 5), che fa riferimento ad un presupposto, ossia “the absurd notion that invented examples can actually represent the language better than real ones”, a cui fanno eco Glisan e Drescher (1993: 24): “if grammar is to be taught for communicative purposes, the structures presented should reflect their use in current-day native speaker discourse”.

A questo proposito, Römer (2004) si interroga su quale varietà linguistica venga proposta agli apprendenti di una lingua tramite i manuali, proponendo di confrontare il cosiddetto “school English”, ossia la varietà utilizzata nei manuali, con l’inglese della vita quotidiana (“real English”):

⁸ Si fa riferimento ai discorsi (non pubblicati) tenuti da John Sinclair e Henry Widdowson nell’ambito del TaLC 2002.

to help us answer two crucial questions related to language teaching: “Do we teach our pupils authentic English, i.e., do we confront them with the same type of English we are likely to be confronted with in natural communicative situations?” and “What can we do to improve EFL teaching materials?” (Römer 2004:152)

A conclusione di tale riflessione, Römer si schiera a favore dell’uso di frasi tratte da situazioni comunicative reali in contrapposizione all’uso di testi creati ad hoc: “I would like to claim that there is a need for more authentic, naturally produced, non-invented examples in EFL teaching” (ivi:155), illustrando come la didattica delle lingue possa trarre beneficio dall’uso di testi autentici.

Innanzitutto, uno dei vantaggi degli esempi reali è che essi si sono verificati davvero, coerentemente all’affermazione di Sinclair sulla necessità della didattica delle lingue di presentare solo esempi reali (“present real examples only”), poiché la lingua non può essere inventata, ma solo catturata: “language cannot be invented; it can only be captured” (1997: 31).

Inoltre, sempre secondo Römer (2004: 154) l’uso dei testi autentici favorirebbe la naturalezza dell’espressione linguistica degli apprendenti: “Another reason why we might want to replace invented examples with real ones is that we may “hinder the development of fluency by excluding data samples that fluent native speakers actually say” (de Beaugrande 2001: 39)”. Tale affermazione trova conferma in Mauranen (2004: 208): i corpora composti da testi autentici (nello studio in questione, i corpora della lingua orale) “can achieve high authenticity, serve as communication aids, and provide irreplaceable models of the target language”.

Pertanto, sviluppare materiali didattici basati sulla lingua autentica significa avvicinare i tratti linguistici insegnati a quelli che si osservano nell’uso della lingua (Römer 2004: 161) ed è proprio questo aspetto a motivare sempre più i docenti ad inserire l’uso dei corpora nella didattica, come affermato da Aston, Bernardini e Stewart (2004: 12):

Authenticity [...] refers to a piece of text being “attested”, having occurred as part of genuine communicative (spoken or written) interactions. According to Hoey (this volume), exposure to authentic data is crucial since “only authentic data can preserve the collocations, the colligations, semantic associations of the language (p. 37). Indeed, it is this belief that motivates more and more teachers to introduce corpora into their classrooms

Sulla base di queste riflessioni, si ritiene che i corpora possano essere utili ai fini di questa ricerca come fonti per il reperimento di testi autentici che contengano i fenomeni dell'italiano dell'uso medio e che fungano da base per la progettazione di attività didattiche, come verrà illustrato successivamente nei capitoli 5 e 6.

CAPITOLO IV

OBIETTIVI E METODI DELLA RICERCA

Dopo aver descritto nel capitolo precedente alcuni concetti base della linguistica dei corpora e dopo aver illustrato alcuni ambiti di applicazione dei corpora alla didattica delle lingue, tra cui la realizzazione di materiali didattici, in questo capitolo verranno illustrati gli obiettivi e i metodi della ricerca. Dopo aver esposto le domande alla base di questa ricerca, verranno presentati i vari corpora disponibili dell'italiano contemporaneo scritto e i criteri di selezione della risorsa da utilizzare. Successivamente si illustrerà il metodo adottato per l'interrogazione del corpus tramite la progettazione di alcune stringhe di ricerca, di cui verrà data una descrizione. Infine, verrà presentato il metodo adottato per la progettazione e sperimentazione delle unità didattiche e per la realizzazione di un questionario per la raccolta delle opinioni degli studenti sulle attività svolte, di cui si descriveranno la struttura, il contenuto e le caratteristiche principali della piattaforma online su cui è stato collocato.

4.1. Domande di ricerca

Come illustrato nei capitoli precedenti, la diffusione della lingua italiana su tutto il territorio nazionale, fino a diventare la lingua ufficiale dello stato, è frutto di un lungo processo iniziato gradualmente con l'Unità d'Italia e che si è intensificato nella seconda metà del Novecento, soprattutto tramite l'istruzione obbligatoria e i mezzi di comunicazione di massa. Dovendo adattarsi alle esigenze pratiche della comunicazione nella vita quotidiana, la varietà standard della lingua, fino ad allora impiegata unicamente nella scrittura di registro formale, sotto la pressione della lingua parlata subì una graduale trasformazione che fece emergere tratti linguistici nuovi o già esistenti da secoli ma sanzionati dalla grammatica tradizionale, in quanto ritenuti sub-standard.

Sebbene tali tratti linguistici siano ampiamente utilizzati dalla comunità dei parlanti (Sabatini 1990, 2016), la loro trattazione nei manuali per l'insegnamento della lingua italiana a stranieri è molto limitata. Uno studio condotto da Słapek (2017) evidenzia, infatti, che su 32 pubblicazioni delle maggiori case editrici operanti nel settore dell'insegnamento della lingua italiana a stranieri, quali AlmaEdizioni, CDL Edizioni, Edilingua, Eli, Guerra Edizioni, Hoepli, Le Monnier e Loescher, solo 2 pubblicazioni

destinate ai livelli avanzati affrontano alcuni tratti dell'italiano dell'uso medio: “Nuovo Magari C1- C2” della casa editrice Alma Edizioni, che tratta la frase scissa nel volume dedicato al livello C1, e “Affresco italiano C2” della casa editrice Le Monnier, che affronta la dislocazione a sinistra e a destra, la frase scissa, l'uso del *c'è* presentativo, l'uso di *lui*, *lei* e *loro* in funzione di soggetto, l'uso di *gli* con i valori di *a loro* e *a lei*, l'uso di *cosa* come pronome interrogativo in concorrenza con *che* e *che cosa* (dati tratti da Słapek 2017).

Poiché “per il docente di italiano L2 il problema fondamentale è quello di evitare l'eccessivo scollamento tra lingua utilizzata in aula e quella effettivamente presente nelle situazioni comunicative quotidiane.” (Diadori, Palermo, Troncarelli 2015: 227) e data la presenza molto limitata dei tratti che caratterizzano il neostandard nella manualistica attualmente disponibile, la presente ricerca ha come punto di partenza il divario tra l'uso linguistico effettivo e l'uso prescritto dalle grammatiche dei tratti linguistici precedentemente nominati. Pertanto le domande della ricerca sono le seguenti:

1. Quali sono i corpora rappresentativi della lingua italiana scritta di media formalità degli ultimi 20 anni in cui è possibile verificare la diffusione dei tratti linguistici identificati da Sabatini (1990)?
2. Sulla base dell'interrogazione corpus prescelto, quali sono i tratti linguistici più frequenti e, dunque, più diffusi nei testi autentici rappresentativi della varietà scritta di media formalità?
3. A fronte all'attuale presenza limitata, se non assenza, dei tratti dell'italiano neostandard nella manualistica L2/LS (cfr. Słapek 2017), su quali dei tratti linguistici ricercati del corpus è rilevante progettare dei materiali per la didattica dell'italiano L2/LS?
4. Quali sono le opinioni dei partecipanti alla sperimentazione rispetto all'approfondimento dell'uso neostandard dei tratti linguistici selezionati?

Come è stato illustrato nei punti 1 e 2, si ipotizza l'utilizzo di un corpus di lingua italiana scritta per reperire i tratti linguistici identificati da Sabatini (1990) come risorsa di verifica dell'ipotesi circa l'effettiva reperibilità – e in quale misura – dei tratti dell'italiano dell'uso medio nei testi scritti di media formalità. A tale proposito nella sezione 4.2 verranno presentate alcune risorse prese in esame per tale scopo.

4.2. Scelta del corpus

In questa sezione verrà tracciata una panoramica delle risorse potenzialmente utili per documentare empiricamente la diffusione dei tratti linguistici identificati da Sabatini (1990) nei testi rappresentativi della varietà scritta contemporanea di media formalità.

Nello specifico, verranno presi in esame vari corpora di lingua italiana attualmente disponibili, al fine di selezionare le risorse più adeguate allo studio dei tratti del neostandard. È bene precisare che tale panoramica non si prefigge l'obiettivo di coprire la totalità dei corpora di lingua italiana esistenti, bensì di presentare una selezione di alcune risorse ritenute pertinenti rispetto agli obiettivi generali della presente ricerca, in linea con quanto affermato da McEnery e Hardie (2012: 2): “the corpus data we select to explore a research question must be well matched to that research question.”

Pertanto, ci si concentrerà esclusivamente sui corpora sincronici dell'italiano scritto, ossia sulle raccolte di testi scritti prodotti all'interno della stessa finestra temporale e che, pertanto, consentano la descrizione di uno stadio ben preciso della lingua: nel caso specifico, della varietà scritta dell'italiano ricorrente nell'uso comune dei parlanti nativi negli ultimi vent'anni. Di conseguenza, saranno esclusi dalla descrizione i corpora diacronici, utili ad uno studio dell'evoluzione dei fenomeni linguistici nel corso di un arco di tempo più ampio, i corpora dell'italiano parlato e i *learner corpora* dell'italiano L2/LS.

4.3. Corpora disponibili dell'italiano contemporaneo scritto

Dopo aver illustrato gli obiettivi e le domande di ricerca, in questa sezione verrà svolta una panoramica delle risorse attualmente disponibili per lo studio della varietà scritta dell'italiano neostandard.

Nella descrizione dei corpora e nella loro valutazione ai fini della presente ricerca verranno presi in considerazione i seguenti parametri: 1) le date di pubblicazione dei testi presenti nei corpora, privilegiando i testi prodotti negli anni Duemila fino ai nostri giorni, allo scopo di ottenere materiali rappresentativi dell'uso attuale della lingua¹; 2) gli eventuali aggiornamenti dei contenuti e la cadenza con cui vengono svolti; 3) le

¹ Ove disponibili, sono state riportate le informazioni relative alle date di *pubblicazione* dei testi. In caso contrario, sono state riportate solo le informazioni relative al periodo in cui è stata svolta la *raccolta* dei testi, che consentono perlomeno di fissare con certezza un limite oltre il quale non sono stati raccolti (e dunque prodotti) i testi che sono entrati a far parte del corpus.

dimensioni generali della risorsa²; 4) i generi testuali a cui appartengono i testi; 5) i criteri di accessibilità della risorsa (liberamente accessibile online o sottoposta a restrizioni di registrazione ed eventuale abbonamento). Un ulteriore punto a favore verrà considerato il rilascio del corpus con licenze *Copyleft*, in vista di un possibile reimpiego nella prassi didattica dei risultati delle ricerche.

Le informazioni sui corpora sono state raccolte tra gennaio e maggio del 2018 e sono state nuovamente controllate nel dicembre del 2020. Eventuali differenze rispetto a quanto di seguito illustrato sono attribuibili a successivi aggiornamenti e modifiche dei corpora o dei siti che li ospitano.

Sulla base delle informazioni raccolte verrà innanzi tutto fornito un elenco alfabetico dei corpora sincronici dell'italiano scritto attualmente disponibili e per ciascuno di essi verranno messe in luce le caratteristiche che li rendono adeguati o meno agli scopi di questa ricerca, fino a giungere alla selezione del corpus ritenuto più idoneo. Quest'ultimo verrà descritto in un'apposita sezione esterna all'elenco alfabetico, in cui verranno motivate le ragioni per le quali possa essere considerato preferibile rispetto alle altre risorse citate. Infine, verranno fornite tre tabelle riassuntive dei dati raccolti rispetto ai parametri sopra indicati.

a) Araneum Italicum

Aranea è un insieme di corpora paragonabili realizzati da Vladimír Benko (Benko 2014a, 2014b, 2016) nell'ambito di un progetto che ha coinvolto la cattedra UNESCO di *Translation Studies* presso l'Università Comenio di Bratislava e l'Istituto di Linguistica *L.*

² A questo proposito Sinclair (1996: 6): “The default value of Quantity is large. A corpus is assumed to contain a large number of words. The whole point of assembling a corpus is to gather data in quantity.” Affermazioni concordi sulla necessità di corpora di grandi dimensioni per gli studi sul lessico sono reperibili, ad esempio, anche in Svartvik (1992); Marengo (1996); Bolwler, Pearson (2002), Mukherjee (2002), McEnery (2003) e Baker, Haride, McEnery (2006). Per quanto riguarda, invece, le dimensioni minime che un corpus deve avere affinché sia adeguato per un'analisi statistica “Un criterio empirico suggerito dagli analisti (Bolasco, 1999, p. 203) è di osservare la *type/token ratio*: quando le parole distinte superano il 20% delle occorrenze il corpus non si può considerare sufficientemente esteso per un'analisi quantitativa. [...] Come indicazione generale, un corpus di 15.000 occorrenze si può considerare di “piccola dimensione”; un corpus di 50.000-100.000 occorrenze di “media dimensione” e un corpus maggiore di 200.000 occorrenze di “grande dimensione”. Un corpus sufficientemente grande (oltre le 500.000 occorrenze) può costituire una base per la costruzione di un lessico di frequenza rappresentativo di un linguaggio purché i testi siano abbastanza rappresentativi della sua eterogeneità.” (Giuliano, La Rocca 2008: 208). Sinclair (2005) sostiene in proposito che “The minimum size of a corpus depends on two main factors: 1) the kind of query that is anticipated from users; 2) the methodology they use to study the data. There is no maximum size.” Il che lascia intendere che “the sustainability of the sample depends on the specific study that is undertaken, and that there is no such thing as the best, optimum, sample size as such” (de Haan, 1992: 3).

Štúr presso l'Accademia Slovaca delle Scienze. I diversi corpora prodotti coprono varie lingue, incluso l'italiano, parlate e/o insegnate in Slovacchia e nei paesi confinanti e sono caratterizzati da denominazioni latine che identificano sia la lingua di riferimento (es. Anglicum, Finnicum, Italicum ecc.), sia le dimensioni della risorsa, poiché per ciascuna lingua esistono due versioni del corpus di dimensioni diverse, denominate rispettivamente Maius e Minus. Dunque, nel caso specifico della lingua italiana, i corpora sono denominati Araneum Italicum Maius e Minus.

Dal 2014 gli Aranea sono consultabili online alle condizioni valide al momento della scrittura di questo capitolo tramite i seguenti programmi di interrogazione di corpora: NoSketch Engine (una versione open source del software commerciale SketchEngine); il portale dedicato ai Web Corpora e alla Corpus Linguistics dell'Università Comenio di Bratislava³ (con accesso ospite che permette di consultare una versione ridotta dei corpora pari ad un massimo di 100 milioni di parole e un accesso completo previa registrazione al sito); dal sito KonText⁴ (con libero accesso a Italicum Maius e Minus, sebbene sia necessaria la registrazione al sito per accedere a tutte le funzioni della piattaforma) e, infine, Sketch Engine⁵, che concede l'accesso al solo Italicum Maius (in modalità gratuita per soli 30 giorni o, in alternativa, attraverso varie opzioni di abbonamento).

I testi sono stati raccolti tramite SpiderLing, un programma specializzato nella raccolta di testi dal web finalizzata alla formazione di corpora. Sono stati annotati per categoria grammaticale (*POS-tagging*, ossia *part-of-speech tagging*⁶) tramite TreeTagger, un software open source che permette di analizzare un testo ed ottenerne l'analisi grammaticale, utilizzando l'Araneum Universal Tagset, e lemmatizzati.

I corpora Araneum Italicum Maius e Minus sono corpora sincronici realizzati con testi raccolti tra l'ottobre e il dicembre 2014; sono costituiti rispettivamente da 1.2 miliardi di *token*⁷ e 119 milioni di *token*, che rappresentano le dimensioni standard delle versioni Maius e Minus di tutti i corpora Aranea:

³ ucts.uniba.sk

⁴ kontext.korpus.cz

⁵ www.sketchengine.co.uk

⁶ Associazione di specifici attributi informativi, in questo caso di carattere morfosintattico, agli elementi che compongono un testo elettronico (Cfr. Barbera 2015: 120).

⁷ "A token is any instance of a particular wordform in a text; comparing the number of tokens in the text to the number of types of tokens – where each type is a particular, unique wordform- can tell us how large a range of vocabulary is used in the text" (McEnery e Hardie 2012: 50)

For on-line use all corpora are sampled to receive the uniform 1.2-billion-token Maius size (i.e., approximately 1 billion of words), as well as the 120-million-token Minus size (for educational purposes). (Benko 2016: 4247)

I corpora possono essere scaricati per scopi didattici e di ricerca. Dal punto di vista delle dimensioni delle risorse e dell'attualità dei contenuti, i corpora Aranea di lingua italiana appaiono potenzialmente utili agli scopi di questa ricerca, tuttavia sono presenti alcune criticità.

Innanzitutto, non essendo state reperite informazioni circa la catalogazione in base a generi testuali o aree tematiche dei testi contenuti nei corpora, non è possibile valutare se tali materiali siano adeguati allo studio dei tratti del neostandard nell'italiano scritto di media formalità.

Inoltre, non sono presenti informazioni circa l'eventuale rilascio di licenze d'uso dei testi raccolti:

The source versions of the corpus data are available for download (for research and educational purposes). Note, however, that the copyright status of the data is not clear and users from countries where this might cause legal problems will have to solve this issue themselves. (Benko, 2016: 4248)

Questo aspetto costituirebbe un ostacolo nell'eventuale reimpiego con finalità didattiche o di ricerca dei testi contenuti nei corpora. Per tali ragioni le risorse Aranea, pur presentando alcuni punti di forza, sono state valutate come non pienamente adeguate agli scopi della presente ricerca.

b) Athenaeum Corpus

Athenaeum Corpus (Barbera 2007) è una risorsa prodotta dal gruppo di ricerca di Manuel Barbera e Carla Marengo dell'Università di Torino. È un corpus realizzato con testi prodotti dall'Università di Torino, tratti principalmente dal magazine universitario *L'Ateneo*, dalla newsletter *Dall'Università* e da circolari e e-mail amministrative, catalogati in base a generi testuali e aree tematiche. La risorsa mira a documentare la produzione scritta di un ateneo italiano nell'intera gamma delle sue attività, dagli ambiti di ricerca a quelli amministrativi.

I testi sono annotati rispetto alle parti del discorso, lemmatizzati e classificati per argomento e genere testuale. Il corpus consta di 306 mila *token* e dispone della licenza *Creative Commons Share Alike*⁸. È liberamente accessibile online⁹, attraverso una maschera di ricerca attiva dal 2004 e analoga a quella dei corpora NUNC (Cfr. sezione I di questo elenco), tuttavia non sono state reperite informazioni circa la conclusione del progetto, che si ipotizza sia ancora in corso, né sulle date di produzione dei testi in esso contenuti. Data la varietà di lingua che si propone di osservare (scritto accademico) e le ridotte dimensioni, questa risorsa non risulta adeguata agli scopi di questa ricerca.

c) CLIPS

CLIPS (Corpora e Lessici dell'Italiano Parlato e Scritto) è il secondo degli otto progetti del Cluster C18 “Linguistica computazionale: ricerche monolingui e multilingui” (Legge 488), finanziato dal Ministero dell'Istruzione, dell'Università e della Ricerca (MIUR) dal 1999 al 2004, avente l'obiettivo di realizzare strumenti per lo studio e per il trattamento automatico dell'italiano scritto e parlato. È consultabile online¹⁰ previa registrazione gratuita. Nella pagina dedicata al progetto CLIPS¹¹ sono reperibili informazioni sul solo corpus di italiano parlato, mentre per la sezione Lessici e Scritto dalla pagina del progetto si viene reindirizzati ad una pagina vuota dell'Istituto di Linguistica Computazionale di Pisa, il che non consente di valutare l'adeguatezza o meno della risorsa rispetto agli scopi della ricerca.

d) COLFIS

COLFIS, Corpus e Lessico di Frequenza dell'Italiano Scritto (Laudanna *et al.* 1995), è stato realizzato da un gruppo di lavoro coordinato da Pier Marco Bertinetto (Scuola Normale Superiore, Pisa) che ha coinvolto, oltre alla Scuola Normale, il CNR e le Università di Salerno, Genova e L'Aquila. È un corpus bilanciato di oltre 3 milioni di parole, i cui contenuti rispecchiano le tendenze di lettura degli italiani ricavate da un'indagine ISTAT del 1993, allo scopo di

costruire una banca di dati che riproducesse nel modo più fedele possibile la frequenza reale delle parole nell'esperienza del parlante, e che dunque si avvicinasse il più possibile al lessico

⁸ Cfr. <https://iris.unito.it/handle/2318/120913#.WqaOuExFx1s>

⁹ www.corpora.unito.it/ateneo/ateneo_.php

¹⁰ www.clips.unina.it/it/corpus.jsp

¹¹ www.clips.unina.it/it

mentale di un parlante di media cultura, piuttosto che ad un dizionario esaustivo della lingua italiana. (Bambini, Trevisan 2012: 2)

Include testi tratti dai quotidiani «La Repubblica», «La Stampa», «Il Corriere della Sera» del periodo 1992-1994¹², periodici e libri (inclusi anche quelli letti per motivi scolastici o professionali)¹³. È composto da 3,7 milioni di *token*, così suddivisi: quotidiani (1,8 milioni), periodici (1,3 milioni), libri (655 mila).

Il corpus è stato lemmatizzato e annotato rispetto alle parti del discorso, riconducendo ogni parola alla forma base (o entrata del dizionario) dapprima attraverso una lemmatizzazione automatica, utilizzando un lemmatizzatore IBM già impiegato per il corpus LIP¹⁴, e successivamente svolgendo alcuni interventi di correzione manuale.

Il lessico di frequenza ricavato dal corpus ha due sottoparti: il formario e il lemmario. Il formario riporta la frequenza di ciascuna forma presente nel corpus, senza distinguere fra i diversi lemmi a cui la forma si può ricondurre (ad esempio: contiene la frequenza di *porto*, ma senza distinguere tra la forma singolare del sostantivo *porto*, oppure la prima persona singolare dell'indicativo del verbo *portare*). Il lemmario riporta, invece, le frequenze delle forme lemmatizzate, cioè ricondotte al lemma di provenienza e contiene, inoltre, la frequenza delle parole sintagmatiche (dette anche *polirematiche*, ossia elementi lessicali formati da più di una parola, ad esempio *occhiali da sole*, *carta di credito*, ecc.).

La pagina dell'Istituto di Linguistica Computazionale (ILC) del CNR di Genova¹⁵ contiene informazioni sul progetto e un link¹⁶ da cui scaricare il corpus, che risultava liberamente accessibile online nel gennaio 2018. Tuttavia, ad un successivo controllo nel dicembre 2020 la pagina da cui scaricare il corpus non è più risultata raggiungibile. Lo stesso vale per il link di EsploraCoLFIS¹⁷ (Bambini, Trevisan 2012), un'interfaccia web sviluppata per facilitare la consultazione del corpus, basata sul motore di ricerca di testo

¹² <http://linguistica.sns.it/COLFIS/Descrizione.htm>

¹³ “I testi inclusi nella banca dati sono [...] quelli più letti nel periodo di tre anni 1992-1994” (Laudanna et al., 1995: 106-108), sulla base dei dati ISTAT 1993 e, nello specifico: Audipress 1993-1994 sulla lettura, Accertamento Diffusione Stampa 1992-1993 e FIEG 1990-1992 per periodici e quotidiani; classifiche dei libri più venduti dal 1992 al 1994 (Inserito *Tuttolibri*, *La Stampa*) e dati ISTAT sulla produzione editoriale del 1993 per i libri.

¹⁴ Lessico di frequenza dell'italiano parlato (LIP), curato da De Mauro, Mancini, Vedovelli e Voghera (1993), in collaborazione con la Fondazione IBM Italia, è una raccolta di trascrizioni di registrazioni svolte tra il 1990 e il 1992 a Milano, Firenze, Roma e Napoli, pari a quasi 57 ore di parlato. È costituito da 469 testi, contenenti complessivamente circa 490.000 parole. I lemmi sono consultabili in base all'ordine alfabetico e in base alla loro frequenza. Al volume si accompagnano due dischetti contenenti le trascrizioni di tutti i testi del corpus, consentendo anche una consultazione in formato audio.

¹⁵ www.istc.cnr.it/grouppage/colfis

¹⁶ www.ge.ilc.cnr.it/strumenti.php

¹⁷ <http://esploracolfis.sns.it/EsploraCoLFIS>

Apache Lucene, che permetteva di ricercare i dati contenuti nel corpus lemmatizzato, le statistiche del lemmario e del formario, e di accedere ad alcuni dei testi originali, poiché per alcuni libri non era stata concessa l'autorizzazione alla trascrizione nel corpus di alcune loro parti. Tali informazioni erano disponibili nel gennaio 2018 alla pagina dedicata ai libri inseriti nella costruzione del corpus¹⁸, tuttavia, anche quest'ultima è risultata non più raggiungibile nel dicembre 2020.

Dunque, in quei casi non è possibile accedere ai testi originali partendo dai risultati dell'interrogazione del corpus. Questi ultimi, però, sono comunque stati utilizzati per la realizzazione del lessico di frequenza e del relativo formario: ciò significa che esiste, quindi, una discrepanza tra i dati del lessico di frequenza e l'interrogazione del corpus lemmatizzato e non lemmatizzato. In tutti gli altri casi in cui le opere sono invece state autorizzate, è possibile scaricare l'intero testo originale.

Tra i punti di forza del CoLFIS vi sono sicuramente il bilanciamento delle fonti, la funzionalità della maschera di ricerca e la libera accessibilità online. Tuttavia, poiché i dati inseriti nel corpus risalgono esclusivamente agli anni '90, l'utilizzo di questo corpus condurrebbe a risultati non più attuali e inadeguati agli scopi della presente ricerca.

e) CorDIC

I corpora CorDIC (Corpora Didattici Italiani di Confronto) sono stati realizzati nell'ambito del LABLITA (Laboratorio Linguistico Italiano dell'Università di Firenze), che dal 1973 si occupa della realizzazione e gestione di corpora con lo standard di trascrizione CHAT (Progetto Childes cfr. Bortolini e Pizzuto, 1997 (a c. di), Mac Whinney, 1997), allo scopo di offrire uno strumento didattico per il confronto tra le varietà dell'italiano scritto e parlato.

CorDIC è suddiviso in due sezioni indipendenti: il corpus CorDIC-scritto e il corpus CorDIC-parlato, pubblicati in rete a cura di Lorenzo Gregori (Università di Firenze) e Alessandro Panunzi e liberamente consultabili online tramite il sito del LABLITA¹⁹, che, come riportato sulla pagina stessa, è un'applicazione web che utilizza la piattaforma open-source messa a disposizione dal progetto NoSketch Engine. Ciascun corpus contiene circa 500.000 parole distribuite su un totale di 200 testi (con una media di 2.500 parole per ogni testo) e sono, dunque, comparabili dal punto di vista delle dimensioni.

¹⁸ <http://linguistica.sns.it/CoLFIS/Testi/libri/libri.htm>

¹⁹ corporadidattici.lablita.it

Il corpus CorDIC-scritto contiene 200 testi classificati in 5 tipologie (*arti, burocratico, creativo, economia e stampa*), ciascuna delle quali è costituita all'incirca dallo stesso numero di parole (100.000) e dallo stesso numero di testi (40). I testi sono stati scaricati dalla rete tramite la piattaforma RIDIRE-CPI, realizzata per il progetto RIDIRE e poi selezionati manualmente con un lavoro di analisi e pulizia per assicurare la loro adeguatezza all'interno del corpus.

Nonostante il CorDIC-scritto sia ben bilanciato al suo interno e sia consultabile tramite NoSketch Engine, non è stato possibile valutarne appieno la conformità agli obiettivi del progetto, poiché riguardo ai contenuti, non sono state reperite informazioni né circa le date in cui sono stati prodotti, né circa il loro status del *copyright*.

f) CorIS/CoDIS

CorIS/CoDIS (Rossini Favretti, Tamburini, De Santis 2002) è il risultato di una ricerca iniziata nel 1998, coordinata da Rema Rossini Favretti e svolta presso l'Università di Bologna. Il CorIS (CORpus di Riferimento dell'Italiano Scritto), come indica il nome, è stato elaborato con l'obiettivo di realizzare un corpus generale di riferimento per l'analisi dell'italiano scritto.

Si tratta di una raccolta di testi in formato elettronico, selezionati come rappresentativi dell'italiano attuale, il cui nucleo originario era collocabile approssimativamente tra gli anni '80 e '90 del Novecento e classificabile, come affermato sulla stessa pagina di riferimento del progetto, come testi appartenenti all'italiano "scritto-scritto" sulla base della terminologia elaborata da Nencioni (1976), ossia testi scritti nei quali non vi è alcun intento di avvicinarsi allo stile del parlato.

È costituito da 150 milioni di *token* e viene aggiornato con cadenza biennale (la più recente, al momento della scrittura, nel 2017), inserendo nuovi materiali in un corpus di monitoraggio inglobato al suo interno (Rossini Favretti, Tamburini, De Santis 2002). A questo proposito, il CorIS è strutturato in sette sottocorpora: sei che rappresentano le diverse varietà dell'italiano scritto (stampa, narrativa, prosa accademica, prosa giuridico-amministrativa, miscellanea ed *ephemera*) e il già nominato corpus di monitoraggio. Per ogni risultato di ricerca viene fornita l'indicazione del sottocorpus di provenienza, ma non del testo di provenienza.

Dal CorIS è derivato il CORpus Dinamico dell'Italiano Scritto (CoDIS), così definito proprio perché può essere dinamicamente adattato alle diverse esigenze di ricerca. È

costituito dai sei sottocorpora delle varietà dell'italiano scritto presenti nel CorIS ma, a differenza di quest'ultimo, offre all'utente la possibilità non solo di selezionarne contemporaneamente più di uno per volta, ma anche di sceglierne le dimensioni funzionali agli scopi di ricerca fra le quattro dimensioni indicate. Sia il CorIS che il CoDIS sono liberamente consultabili online²⁰.

I corpora sono stati lemmatizzati e annotati rispetto alle parti del discorso attraverso il CORISTagger (Tamburini 2000), un programma di *tagging* appositamente elaborato per la lingua italiana nell'ambito del progetto CORIS.

Dal punto di vista del copyright dei testi, in Rossini Favretti, Tamburini, De Santis (2002) vengono messe in evidenza le difficoltà legate all'acquisizione dei diritti al riutilizzo dei testi raccolti attraverso complesse trattative con gli editori che, come affermato nella pubblicazione sopra menzionata, non sempre hanno dato i loro frutti, nonostante la garanzia di un uso non commerciale, funzionale agli scopi della ricerca accademica, e l'adozione di misure di sicurezza contro la pirateria. Non essendo reperibili al momento della scrittura ulteriori informazioni sull'esito di tali trattative, si può dedurre che alcuni editori non abbiano concesso i permessi e che, pertanto, è probabile che alcuni testi o siano stati esclusi dal corpus, o vi siano stati inseriti, ma privi delle licenze d'uso.

I punti di forza di questa risorsa sono sicuramente l'attualità dei materiali, sottoposti a costante aggiornamento biennale, e la libera accessibilità in rete. Tuttavia, tra i fattori che hanno determinato la sua esclusione vi sono: le dimensioni nettamente inferiori a quelle di altre risorse incluse in questa trattazione (130 milioni di *token*); il fatto che alcuni tra i subcorpora disponibili (stampa, narrativa, prosa accademica, prosa giuridico-amministrativa, miscellanea ed *ephemera*) si basino su generi testuali caratterizzati da un registro linguistico più elevato rispetto alla media formalità che caratterizza il neostandard e, infine, l'impossibilità di accertare con certezza lo status del copyright per la totalità dei testi contenuti, che rappresenta un fattore sfavorevole ad un loro possibile reimpiego con finalità didattiche. Inoltre, la presenza di testi che spaziano tra gli anni '80 del Novecento e il 2017 permette di ipotizzare la presenza di effetti diacronici sulla variazione lessicale, il che contrasta con l'obiettivo generale di uno studio sincronico della varietà linguistica più recente e maggiormente in uso.

²⁰ <http://corpora.dslo.unibo.it/TCORIS>

g) ItTenTen

ItTenTen è un corpus che fa parte della *TenTen Corpus Family* (Jakubíček *et al.*, 2013), un set di corpora elaborato dall'azienda Lexical Computing e dall'Università Masaryk (Repubblica Ceca), in cui rientrano oltre 30 corpora rappresentativi di altrettante lingue. La versione del 2016 del corpus italiano, itTenTen16, contiene 5,8 miliardi di *token*. Come per tutti i corpora del set TenTen, è costituito da testi scaricati dal web tra maggio e agosto 2016 tramite SpiderLing. Tutti i prodotti TenTen sono consultabili tramite Sketch Engine²¹ e la piattaforma ne garantisce l'accesso gratuito per 30 giorni o, in alternativa, attraverso varie opzioni di abbonamento.

Tutti i testi sono stati annotati per categoria grammaticale e lemmatizzati tramite Tree Tagger utilizzando i parametri di Baroni *et al.* (2009)²². I punti di forza di questa risorsa sono la grande quantità di dati contenuti (5,8 miliardi di *token*) e la loro attualità (2016). Tuttavia, come nel caso dei corpora Aranea, non sono state reperite informazioni riguardanti la catalogazione dei testi in base a generi testuali o aree tematiche di appartenenza, né circa l'eventuale rilascio di licenze d'uso dei testi raccolti. Tali lacune non consentono di valutare pienamente se, al di là della loro attualità, i contenuti del corpus possano essere adatti agli scopi di questa ricerca e se sia possibile un loro reimpiego con finalità didattiche o di ricerca nel rispetto dei diritti d'autore.

h) itWaC

itWaC, è un corpus da 2 miliardi di *token*, che rientra nel progetto WaCky (Web-as-Corpus Kool Ynitiative; cfr. Baroni, Bernardini 2006, Baroni et alii 2009) avente come obiettivo la realizzazione di grandi corpora di italiano, inglese, francese e tedesco a partire dal Web.

È stato realizzato con testi tratti dal web raccolti attraverso un procedimento definito *seeds crawling* (letteralmente: raccolta di semi) che consta di due fasi. Nella prima fase i *seed*, ossia i termini che si ritengono rappresentativi di un certo ambito²³, erano costituiti da parole chiave di uso comune combinate casualmente (ad esempio: cane + casa) e ciascuna combinazione è stata utilizzata come stringa di ricerca in *Yahoo!*. Nello specifico, le parole utilizzate come *seed* sono parole a media frequenza tratte in parte dal corpus "la Repubblica" (cfr. sezione *i*, a seguire) e in parte da una lista di parole appartenenti al

²¹ Cfr. "The corpora are all available for research at <http://www.sketchengine.co.uk>" (Jakubíček *et al.* 2013 :126-127) e www.sketchengine.eu/documentation/tenten-corpora

²² Cfr. www.sketchengine.co.uk/ittenten-italian-corpus/#toggle-id-1

²³ "Terms that are expected to be typical of the domain of interest" (Baroni e Bernardini 2004: 1313).

vocabolario italiano di base. La scelta di queste due fonti si basa sull'osservazione (Baroni, Ueyama 2006) che le ricerche tramite parole tratte da giornali hanno come risultato testi appartenenti alla sfera pubblica, mentre le ricerche tramite parole appartenenti al vocabolario di base conducono a blog e a testi di carattere più personale, il che permette di ottenere un corpus variegato al suo interno. Dalle ricerche effettuate in *Yahoo!* con le parole combinate casualmente è stata ottenuta la lista con le URL delle pagine contenenti quelle parole.

Successivamente, utilizzando BootCaT (Baroni, Bernardini 2004), un programma per la costruzione di corpora, le pagine ottenute per ciascuna ricerca sono state dapprima raccolte automaticamente tramite una procedura detta *crawling*, limitata al solo dominio .it., poi sono state formattate come testo e utilizzate per la costruzione di un primo corpus. La procedura è stata, poi, ripetuta: i testi ritenuti rilevanti nella prima fase di ricerca sono stati a loro volta reimpiegati come base per l'estrazione di nuove parole da utilizzare come *seed* in una nuova fase di ricerca su *Yahoo!*, proseguendo, così la procedura di *crawling*. In questo modo, analogamente a quanto già descritto, le nuove URL sono state formattate come testo ed è stato creato un nuovo corpus. La procedura di estrazione di *seed* e di creazione di un nuovo corpus può essere ripetuta infinite volte, fino a quando il corpus non raggiunge una certa dimensione desiderata.

Il corpus è stato POS-tagato tramite TreeTagger, mentre la lemmatizzazione delle singole parole è stata realizzata tramite la risorsa morfologica Morph-it! (Zanchetta, Baroni 2005).

È liberamente consultabile e scaricabile tramite la pagina del progetto²⁴, che mette a disposizione i dati relativi ai corpora, gli strumenti utilizzati nella loro costruzione (incluse le *seeds lists* e le URL) e altre risorse, e dalla pagina del centro di ricerca CoLiTec (Corpora, Linguistica, Tecnologia)²⁵. Per quanto riguarda lo status del copyright dei testi, poiché nella costruzione del corpus non è stata imposta la selezione dei soli testi aventi licenze di tipo *Copyleft*, alla luce delle grandi dimensioni del corpus non è possibile accertare lo status del copyright di ciascun testo contenuto, né di richiederne la licenza d'uso:

The copyright issue remains a thorny one: there is no easy way of determining whether the content of a particular page is copyrighted, nor is it feasible to ask millions of potential

²⁴ <http://wacky.sslmit.unibo.it/doku.php>

²⁵ <https://corpora.dipintra.it/>

copyright holders for usage permission. However, our crawler does respect the download policies imposed by website administrators (i.e. the robots.txt file), and the WaCky website contains information on how to request the removal of specific documents from our corpora. (Baroni *et al.* 2009: 224)

Pertanto, pur presentando il vantaggio delle grandi dimensioni, della varietà delle fonti e dell'accessibilità in rete, la risorsa non risulta pienamente adeguata agli scopi di ricerca sotto due aspetti fondamentali: essendo stata realizzata tra il 2005 e il 2007, i suoi contenuti risultano più datati rispetto a quelli di altre analoghe risorse citate in questa trattazione; inoltre, l'impossibilità di accertare lo status del copyright dei suoi contenuti non consentirebbe un possibile reimpiego con finalità didattiche dei materiali in essa contenuti.

i) la Repubblica

Il corpus “la Repubblica”²⁶ (Baroni *et al.* 2004) è stato realizzato presso la allora Scuola Superiore di Lingue Moderne per Interpreti e Traduttori (ora Dipartimento di Interpretazione e Traduzione) dell'Università di Bologna (Campus di Forlì) da Marco Baroni ed il suo gruppo di ricerca. È un corpus di italiano giornalistico, liberamente consultabile online dalla pagina già citata del centro di ricerca CoLiTec (Corpora, Linguistica, Tecnologia) tramite NoSketch Engine.

Il corpus contiene gli articoli delle annate dal 1985 al 2000 del quotidiano omonimo da cui trae il nome. I testi sono stati taggati per genere testuale (reportage e commento)²⁷ e per argomento (cultura, economia, istruzione, meteo, notizie, politica, religione, scienza, società e sport)²⁸. È stato POS-taggiato tramite TreeTagger, mentre la lemmatizzazione è stata realizzata tramite Morph-it!. È possibile effettuare ricerche per singole parole o locuzioni (*phrase*), POS, oppure restringere la ricerca a parti del corpus (es. titoli *vs* corpo degli articoli), selezionare sottocorpora sulla base di informazioni contenute in alcuni *tag* (ad esempio: selezionare tutti gli articoli di un certo autore, pubblicati in un certo mese ecc.). Presenta dimensioni notevoli, poiché contiene 326 milioni di *token* (Baroni *et al.* 2004), tuttavia presenta alcune criticità: i suoi contenuti provengono tutti da un'unica

²⁶ https://corpora.dipintra.it/public/run.cgi/first_form

²⁷ Nel corpus sono utilizzate le denominazioni in inglese: *news report* e *comment*.

²⁸ Come sopra: *culture, economics, education, weather, news, politics, church, science, society* e *sport*.

fonte, rientrano all'interno della sola prosa giornalistica, inoltre risultano ormai datati rispetto agli scopi di questa ricerca e, infine, non sono stati rilasciati con licenza *Copyleft*.

j) LexALP

LexALP²⁹ (Lyding *et al.* 2006), è il risultato di un partenariato che include l'Accademia Europea di Bolzano (EURAC research) come partner capofila e varie università, istituti di ricerca e istituzioni pubbliche. Il progetto, svoltosi dal gennaio 2005 al febbraio 2008, ha avuto come obiettivo un raffronto contrastivo mirato all'armonizzazione dei linguaggi giuridici utilizzati negli stati dell'arco alpino (Austria, Italia, Francia, Germania, Slovenia e Svizzera), nel diritto europeo e internazionale, con particolare riferimento alla Convenzione delle Alpi (1995). Ciò ha portato alla realizzazione di corpora monolingue paralleli che coprono il francese, l'italiano il tedesco e lo sloveno, basati su circa 3000 documenti legali degli otto sistemi giuridici sopra nominati, realizzati e gestiti tramite CWB (Corpus Work Bench), un database terminologico, un database bibliografico e ulteriori strumenti linguistici per la lingua francese, italiana, tedesca e slovena, liberamente consultabili gratuitamente online. Poiché i generi testuali in cui rientrano i materiali contenuti non sono adeguati agli scopi della presente ricerca, si è deciso di non adottare il corpus italiano del progetto LexALP.

k) LIF

Il Lessico di frequenza della lingua italiana contemporanea (LIF) è stato realizzato nel 1971 al CNUCE (Centro Nazionale Universitario di Calcolo Elettronico) di Pisa (cfr. Bortolini *et al.* 1971). È stato il primo grande progetto di realizzazione di un lessico di frequenza per la lingua italiana non incentrato in maniera specifica su un singolo autore o su testi letterari. I testi contenuti nel corpus si collocano tra il 1947 e il 1968, sono tratti da 5 fonti diverse (teatro, romanzi, cinema, periodici, sussidiari) con l'obiettivo di rappresentare diversi generi testuali, alcuni dei quali con varietà di lingua vicine al parlato. Il corpus contiene 500.000 parole, dalle quali sono stati estratti circa 5.000 lemmi ordinati per frequenza e in ordine alfabetico. A causa della datazione non più attuale dei suoi testi e del fatto che la risorsa non sia disponibile per la consultazione né offline né online, il LIF è stato valutato come inadatto agli scopi della presente ricerca.

²⁹ http://lexalp.eurac.edu/projects/corpus_en.htm
http://lexalpapps.eurac.edu:8080/htdocs2/lexalp/corp_lexalp/search_corp.php

D) NUNC

Il corpus NUNC (*Newsgroups UseNet Corpora*) in lingua italiana fa parte di un progetto dell'Università di Torino, coordinato da Manuel Barbera (cfr. Barbera, Marengo 2008), avviato nel 2003 e tutt'ora in corso, che ha portato alla realizzazione di corpora generici e specialistici in cinque lingue³⁰, tra cui l'italiano, tramite CWB (Corpus Workbench), un insieme di strumenti open source mirati alla gestione e all'interrogazione di corpora di grandi dimensioni. NUNC italiano è un web corpus generico, realizzato con pagine web di conversazioni scritte tratte da "newsgroups" di cinema, cucina, fotografia e motori presenti in rete negli anni 2002-2004:

Un newsgroup è un forum telematico a libero accesso, gratuito, disponibile su Internet, che si manifesta nella forma di testi scritti, ed il cui funzionamento è assai semplice: ogni utente scrive un messaggio, il post, e lo invia ad una specie di "bacheca elettronica" mantenuta presso una rete di server (i newserver che costituiscono UseNet), dai quali gli altri utenti del gruppo possono scaricarlo, leggerlo e rispondervi, costruendo anche articolate catene (thread) di botte e risposte. (Barbera 2007: 8)

Il corpus è suddiviso in due sezioni, costituite rispettivamente da 127 milioni e 109 milioni di *token*. È liberamente consultabile online³¹, dove è possibile impostare la ricerca per parola, lemma e POS, visualizzare il contesto scegliendo in un menu l'opzione desiderata da un minimo di 5 parole ad un massimo di 200 e di visualizzare dai 20 ai 1000 risultati. Per ciascuna delle aree tematiche sopra nominate è, inoltre, disponibile un corpus specialistico sul tema. Al momento della scrittura non sono state reperite informazioni che chiariscano se i testi contenuti nel corpus NUNC siano gli stessi utilizzati per costruire i corpora specialistici sui diversi temi.

Come tutti i corpora del progetto NUNC, l'interfaccia di ricerca utilizzata per il corpus generico italiano è il CQP (Corpus Query Processor), sviluppato dall'*Institut für Maschinelle Sprachverarbeitung* di Stoccarda. È dotato di un breve prontuario di istruzioni all'uso della maschera di ricerca che rende facoltativa la conoscenza del linguaggio di CQP.

³⁰ Corpus generici di lingua francese, inglese (varietà australiana), italiana e spagnola e corpus specialistici nelle stesse lingue, con l'aggiunta della lingua tedesca.

³¹ www.corpora.unito.it/index_nunc.php

I punti di forza di questa risorsa sono sicuramente l'attualità dei materiali, trattandosi di un progetto avviato nel 2003 e ancora in corso, le dimensioni, seppur inferiori a quelle di altre risorse già citate, e la libera accessibilità in rete. Tuttavia, la questione dello status del *copyright* è articolata come segue:

tutti i NUNC sono stati licenziati da bmanuel.org (che ne è il soggetto proprietario) con licenza Creative Commons Share Alike [...], il che ne garantisce la libera diffusione a chiunque (ad esempio a corpora.unito.it, che ne è il distributore principale) purché i risultati derivati dal lavoro su questi corpora siano rilasciati con la medesima licenza (copyleft): siano cioè resi pubblici alle stesse condizioni. (Barbera 2007 in Barbera 2013: 100)

I singoli testi utilizzati per costruire i corpora non sono stati rilasciati con licenza Copyleft e il loro utilizzo si basa esclusivamente sull'assunto della riutilizzabilità di materiali di pubblico dominio:

UseNet per definizione e tradizione è il regno del pubblico dominio, quindi [la disponibilità legale del materiale] sembrerebbe una ovvia assunzione; in realtà, se lo si dovesse sostenere legalmente, le cose potrebbero non essere così pacifiche (talvolta si è ricorso ad un cosiddetto "diritto implicito"), ma dato che il comune sentire sostiene comunque la nostra *bonam fidem*, e che non vi sono ad ogni buon conto interessi rilevanti lesi, è certo assai improbabile che contestazioni significative possano essere sollevate. (*ibidem*)

Riassumendo, i punti di forza di questa risorsa sono la libera accessibilità in rete, il fatto che sia sufficientemente grande da essere preso in esame come risorsa rappresentativa dell'italiano contemporaneo, e il fatto che i suoi testi, essendo tratti da conversazioni sul web, rispecchiano una varietà linguistica contemporanea e reale, variegata per registri e temi. A determinarne, tuttavia, l'esclusione dalla presente ricerca è l'inattualità dei materiali contenuti, che si collocano tra il 2002 e il 2004, nonché la complessa questione relativa alle licenze d'uso, ove le licenze *Copyleft* si applicano solo all'intera risorsa, ma non ai singoli testi contenuti.

m) Perugia Corpus

Il Perugia Corpus (PEC) (Spina 2014) è un corpus realizzato all'Università per Stranieri di Perugia tra il 2011 e il 2012, col l'obiettivo di costituire un riferimento per l'italiano

contemporaneo, nelle varietà scritte e parlate, ove lo scritto rappresenta l'85% del totale del PEC, ed il parlato il restante 15%. È costituito da oltre 26 milioni di parole, distribuite in 10 differenti sezioni, che corrispondono ad altrettanti generi testuali. Ogni genere è, poi, a sua volta suddiviso in una serie di sottogeneri, per un totale di 43 tipologie testuali.

Nello specifico, per l'italiano scritto, al quale si farà riferimento in questa descrizione, sono disponibili i seguenti generi, qui elencati per ordine decrescente di *token*: testi tratti dal web (oltre 7 milioni), stampa (oltre 5 milioni), letteratura (oltre 3 milioni), saggi (oltre 2 milioni) e seguono, poi, con oltre 1 milione di *token* lo scritto scolastico, amministrativo e accademico.

Per quanto riguarda i sottogeneri, le sezioni dell'italiano scritto sono così suddivise: letteratura: 60 romanzi pubblicati tra il 1990 e il 2012 da 45 autori diversi; saggistica: saggi pubblicati da autori italiani tra il 1990 e il 2012 e che rientrano in quattro aree tematiche (attualità, biografia, politica e tempo libero); stampa: articoli del 2011 tratti da «L'Espresso» e del 2012 tratti sempre da «L'Espresso», dal «Corriere della Sera» e da «Il Sole 24 ore».

Tali aree rientrano nelle seguenti tipologie, elencate per ordine decrescente di *token*: cronaca, politica, esteri, cultura, economia, editoriali, sport, spettacolo e lettere; scritto accademico: questa sezione coincide interamente, seppur con alcune integrazioni, con il Corpus di Italiano Accademico (Spina 2010b) che comprende quattro sottosezioni (tesi di laurea, dispense, manuali e articoli scientifici), a loro volta suddivise in tre macroaree tematiche (umanistica, giuridico-economica e scientifica); scritto scolastico: include 4.054 temi svolti tra il 2010 e il 2011 da studenti delle scuole secondarie inferiori e superiori, estratti automaticamente dal sito web di Repubblica Scuola³²; scritto amministrativo: è costituito da testi di leggi (europee, statali, regionali) per il 75% e da regolamenti e documenti amministrativi più brevi per il restante 25%; web: è la più ampia sezione del corpus, suddivisa in testi di interazione (tratti da blog, forum, chat e social network) e testi di riferimento (tratti da Wikipedia). Per i testi tratti dai blog sono stati selezionati solo i post, senza i commenti, da una cinquantina di blog di genere personale, giornalistico o aziendale. I testi di Wikipedia sono stati, invece, selezionati in modo casuale dalla versione italiana integrale nel gennaio 2012. Nella categoria dei social network sono inclusi 24.424 post tratti da profili di Facebook e di Twitter in base alle già nominate tre tipologie: personale, politico e aziendale.

³² <http://scuola.repubblica.it>

Come riportato da Spina (2014), nella composizione del corpus si è scelto di privilegiare la differenziazione dei generi testuali, includendo anche il parlato, a discapito delle dimensioni del corpus.

Il PEC può dunque essere considerato un corpus di riferimento “low cost”, di dimensioni contenute ma con una buona rappresentatività delle diverse varietà scritte e parlate dell’italiano. Le dimensioni contenute del PEC presentano inoltre due vantaggi: permettono di gestire, in fase di interrogazione, quantità di risultati più maneggevoli (Hundt e Leech 2012), e consentono di ottenere una buona accuratezza nell’annotazione. (Spina 2014: 354)

Il corpus è stato POS-taggiato tramite TreeTagger con un tagset creato *ad hoc*, ispirato a quello descritto in Baroni *et al.* (2004), ed è stato lemmatizzato. I testi che compongono il corpus sono, inoltre, stati etichettati a livello di genere testuale (Spina 2010a: 3203).

Da giugno 2015 il corpus è accessibile in rete dal sito dell’Università per Stranieri di Perugia³³, previa registrazione al sito. È consultabile tramite CQPWeb, un software per la gestione di corpora testuali, che permette di compiere ricerche semplici e complesse sull’intero corpus, su singoli generi o singole tipologie testuali, per anno o per canale (scritto e parlato), e di gestire i dati attraverso concordanze, liste di frequenza ed estrazione di collocazioni. “Non è invece prevista, per motivi di copyright, la disponibilità dei testi che compongono il corpus” (*ibidem*: 357, nota 5)

In base a quanto già affermato, un punto di forza di questa risorsa è sicuramente la rappresentatività delle diverse varietà della lingua scritta e parlata. Il suo impiego va, tuttavia, escluso a causa delle dimensioni contenute, nettamente inferiori a quelle di altre risorse già citate, e della datazione dei materiali, che si collocano tra il 1990 e il 2012. Quest’ultimo aspetto potrebbe determinare, come nel caso dei già citati corpora CorIS e CoDIS (Cfr.sezione *f*), la presenza di effetti diacronici sulla variazione lessicale, entrando in contrasto con gli obiettivi di uno studio sincronico del neostandard. Inoltre, la scelta di questo corpus comporterebbe l’impossibilità di reimpiegarne i contenuti per motivi di copyright.

n) RIDIRE

Il corpus RIDIRE (Panunzi *et al.* 2014) è il prodotto principale di un progetto “Le risorse dinamiche dell’italiano. Creazione, diffusione e mantenimento di una infrastruttura di rete

³³ www.unistrapg.it/cqpweb

per potenziare l'insegnamento della lingua italiana nel mondo. (RBNE075J8Z)" finanziato dal Fondo Italiano per la Ricerca di Base (FIRB) e diretto da Emanuela Cresti dalla Società Internazionale di Linguistica e Filologia Italiana (SILFI).

È stato realizzato con una procedura di *web-crawling* "mirato", in cui al processo di *crawling* (cfr. sezione *h* di questo capitolo) sono stati applicati criteri di bilanciamento modellati su quelli adottati per il British National Corpus (Cresti, Panunzi 2013: 63), con l'obiettivo di mostrare l'utilizzo della lingua italiana nel web in vari contesti. Il risultato è un corpus che contiene circa 1,5 miliardi di *token* da testi estratti dal web italiano tra il 2009 e il 2013, suddiviso in dodici subcorpora corrispondenti ciascuno a un dominio funzionale, in cui la lingua è caratterizzata dalla funzione con cui viene usata (informazione, economia e business, amministrazione e legislazione) o semantico, in cui la lingua è caratterizzata dall'insieme degli argomenti per cui viene usata (sport, architettura e design, cucina, cinema, musica, moda, arti visive, religione, letteratura e teatro) (cfr. Budinich 2016: 32).

Per quanto riguarda la varietà di lingua disponibile nel corpus,

se prendiamo in considerazione i dieci domini semantici di RIDIRE, definiti dal macro-argomento da cui prendono il nome, possiamo certamente escludere la possibilità di considerarli come campioni di lingue speciali in senso stretto, dato che non paiono caratterizzati dalla presenza di un vero e proprio lessico specialistico; rimane allora la possibilità di considerarli rappresentativi, per lo meno in linea di principio, di lingue speciali in senso ampio. (*ibidem*: 34)

È possibile accedere liberamente³⁴ alle funzioni di ricerca della area di query per selezionare le funzioni disponibili (liste di frequenza, concordanze, pattern e sketch), mentre è richiesto il log-in per le funzioni di gestione, monitoraggio e aggiornamento del corpus. Nell'area di query è possibile attivare la ricerca o per tutti i domini sopra elencati, oppure per uno solo di questi.

Il corpus è stato lemmatizzato e POS-tagato usando il TreeTagger con una versione semplificata dei parametri di Morph-it! (Zanchetta, Baroni 2005).

I punti di forza di questa risorsa sono sicuramente le sue dimensioni, la varietà dei domini coperti, i materiali sufficientemente recenti (la raccolta è terminata nel 2013), seppur non attualissimi, e la libera accessibilità in rete. Tuttavia, com'è già stato

³⁴ www.ridire.it/it.drwolf.ridire/home.seam

menzionato, la maschera di ricerca presenta alcuni limiti nella selezione dei materiali e mostra di avere un numero ridotto di funzionalità, soprattutto in confronto, ad esempio, con altre interfacce, ad esempio Sketch Engine. Inoltre, alcuni dei subcorpora disponibili (ad esempio, amministrazione e legislazione, economia e *business*, religione ecc.) è ipotizzabile che contengano tipologie di testi caratterizzati da un registro linguistico più elevato rispetto alla media formalità che caratterizza il neostandard e l'impossibilità di selezionare solo alcuni subcorpora, escludendone altri, appare poco funzionale agli scopi della ricerca.

o) SMS

SMS (SMS Monitor Studies) è un corpus non accademico di monitoraggio che consta al momento di soli 1.395 SMS, nato come iniziativa informale di un ricercatore non affiliato³⁵. È una raccolta di testi che ha l'obiettivo di rappresentare la varietà di italiano scritto utilizzata negli SMS e che offre a chiunque sia interessato al progetto la possibilità di arricchire il database con l'inserimento di nuovi materiali.

La pagina di riferimento³⁶ mette a disposizione alcuni dati sociodemografici degli autori degli SMS, ma non contiene informazioni sulla data di inizio del progetto, che apparentemente sarebbe ancora in corso, né sul periodo a cui risalgono i testi raccolti. Non sono disponibili informazioni sul tipo di *tagging* dei dati.

I tentativi di effettuare ricerche di concordanze tramite la piccola maschera di ricerca presente sulla pagina sopra menzionata non hanno dato alcun risultato, poiché dopo aver lanciato la ricerca per concordanze si torna indietro alla pagina iniziale del sito. Per accedere alla modalità di interrogazione avanzata è necessario contattare il gestore e tale operazione è subordinata alla collaborazione da parte dell'utente all'incremento dei dati del corpus attraverso l'inserimento di almeno una decina di SMS nel database³⁷.

Pur offrendo un punto di vista interessante per studiare i tratti linguistici dell'italiano contemporaneo, la quantità estremamente ridotta dei testi raccolti (1395) renderebbe assai poco significativi e rappresentativi i risultati di qualunque tipo di ricerca, senza contare che la tipologia di testi raccolti, gli SMS, non corrisponde più al giorno d'oggi al principale

³⁵ www.e-allora.net

³⁶ www.e-allora.net/SMS/ms_index.php

³⁷ *Ibidem*

strumento di comunicazione scritta tramite cellulari, che si svolge per la maggior parte attraverso modalità diverse di messaggistica istantanea, quali la chat³⁸.

4.4. Corpus selezionato per l'indagine: PAISÀ

È un corpus realizzato nell'ambito del progetto PAISÀ (Piattaforma per l'Apprendimento dell'Italiano Su corpora Annotati) che ha coinvolto dal 2009 al 2012 l'Università di Bologna, il CNR di Pisa, l'Accademia Europea di Bolzano e l'Università di Trento, avente l'obiettivo di facilitare l'accesso a una grande quantità di testi scritti in italiano contemporaneo da parte degli utenti del web che desiderano migliorare le proprie conoscenze linguistiche e culturali (cfr. Lyding *et al.* 2014: 36). Nello specifico, i destinatari principali del progetto sono gli emigrati sia di seconda generazione, che pur avendo l'italiano come L1 ne fanno un uso molto limitato, sia di terza generazione, per i quali l'italiano è una seconda lingua (L2).³⁹

I testi che costituiscono il corpus PAISÀ sono rappresentativi dell'uso contemporaneo della lingua italiana, selezionati in maniera automatica da internet **tra settembre e ottobre del 2010**⁴⁰ e liberamente disponibili, poiché protetti da **licenze Creative Commons** di tipo *CC-Attribuzione*, *CC-Attribuzione-Condividi allo stesso modo*, *CC-Attribuzione-Non commerciale*, o *CC-Attribuzione-Non commerciale-Condividi allo stesso modo*⁴¹. A questo proposito:

The main novelty of the PAISÀ web corpus is that it exclusively draws on Creative Commons licensed data, provides advanced linguistic annotations with respect to corpora of comparable

³⁸ “WhatsApp – secondo i dati dell'Osservatorio sulle comunicazioni pubblicati dall'Agcom - ha cannibalizzato i vecchi messaggi di testo, facendone crollare i numeri in modo dirompente. Il traffico sms nel 2016 ha continuato a ridursi del 27,7% rispetto all'anno precedente, toccando 17,8 miliardi di messaggi e portando il passivo a -75% rispetto al livello massimo raggiunto nel 2012 (72,2 miliardi di sms). Il trend è sempre uguale da 4 anni: una decrescita senza scampo. L'invio di messaggi di testo attraverso la rete telefonica è ormai stato soppiantato dalla modalità online. [...] WhatsApp e simili (da Line a WeChat) hanno vinto per grafica, contenuti e potenzialità: l'invio di foto, video, emoticons e audio ha cambiato le nostre abitudini. Il vecchio sms è invece rimasto legato ad un modo di comunicare che appartiene ormai al passato.” Dall'articolo *Ecco come WhatsApp ha seppellito gli sms: -75% in cinque anni* (Il Sole24, 28/12/2016): <https://st.ilsole24ore.com/art/tecnologie/2016-12-28/ecco-come-whatsapp-ha-seppellito-sms-75percento-cinque-anni-195909.shtml?uuiid=ADnEyxLC> (consultato il 02/07/2018).

³⁹ Cfr. www.corpusitaliano.it/it/contents/PAISÀ.html

⁴⁰ “All documents contained in the PAISÀ corpus date back to Sept/Oct 2010” (Lyding *et al.* 2014: 37). Contrariamente a quanto si potrebbe dedurre da tale affermazione, i testi non *risalgono* al 2010, bensì sono stati *raccolti* tra settembre e ottobre 2010, come è stato chiarito dal confronto diretto con due dei ricercatori coinvolti nel progetto PAISÀ (Sara Castagnoli e Marco Brunello). Questi ultimi hanno chiarito che non sono state poste restrizioni temporali nella raccolta di testi disponibili con licenze *Creative Commons*, tra i quali i più recenti risalgono al 2010.

⁴¹ Per approfondire le caratteristiche di ciascuna licenza, si rinvia a: <http://creativecommons.org/licenses/>

size and corpora of web data, and invests in a carefully designed query interface, targeted at different user groups. In particular, the integration of richly annotated language content with an easily accessible, user-oriented interface makes PAISÀ a unique and flexible resource for language teaching. (Lyding *et al.* 2014: 36)

La raccolta dei testi è avvenuta attraverso due procedimenti (cfr. *ibidem*: 37-38). Il primo è il già citato *seeds crawling*, ispirato al progetto WaCky (Baroni *et al.* 2009) (cfr. sezione *h* di questo capitolo): tramite BootCaT, un programma di costruzione di corpora, è stato utilizzato il motore di ricerca Yahoo! per ricercare combinazioni casuali di parole di uso comune, nello specifico 50.000 coppie di parole tratte dal Vocabolario di Base della Lingua Italiana di Tullio De Mauro (1997 [1980]), limitando la raccolta alle pagine dal dominio .it e con le licenze *Creative Commons* precedentemente menzionate. I risultati ottenuti hanno permesso di identificare le URL dei documenti da scaricare per costruire il corpus⁴². Successivamente sono state rimosse dalla lista delle URL tutte le pagine che il motore di ricerca aveva classificato erroneamente come protette da licenze *Creative Commons* (individuate in base ad una *black list* di siti realizzata manualmente durante la realizzazione delle versioni precedenti e sperimentali del corpus).

Il secondo procedimento consiste nell'integrazione nel corpus di altri testi protetti da licenza *Creative Commons* tratti dalle versioni italiane di Wikipedia, Wikinews, Wikisource, Wikibooks, Wikiversity e Wikivoyage. Come riportato in Lyding *et al.* (2014: 38), ciò è stato reso possibile dall'iniziativa della Wikimedia Foundation nel 2009 di rendere pubblici i propri contenuti dietro licenza dapprima GNU e in seguito *CC-Attribuzione-Condividi allo stesso modo*, permettendo così di inserire nel corpus una grande quantità di testi delle versioni italiane dei siti sopra nominati.

Successivamente si è proceduto a scaricare il contenuto degli indirizzi URL, a ripulire le pagine attraverso strumenti KrdWrd⁴³ o, nel caso specifico delle pagine di Wikipedia, attraverso il Wikipedia Extractor in combinazione con uno script che permette di separare i singoli documenti. Dopodiché sono stati rimossi dai contenuti i file vuoti e i file contenenti meno di 150 parole

⁴² Né in Lyding *et al.* (2014) né sulla pagina web del corpus risulta che in una fase successiva sia stato fatto il crawling a partire dalle pagine trovate.

⁴³ “The KrdWrd Project ran from 2008 to 2011. The mission statement was: *provide tools and infrastructure for acquisition, visual annotation, merging and storage of web pages as parts of bigger corpora; develop a classification engine that learns to automatically annotate pages, and provide visual tools for inspection of results.* Basically, it was an infrastructure for research into web page cleaning.” (Cfr. <https://krdwrld.github.io>)

I procedimenti illustrati hanno permesso di realizzare un corpus costituito da 250 milioni di *token* e i testi che lo compongono hanno una lunghezza media di 637 parole.

Per quanto riguarda invece le fonti dei testi raccolti,

The corpus contains approximately 388,000 documents. [...]The documents come from several web sources which, at the time of corpus collection, provided their content under Creative Commons license [...]. About 269,000 texts are from Wikimedia Foundation projects [...].The remaining 119,000 documents come from guide.supereva.it (ca. 19,000), italy.indymedia.org (ca. 10,000) and several blog services from more than another 1,000 different sites (e.g. www.tvblog.it (9,088 pages), www.motoblog.it (3,300), www.ecowebnews.it (3,220), and www.webmasterpoint.org (3,138)). (Lyding *et al.* 2014: 37)

Ai testi sono poi stati aggiunti diversi livelli di annotazione sia sul piano linguistico, sia sul piano dei metadati. Le annotazioni sul piano linguistico nel formato CoNLL (Buchholz, Marsi 2006) includono l'annotazione per parti del discorso tramite l'ILC-POS-Tagger del CNR di Pisa, a cui è stato applicato il TANL *fine-grained pos tagset*, la lemmatizzazione, il *parsing* (ossia, l'analisi sintattica, tramite il DeSR Dependency Parser dell'Università di Pisa, a cui è stato applicato l'ISST-TANL Dependency Tagset), le informazioni morfosintattiche, le relazioni di dipendenza, eccetera. Le annotazioni sul piano dei metadati (ossia, argomento, funzione e genere testuale) sono state svolte in base alla letteratura su generi e funzioni testuali nel web (ad esempio, Santini 2005, Sharoff 2006, Rehm *et al.* 2008, Santini *et al.* 2010, Sharoff 2010). Quest'ultima classificazione è stata realizzata sia per favorire il giusto bilanciamento nella composizione del corpus rispetto a argomenti, funzioni e tipologie testuali ed eventualmente intervenire per bilanciarlo scaricando in maniera mirata ulteriori testi, sia per fornire agli utenti una vasta gamma di varietà testuali e consentire modalità di ricerca avanzate (cfr. Borghetti, Castagnoli, Brunello, 2010). È possibile consultare liberamente online il corpus⁴⁴ ed anche scaricarlo gratuitamente sia la versione annotata che una versione non annotata. Ad ogni documento del corpus è stato attribuito l'indirizzo URL sorgente, il che permette di risalire alle pagine web da cui sono stati tratti i testi, qualora queste siano ancora attive.

Nella sezione 4.5 verranno illustrate le ragioni che hanno portato all'individuazione del corpus PAISÀ come il più idoneo per gli scopi della presente ricerca attraverso un confronto con alcune delle risorse precedentemente presentate.

⁴⁴ www.corpusitaliano.it/it/access/simple_interface.php

4.5. Confronto tra i corpora e individuazione del corpus da utilizzare

Al fine di illustrare le ragioni per le quali il corpus PAISÀ possa essere considerato preferibile rispetto alle altre risorse citate, si farà riferimento alle seguenti tabelle (4.1 e 4.2) e all'appendice 1 che permettono di effettuare un rapido confronto delle risorse incrociando i dati raccolti.

Poiché al centro di questa ricerca vi è la varietà neostandard attualmente in uso, il criterio principale in questa fase finale di confronto tra le risorse citate sarà l'attualità dei materiali. Pertanto, innanzitutto nella tabella 4.1 verranno riportati i primi cinque corpora contenenti i materiali più attuali. In secondo luogo, per le cinque risorse più attuali verrà preso in considerazione il rapporto attualità/dimensioni (tabella 4.2). Infine, l'appendice 1 riassumerà tutti i dati raccolti, al fine di fornire un quadro generale.

Osservando la tabella 4.1, dal punto di vista dell'attualità al primo posto vi è **CorIS/CoDIS**, il cui ultimo aggiornamento risale al 2017. Nel suo complesso i suoi contenuti rappresentano una notevole variazione diacronica, essendo stati raccolti dagli anni '80 in poi. Dal punto di vista delle dimensioni, CorIS/CoDIS non si distingue per l'estensione, poiché occupa la settima posizione con 150 milioni di *token* (cfr. tabella 4.2). Risulta, inoltre, sprovvisto delle licenze *Copyleft*.

Al secondo posto rispetto all'attualità dei materiali vi è **ItTenTen**, con l'ultima versione risalente al 2016, che presenta l'estensione maggiore rispetto a tutte le risorse esaminate (5,8 miliardi), tuttavia non essendo state reperite informazioni sulle fonti dei testi che lo compongono, né sullo status del copyright, la risorsa è stata valutata come non pienamente adeguata agli scopi della presente ricerca.

La stessa riflessione è applicabile all'**Araneum Italicum Maius**, che si colloca al terzo posto per l'attualità dei materiali (i più recenti risalgono al 2014) e al quarto per le dimensioni (1,2 miliardi), per il quale non sono disponibili informazioni sulle fonti dei testi che lo compongono, né circa le licenze d'uso dei testi.

Al quarto posto, dal punto di vista dell'attualità si colloca il corpus **RIDIRE**, per il quale la raccolta di testi si è conclusa nel 2013. Pur occupando la seconda posizione rispetto alle dimensioni (1,5 miliardi), presenta due criticità, ossia l'impossibilità di escludere dalle impostazioni di ricerca i singoli subcorpora non adeguati allo studio della varietà dell'uso medio e l'assenza, anche in questo caso, di informazioni sullo status del *copyright* dei testi.

Con la sua quinta posizione rispetto all'attualità dei materiali (i più recenti risalgono al 2010) e la quarta posizione per le dimensioni (250 milioni) ad una prima osservazione il corpus **PAISÀ** non emerge rispetto alle altre risorse già citate, più estese e aggiornate fino ai nostri giorni. Tuttavia, i fattori che ne hanno determinato la scelta sono innanzitutto la presenza di licenze *Creative Commons*, che ne consentono il reimpiego dei risultati nella successiva fase di progettazione delle attività didattiche. In secondo luogo, sono state prese in considerazione le finalità didattiche per l'apprendimento dell'italiano con cui è stato progettato. Infine, la datazione dei suoi materiali⁴⁵, pur non essendo attualissima, è sufficientemente recente da non compromettere lo studio sull'italiano contemporaneo.

Dopo aver identificato il corpus PAISÀ come la risorsa di riferimento per la ricerca, nelle sezioni seguenti si descriveranno la procedura di progettazione delle stringhe di ricerca funzionali all'interrogazione del corpus e la scelta dei materiali utili per la progettazione delle unità didattiche e della piattaforma online su cui collocare tali unità.

Tabella 4.1: Classificazione dei corpora in base all'attualità dei contenuti.

Corpora	Date dei contenuti
1. CorIS/ CoDIS	Prodotti dagli anni '80 al 2017
2. ItTenTen	Scaricati dal web nel 2016
3. Araneum Italicum (Maius e Minus)	Scaricati dal web nel 2014
4. RIDIRE	Scaricati dal web dal 2009 al 2013
5. PAISÀ	Scaricati dal web nel 2010

Tabella 4.2: Classificazione dei corpora in base alle dimensioni

Corpora	Dimensioni della risorsa (n.ro di <i>token</i>)	Date dei contenuti
1. ItTenTen	5.8 miliardi	Scaricati dal web nel 2016
2. RIDIRE	1,5 miliardi	Scaricati dal web dal 2009 al 2013
3. Araneum Italicum Maius	1,2 miliardi	Scaricati dal web nel 2014
4. PAISÀ	250 milioni	Scaricati dal web nel 2010
5. CorIS/ CoDIS	150 milioni	Prodotti dagli anni '80 al 2017

⁴⁵ Richiamando quanto già espresso alle note n.1 e 40 di questo capitolo, i testi più recenti contenuti nel corpus risalgono al 2010.

4.6. Progettazione delle stringhe di ricerca

Una volta selezionato il *corpus* PAISÀ come risorsa di riferimento per la ricerca, le attività successive si sono concentrate sulla sua consultazione, al fine di reperire al suo interno testi contenenti i tratti linguistici elencati da Sabatini (1990). Nello specifico, ciò ha comportato l'elaborazione di stringhe di ricerca in linguaggio CQL (§ 4.7.1) funzionali al reperimento nel *corpus* delle occorrenze dei suddetti tratti linguistici (§4.7.2), l'interrogazione del *corpus* e il salvataggio dei risultati ottenuti (§ cap. 5) e, infine, il calcolo della loro ricorrenza all'interno del corpus, al fine di determinare i fenomeni linguistici più frequenti (§ cap. 5).

4.6.1. Il linguaggio CQL

CQL (*Corpus Query Language*) è un linguaggio di interrogazione di corpora che permette di ricercare strutture grammaticali o lessicali complesse, funzione indispensabile per il reperimento dei tratti linguistici elencati da Sabatini (1990). È stato sviluppato dall'*Institut für Maschinelle Sprachverarbeitung* dell'Università di Stoccarda nei primi anni '90, dove è stato impiegato nel programma di interrogazione di corpora IMS Open Corpus Workbench (Christ 1994a, 1994b). Un'estensione del linguaggio originale, rispetto al quale presenta alcune differenze⁴⁶, è utilizzata sia in Sketch Engine che nella sua versione *open source* NoSketch Engine: quest'ultima è stata impiegata nell'ambito di questo progetto per la consultazione del corpus PAISÀ, in seguito ad alcune difficoltà tecniche riscontrate con l'uso dell'interfaccia originale.

La risorsa mette a disposizione diverse modalità di ricerca. Nella modalità di ricerca base delle concordanze è possibile inserire una parola o una frase breve sequenza di parole; digitando il lemma di una parola, si otterrà tutta la forma flessionale; mentre digitando una forma flessionale si troveranno occorrenze solo per la forma inserita. Nella ricerca avanzata è, invece, possibile scegliere tra le seguenti opzioni: *lemma*, *phrase*, *word*, *character* e *CQL*.

Con l'opzione *lemma*, inserendo nella maschera di ricerca la forma base di una parola, si otterranno le occorrenze di tutte le sue forme flessionali. Ad esempio, digitando un verbo all'infinito, verranno reperite le occorrenze di tutta la sua coniugazione; lo stesso vale

⁴⁶ www.sketchengine.eu/documentation/corpus-querying, dove si descrive la versione realizzata tramite Manatee 2.122, risalente all'aprile 2015.

inserendo un aggettivo o un sostantivo al singolare o una preposizione semplice con cui si possano ottenere delle preposizioni articolate.

L'opzione *phrase* permette di reperire una sequenza di parole nell'ordine esatto in cui sono state digitate.

Word troverà le occorrenze di una parola che rispettano la grafia esatta di come è stata digitata. Tale procedura si utilizza quando si ricerca una parola considerandola solo come una determinata sequenza di grafemi e, dunque, indipendentemente dalla categoria grammaticale a cui appartiene. L'opzione *Character* troverà tutte le parole che contengono il carattere o la sequenza di caratteri digitati. Ad esempio, digitando *sch* si otterranno le occorrenze di tutte le parole che contengono tale sequenza (esempi: *schizofrenia*, *rischio*, *maschio* ecc.)

Infine, l'opzione CQL permette di utilizzare il *Corpus Query Language* (Evert 2009) per effettuare ricerche più complesse, utilizzando alcuni attributi (*word*, *lemma*, *tag*, *lembo* e *lowercase*), a cui possono combinarsi le espressioni regolari (*regular expressions*) e i caratteri jolly (*wild cards*). Di seguito verrà data una breve descrizione solo delle modalità di ricerca utilizzate nella progettazione delle stringhe.

Le opzioni *Word* e *Lemma*, già citate tra le varie modalità di ricerca, possono essere utilizzate anche tramite CQL secondo queste due modalità: ad esempio, la stringa `[word="gatto"]` o semplicemente l'uso delle virgolette come in "gatto" permettono di identificare tutte le occorrenze della parola ricercata rispettando la grafia esatta in cui è stata digitata, senza dover impostare la ricerca tramite la categoria grammaticale (in questo caso come sostantivo). Le stesse considerazioni valgono per la modalità di ricerca *Lemma* in CQL: la stringa `[lemma="andare"]` consente di reperire le occorrenze di tutta la sua coniugazione, senza dover impostare la ricerca tramite la categoria grammaticale *verbo*.

L'attributo *tag* contiene informazioni sulla parte del discorso di ciascun *token* e di solito anche altre informazioni grammaticali o morfologiche come numero, genere, tempo, ecc. I *tag* vengono assegnati automaticamente da un *tagger*. L'uso del *tag* per la ricerca consente di trovare tutte le occorrenze etichettate con la stessa parte del discorso (esempio: `[tag="A"]` identifica gli aggettivi), come verrà illustrato nel dettaglio più avanti.

Le espressioni regolari sono un insieme di simboli speciali che possono essere utilizzati per ricercare modelli anziché caratteri specifici, ad esempio per trovare tutte le parole che iniziano, contengono o finiscono in una specifica sequenza di caratteri. Un esempio, che è stato spesso utilizzato nelle stringhe di ricerca prodotte è `(?i)`. Questa

espressione regolare sta ad indicare la funzione *case insensitive*, che serve a reperire occorrenze con lettera maiuscola e minuscola e quindi i casi in cui la parte del discorso ricercata può trovarsi ad inizio frase o al suo interno.

La *wild card*, invece, è un carattere che all'interno di una stringa rappresenta altri caratteri o sequenze di caratteri. Un esempio è l'uso dell'asterisco (*), che rappresenta genericamente un numero indefinito e variabile di caratteri.

Per apprendere il linguaggio CQL e l'uso del programma NoSketch Engine sono stati utilizzati il manuale di Nesi e Thompson (2014) e la documentazione presente sul sito di Sketch Engine⁴⁷, che hanno consentito di familiarizzare con le diverse modalità di interrogazione del *corpus* e di elaborare delle stringhe di ricerca funzionali al reperimento di strutture grammaticali complesse. Per quest'ultima attività è stata necessaria, inoltre, la consultazione del *tagset* morfosintattico (ISST TANL *tagset*)⁴⁸ e del *tagset* delle dipendenze sintattiche (ISST TANL DEP*tagset*)⁴⁹ utilizzati specificamente per l'annotazione del corpus PAISÀ. L'ISST TANL è un sistema di annotazione per parti del discorso: attraverso un set di 14 *tag* generici (*coarse-grained pos tag*) e 37 *tag* specifici (*fine-grained pos tag*) ad ogni *token* contenuto nel corpus vengono associate informazioni relative alle categorie morfo-sintattiche a cui esso appartiene. Ad esempio, nell'espressione "il mio libro" a *mio* corrisponde sia il *tag* generico che identifica in generale tutti gli aggettivi, ovvero [cpostag="A"] oppure [postag="A.*"], sia il *tag* specifico [postag="AP"], che identifica gli aggettivi possessivi.

Alle parti del discorso possono, inoltre, essere associate delle caratteristiche morfosintattiche (*morpho-syntactic features*, il cui *tag* è *feats*), di cui di seguito a titolo di esempio si riportano alcune tipologie: genere (*tag*= gen), numero (*tag* = num), persona (*tag*= per), modo (*tag* = mod) e tempo (*tag*= tmp)⁵⁰. Ciascuna di esse è ulteriormente esplicitata da una sottocategoria di caratteristiche che forniscono informazioni più dettagliate sulle parti del discorso, a cui corrisponde un set di *tag*. Ad esempio, nella seguente stringa, che identifica i sostantivi plurali, la caratteristica morfosintattica del numero (*feats*=num) è ulteriormente esplicitata dal plurale (p):

```
[postag="S.*" & feats=".*num=p.*"]
```

⁴⁷ www.sketchengine.eu/documentation/cql-basics

⁴⁸ www.corpusitaliano.it/static/documents/POS_ISST-TANL-tagset-web.pdf

⁴⁹ www.italianlp.it/docs/ISST-TANL-DEPtagset.pdf

⁵⁰ Le caratteristiche qui descritte non sono standard, bensì si applicano al *tagset* utilizzato nel corpus PAISÀ.

L'uso dell'ISST TANL DEPtagset ha permesso di progettare delle stringhe di ricerca che tengano conto delle relazioni di dipendenza sintattica. Tra queste si citano a titolo di esempio le relazioni di soggetto, oggetto e di verbo principale, che sono state utilizzate per ricercare i diversi casi di soggetto in posizione rematica rispetto al predicato e del verbo in forma pronominale per indicare partecipazione affettiva, a cui si farà riferimento nel dettaglio nella sezione 4.6.2 (cfr. i punti 7c-h e il punto 9 dell'elenco delle stringhe prodotte).

Quest'ultimo *tagset* si è rivelato una risorsa fondamentale per poter interrogare il corpus rispetto ad alcuni fenomeni grammaticali che potremmo definire più complessi, poiché coinvolgono la sintassi del periodo e non la semplice co-occorrenza di parti del discorso all'interno della stessa frase. Tuttavia, come verrà precisato sia nella sezione 4.6.2, sia nel capitolo 5, è bene fare alcune riflessioni su alcune scelte compiute nella formulazione delle stringhe stesse. È necessario precisare che i risultati ottenuti dall'utilizzo di stringhe che fanno uso dell'annotazione sintattica di ricerca dipendono necessariamente dalla formulazione stessa delle stringhe, la quale a sua volta è determinata dalle opzioni a disposizione nell'uso del linguaggio CQL. Nel caso, ad esempio, della ricerca delle diverse tipologie di soggetto in posizione rematica rispetto al predicato, le stringhe massimizzano le possibilità che la parte del discorso che segue il verbo sia effettivamente il soggetto del verbo che occorre a inizio frase, ma non lo garantisce, dato che imporre tramite il linguaggio CQL la ricerca di un determinato soggetto legato ad un determinato verbo è un'operazione molto complessa. È il caso, ad esempio del punto 7e- Soggetto in posizione rematica rispetto a verbi composti con l'ausiliare *avere*, senza l'oggetto diretto tra participio e soggetto (esempio: *ha cancellato lei la pagina*; cfr. sezione 4.6.2): una prima versione della stringa includeva anche i casi di frasi con soggetto ellittico, ad esempio *ho visto Marco*, in cui il soggetto ellittico è *io*, ma la funzione di soggetto all'interno del corpus è stata assegnata a *Marco*. Ciò ha permesso di ipotizzare che il tagger abbia assegnato la funzione di soggetto (`deprel="subj"`) a qualunque sostantivo o pronome collocato subito dopo il verbo: una procedura che non sempre ha condotto a risultati corretti.

Inoltre, sono state riscontrate alcune criticità del tagset nel caso 7d -Soggetto in posizione rematica rispetto a verbi composti con l'ausiliare *essere* (esempio: *Sono state moltissime le reazioni*): nella produzione della stringa è emerso che il *tagset* non fa distinzione tra participi passati o presenti, dunque la ricerca è stata impostata sul solo *tag* del modo participio. Sempre nella ricerca di questo fenomeno grammaticale sono state

riscontrate delle discrepanze tra alcuni *tag* relativi ai tempi passati dei verbi, riportati nella guida al *tagset*, e quelli poi risultati effettivamente in uso nel corpus, a cui si farà riferimento nel dettaglio nella sezione 4.6.2.

In altri casi, pur non essendo stati riscontrati dei limiti nell'uso del linguaggio CQL, è stato necessario impostare le stringhe trovando un equilibrio tra la creazione di stringhe ampie, che identificassero tutti i casi che rientrano nel fenomeno indagato, e stringhe più mirate e restrittive, che non includessero troppo “rumore”, ossia casi che, pur presentando delle analogie strutturali con il fenomeno indagato, non fossero inerenti. Un esempio è il punto 12c -*Che* pseudorelativo (una sottocategoria del *che* polivalente. Esempio: *li vedo che scendono*. Cfr. sezione 4.6.2), ove la prima versione della stringa⁵¹ includeva tra i risultati anche molte occorrenze del *si* impersonale (esempio: *si vede che hanno qualche domanda*), che chiaramente non rientrano nel fenomeno indagato. Di conseguenza, nella seconda fase di progettazione della stringa si è scelto di escludere il pronome *si* dai pronomi ricercabili, trovando un compromesso tra la perdita di occorrenze corrette e l'esclusione di casi non inerenti⁵².

Un altro esempio dell'esigenza di restringere il campo della ricerca è il caso 9-impiego del verbo in forma pronominale per indicare partecipazione affettiva (esempio: *mi guardo un film*), ove, utilizzando la prima versione⁵³ delle stringhe prodotte, la ricerca restituiva molte occorrenze che non rientrano nell'ambito del fenomeno in questione (esempi: *Mi faccio (del) male; mi sono fatto accompagnare; mi sono visto rimproverare; mi faccio prendere; mi sono visto costretto ecc.*). Pertanto, allo scopo di ridurre il “rumore” tra i risultati, si è preferito restringere la ricerca, modificando la stringa e imponendo come elemento finale la presenza di un gruppo nominale col valore di oggetto⁵⁴, come verrà illustrato nel dettaglio nella sezione 4.6.2.

⁵¹ La prima versione della stringa mirata a reperire i casi di *che* pseudorelativo è tratta da Budassi (2017:165), Rispetto all'originale presenta le sole modifiche relative al *tagset* in uso in PAISÀ:

[postag="P.*"] [lemma="vedere|sentire|guardare|scorgere|notare|osservare|mirare|ascoltare|udire|intendere|percepire"] [word="che"] [postag="V.*"]

⁵² La seconda versione della stringa, modificata rispetto a Budassi (*ibidem*), con l'esclusione del pronome *si*:

[postag="P.*&lemma!="si"] [lemma="vedere|sentire|guardare|scorgere|notare|osservare|mirare|ascoltare|udire|intendere|percepire"] [word="che"] [postag="V.*"] Questa versione esclude dalla ricerca il pronome *si* tramite la seguente parte della stringa: [postag="P.*&lemma!="si"].

⁵³ Un esempio della prima versione della stringa per la prima persona singolare dei verbi riflessivi “affettivi”, con tempi verbali non composti:

[word="(?)mi"] [lemma="mangiare|leggere|bere|fare|guardare|vedere"&feats="*num=s.*" & feats="*per=1.*"]

⁵⁴ Un esempio della seconda versione della stringa per la prima persona singolare dei verbi riflessivi “affettivi”, con tempi verbali non composti:

La riflessione derivante da quanto esposto, e che verrà ripresa successivamente nel corso del capitolo e nel capitolo 5 è che le occorrenze ottenute dipendano notevolmente dalla formulazione della stringhe di ricerca, la quale a sua volta è in alcuni casi influenzata o dai limiti imposti dalle opzioni attualmente a disposizione nell'uso del linguaggio CQL o dalla necessità di ricercare un equilibrio tra la copertura ad ampio spettro di un fenomeno grammaticale e il rischio di reperire anche grandi quantità di casi non inerenti. Alla progettazione stringhe di ricerca e alle scelte sottostanti si farà riferimento nella sezione 4.6.2.

4.6.2 Elaborazione delle stringhe di ricerca

Il punto di partenza per l'elaborazione delle stringhe di ricerca è stata la rilettura della lista dei 14 tratti linguistici elencati da Sabatini (1990) (cfr. Capitolo 1, §1.4.1). Da questa lista è stato escluso per le finalità della ricerca il primo fenomeno elencato, ossia l'uso dei pronomi *lui* e *lei* e *loro* in funzione di soggetto al posto di *egli/ella* ed *essi/esse*, ritenendo che tali forme siano ormai integrate nello standard:

L'uso di *lui*, *lei* e *loro* come pronomi soggetto è stato a lungo sanzionato dai grammatici del passato e oggetto di dibattito nella storia della grammatica italiana fin dal Cinquecento. Oggi non vi sono più dubbi sull'accettabilità di queste forme.⁵⁵

Di conseguenza, partendo dai 13 tratti linguistici rimanenti, sono state elaborate 57 stringhe di ricerca (*query*) in linguaggio CQL. Com'è evidente, il numero di stringhe è nettamente superiore ai tratti elencati. Come verrà di seguito illustrato nei capitoli 4 e 5, l'elaborazione di più stringhe si è resa necessaria principalmente in due casi: per i tratti linguistici articolati in sottofenomeni, che chiaramente vanno cercati separatamente, e nel caso in cui le specifiche caratteristiche sintattiche o morfologiche di un determinato tratto hanno inevitabilmente richiesto la progettazione di più stringhe di ricerca (si veda ad esempio il caso dei verbi in forma pronominale per indicare partecipazione affettiva, per i

```
[word="(?)mi"] [lemma="mangiare|leggere|bere|fare|guardare|vedere"&
feats="*num=s.*" & feats="*per=1.*"] [postag="R.*|A.*|D.*"]{0,2}
[postag="P.*|S.*" &deprel="obj"]
```

⁵⁵ EGLI/LUI, in La grammatica italiana, Istituto dell'Enciclopedia Italiana www.treccani.it/enciclopedia/egli-lui_%28La-grammatica-italiana%29 (consultato il 05/01/2019), versione html, pertanto priva di pagine. Si faccia riferimento, a questo proposito, anche alle motivazioni linguistiche fornite da Francesco Sabatini in risposta ad un quesito sull'uso di lui/egli soggetto posto sulla rivista La Crusca per voi (n.1, ottobre 1990, pp. 6-8), disponibile online alla pagina: www.accademiadellacrusca.it/it/lingua-italiana/consulenza-linguistica/domande-risposte/soggetto (consultato il 18/07/2018)

quali un'unica stringa di ricerca non avrebbe potuto coprire tutte le forme pronominali usate per l'intera coniugazione). Quindi, il numero di stringhe finali è in alcuni casi correlato linguisticamente alla presenza di sottofenomeni, in altri casi dall'oggettiva impossibilità di coprire con un'unica stringa di ricerca tutte le occorrenze di un fenomeno, a causa della sua complessità intrinseca. Ciò ha portato, quindi, il numero finale di tratti indagati a 26.

Qui di seguito si riporta per ciascun fenomeno la descrizione delle stringhe di ricerca prodotte, di cui viene fornita una breve descrizione. L'elenco completo di tutte le stringhe è, inoltre, disponibile in appendice (cfr. Appendice 2).

1) impiego di *gli* nella forma dativale al posto di *le* (a lei) e *loro* (a loro). Allo scopo di reperire occorrenze in cui *gli* può trovarsi ad inizio frase o al suo interno, nella stringa è stata inserita la già menzionata espressione regolare (?i), che permette di ottenere occorrenze sia con lettera maiuscola, sia con lettera minuscola. A seguire vi è un verbo, al fine di individuare tutti i casi in cui *gli* si trova in posizione preverbale (esempi: *gli ho detto*; *gli rispondo*, ecc).

Poiché si è imposto nella stringa che *gli* fosse seguito immediatamente da un verbo e ciò può verificarsi solo nel caso in cui esso funga da pronome (esempio: *gli dico* = pronome + verbo) e non da articolo determinativo maschile plurale (esempio: *gli alberi*), si è ritenuto superfluo progettare una stringa di ricerca più complessa, in cui *gli* venisse indicato specificamente come pronome clitico (esempio, [postag="PC" & word="(?)gli"]). Per tale ragione, la sua ricerca è stata impostata come semplice parola [word="(?)gli"].

La discriminazione tra i casi in cui il pronome *gli* si riferisce secondo lo standard a individui di genere maschile (esempio: *ho visto Marco e gli ho detto...*) o, nel registro colloquiale, a individui di genere femminile singolare (esempio: *ho parlato ad Anna e gli ho detto...*) o di genere maschile/femminile plurale al posto di *loro* (esempio: *ho chiamato i miei amici/ le mie amiche e gli ho detto...*), è stata fatta manualmente, tramite la consultazione delle occorrenze ottenute e la lettura del contesto circostante. Segue la stringa prodotta:

```
[word="(?)gli"] [postag="V.*"]
```

2) impiego del partitivo preceduto da preposizione (esempio: *con degli amici*). Per questo fenomeno è stata progettata una stringa in cui una preposizione semplice, indicata

dal tag E, purché diversa da *prima* e *dopo*, sia immediatamente seguita da una preposizione articolata (tag = EA) composta da *di* + articolo:

```
[postag="E" & word!="prima|dopo"] [postag="EA" & lemma="di"]
```

L'esclusione di *prima* e *dopo* è determinata dall'esigenza di evitare occorrenze come "prima/dopo di + articolo" (esempio: *prima del concerto*, *dopo dei pasti*, ecc.) che non costituiscono casi di partitivo, ma semplici espressioni di tempo;

3) impiego di dislocazioni a sinistra (*quel libro l'ho già letto*), **a destra** (*l'ho già letto, quel libro*) della parte rematica della frase con la ripresa pronominale atona e **costruzioni ad esse assimilabili**, tipo: *di pane non ne ho più*, *a me mi piace leggere*, con uso pleonastico della particella pronominale. Nella progettazione delle stringhe si è tenuto conto solo della dislocazione a sinistra dell'elemento pronominale (*a me mi*) e del cosiddetto "doppio partitivo" (*di pane non ne ho più*) e non di tutti i casi di dislocazione a sinistra e a destra di altri elementi grammaticali, come il gruppo nominale preceduto o meno da preposizione (a sinistra: *il libro, lo ha preso Giacomo; a Giorgio gli porto una bottiglia di vino*; a destra: *lo ha preso Giacomo, il libro; gli porto una bottiglia di vino, a Giorgio*) o il verbo (*piovere, non è piovuto, per imparare, Carla impara*) eccetera. Nel compiere tale scelta si è tenuto conto dell'obiettivo finale del progetto, ossia la progettazione di materiali didattici destinati a non-italofoni nativi. In questa prospettiva, si è ritenuto che i casi di dislocazione dell'elemento pronominale potessero risultare più produttivi dal punto di vista della progettazione didattica⁵⁶ e più utili dal punto di vista degli utenti finali delle unità didattiche, data la complessità del fenomeno, che, nel caso della dislocazione a sinistra, vede coinvolto un raddoppiamento della forma pronominale (forma tonica + atona, ad esempio: *a me mi piace*), rispetto invece agli altri casi di dislocazione sopra citati (esempio: *quel libro l'ho già letto; l'ho già letto, quel libro*, ecc.), nei quali il significato della frase comunque è facilmente deducibile dal contesto. Sulla base di tali motivazioni sono state elaborate le seguenti stringhe di ricerca:

3a) doppio partitivo (esempio: *di pane non ne ho più*): per questo fenomeno è stata progettata una stringa in cui la preposizione *di*, scritta con lettera maiuscola o minuscola:

- 1) non è mai preceduta da qualunque tipo di sostantivo o aggettivo, per evitare occorrenze in cui la preposizione *di* sia parte di un'espressione che esprime il possesso (esempio: *casa*

⁵⁶ In tutta la manualistica dell'italiano L2/LS in commercio al momento dello svolgimento della presente ricerca, pari a 32 pubblicazioni (Słapek 2017), *Affresco italiano C2* (Le Monnier) è l'unico manuale a trattare la dislocazione a destra e a sinistra.

di Anna) o la materia (esempio: *scatola di cartone*) o di un comparativo (esempi: più grande di; maggiore/minore di); 2) è immediatamente seguita da qualunque tipo di sostantivo (purché diverso da *fatto*, *seguito*, *conseguenza*, *norma*, *solito*, per evitare occorrenze come: *di fatto*, *di seguito*, ecc.), poi dalla particella pronominale *ne* e, infine, da un verbo di qualunque tipo. Ne deriva:

```
[postag!="S.*|A.*"] [word="( ?i) di"] [postag="S.*" &
word!="fatto|seguito|conseguenza|norma|solito"] [word="ne"]
[postag="V.*"]
```

3b) dislocazione a sinistra dell'elemento pronominale (*a me mi*, *a te ti*, ecc.). In questo caso è stato necessario progettare 8 stringhe di ricerca per le 7 forme pronominali toniche precedute dalla preposizione *a* e seguite dalla forma atona (*a me mi*, *a te ti*, *a lui gli*, *a lei le*, *a noi ci*, *a voi vi*, *a loro gli*), più un'ottava forma che corrisponde ad un'alternativa della forma *a lei le*, ossia *a lei gli* (esempio: *a lei gli dico*), tipica delle varietà substandard. Di seguito si riporta la prima delle 8 stringhe progettate, ossia la preposizione *a*, scritta con lettera maiuscola o minuscola, seguita immediatamente da *me* e *mi.*, rinviando all'appendice 2 per una visualizzazione completa del fenomeno:

```
[word="( ?i) a"] [word="me"] [word="mi"]
```

4) impiego del connettivo *per cui* (derivato per ellissi da *motivo/fatto/ragione per cui*, oppure per generalizzazione di casi con normali antecedenti nominali) **in funzione di congiunzione causale-consecutiva**. È stata progettata la seguente stringa, ove *per*, scritto con lettera maiuscola o minuscola, è stato ricercato come semplice parola seguita immediatamente da *cui*:

```
[word="( ?i) per"] [word="cui"]
```

5) impiego di *cosa?* interrogativo diretto e indiretto al posto di *che cosa?* Per il fenomeno in questione sono state progettate due stringhe.

Nel caso dell'**interrogativa diretta** è stato imposto che la stringa inizi con qualunque segno di punteggiatura ("F"), per delimitare la fine della frase precedente; in apertura di frase può esservi o meno⁵⁷ un elemento grammaticale, purché sia diverso da qualunque aggettivo ("A"), determinante ("D")⁵⁸, articolo determinativo o indeterminativo ("R"), per evitare di reperire occorrenze simili a *bella cosa*, *questa/quella/alcuna/tale/qualsiasi cosa*, *la cosa*, *una cosa*, ecc. e, allo stesso tempo, non impedire il reperimento di casi in cui *cosa*

⁵⁷ La presenza opzionale è rappresentata dal simbolo del punto interrogativo.

⁵⁸ Per la voce "determinante", si rinvia alla guida:

www.corpusitaliano.it/static/documents/POS_ISST-TANL-tagset-web.pdf

sia preceduto, ad esempio, da congiunzioni, avverbi ecc. (Esempi: *ma cosa dici? E ora cosa fai?* Ecc.). Immediatamente dopo vi è *cosa*, scritto maiuscolo o minuscolo (a seconda che sia posto ad apertura di frase o meno); a seguire è stato imposto che non vi sia immediatamente un punto interrogativo, ma un numero di parole che spaziano da 0 a 10 e, infine, il punto interrogativo. Segue la stringa:

```
[cpostag="F"] [postag!="A.*|D.*|R.*"]? [word="( ?i) cosa"]
[word!="\?" ] {0,10} [word="\?" ]
```

Nel caso dell'**interrogativa indiretta** (esempio: *mi domando cosa...*) in apertura della stringa vi sono i lemmi dei verbi *chiedere* e *domandare*; qualora tra questi verbi e *cosa* siano presenti degli elementi grammaticali (ad esempio: *mi chiedo allora cosa...*), la cui presenza opzionale è simboleggiata dal punto interrogativo, si impone che questi siano diversi sia dai segni di punteggiatura (per evitare di reperire esempi di discorso diretto), sia da *che* (per evitare l'interrogativo standard *che cosa*). Segue la stringa:

```
[lemma="chiedere|domandare"] [cpostag!="F" & word!="che"]?
[word="( ?i) cosa"]
```

6) impiego di *e*, *ma*, e di connettivi testuali come *allora*, *comunque* in posizione iniziale di discorso o di enunciato. In questi casi si è imposto che la stringa inizi con un punto o un punto e virgola, seguiti dalla congiunzione o dal connettivo testuale in questione con lettera maiuscola o minuscola (a seconda che siano preceduti o meno dal punto):

- a) [word="\.| \;"] [word="(?i) E"]
- b) [word="\.| \;"] [word="(?i) Ma"]
- c) [word="\.| \;"] [word="(?i) Allora"]
- d) [word="\.| \;"] [word="(?i) Comunque"]

7) soggetto in posizione rematica rispetto al predicato, talvolta con discordanza di numero con questo (esempi: *non ci sono soldi*; o anche il tipo: *niente soldi!*). Data la complessità del fenomeno in questione, che si articola in numerosi sottocasi, è stato necessario elaborare 8 stringhe di ricerca, partendo dai casi più particolari, caratterizzati dalla discordanza di numero tra soggetto e predicato:

7a) niente seguito da un sostantivo plurale (esempio: *niente soldi*): è stata elaborata una stringa in cui la parola *niente* è seguita da un sostantivo di genere plurale:

```
[word="niente"] [postag="S.*" & feats=".*num=p.*"]
```

7b) sostantivo plurale seguito da *non ce n'è* (esempio: *soldi non ce n'è*): è stata elaborata una stringa in cui un sostantivo di genere plurale è seguito da *non + ce + n'è*, infine, dal verbo *essere* alla terza persona singolare:

```
[postag="S.*"& feats=".*num=p.*"] [word="non"] [word="ce"]
[word="n'"] [lemma="essere" & feats=".*num=s.*"&
feats=".*per=3.*"]
```

7c) Soggetto in posizione rematica rispetto a verbi non composti (esempio: *parla il presidente*): nella costruzione della stringa sono stati esclusi tutti i casi di frasi che nella costruzione non marcata richiedono il soggetto in posizione post-verbale (per il dettaglio, si rinvia al capitolo 2).⁵⁹

Pertanto, è stato imposto che la stringa inizi con un punto, per delimitare la fine della frase precedente; in apertura di frase vi è un verbo con lettera maiuscola, che sia diverso da *essere*, *occorrere*, *servire*, *avere* e *esistere*, avente la funzione di verbo principale e coniugato al modo indicativo. A seguire possono essere presenti in quantità compresa tra 0 e 2 un articolo determinativo o indeterminativo (R.*) e/o un aggettivo (A.*) o, in alternativa all'articolo, un determinante (D.*) (ad esempio: *questo*, *codesto*, *quello*, *quale*, *alcuno*, *certo*, *tale*...) e, infine, un sostantivo o un pronome di qualunque tipo ("S.*|P.*") che ricopra il ruolo di soggetto:

```
[word="\." ] [postag="V.*" & word="[A-Z].+" &
lemma!="essere|occorrere|servire|avere|esistere" &
deprel="ROOT" & feats=".*mod=i.*"]
[postag="R.*|A.*|D.*"] {0,2} [postag="S.*|P.*" &
deprel="subj"]
```

Esempi di occorrenze reperibili con questo tipo di stringa (per pura praticità verranno presentati solo casi con soggetto al singolare): *parla (V.*) Luca (S.*)*; *parla (V.*) lui (P.*)*; *interviene (V.*) il/un (R.*) deputato (S.*)*; *interviene (V.*) il/un (R.*) nostro (A.*) deputato (S.*)*; *recita (V.*) il/un (R.*) simpatico (A.*) artista (S.*)*; *gioca (V.*) questo/quel (D.*) calciatore (S.*)*, *gioca (V.*) questo/quel (D.*) bravo (A.*) calciatore (S.*)*, ecc.. Come si evince dagli esempi, la formulazione di questa stringa non consente, ad esempio, di ricercare occorrenze in cui il sostantivo che funge da soggetto sia preceduto da più di un aggettivo (esempio: *è arrivato il mio carissimo amico*, ove *mio* è un aggettivo possessivo e *carissimo* è un aggettivo qualificativo). Le occorrenze reperite presentano, infatti, un solo

⁵⁹ Trattandosi di una stringa mirata a ricercare le occorrenze di soggetti in posizione rematica rispetto a verbi non composti, i verbi *essere* e *avere* sono stati necessariamente esclusi dalla stringa per evitare di ottenere occorrenze in cui tali verbi fungono da ausiliari dei tempi composti. Inoltre, sono stati esclusi dalla ricerca anche i verbi *occorrere*, *servire* (che possono essere usati entrambi impersonalmente, mentre nel caso del verbo *servire*, esso può essere usato anche transitivamente) ed *esistere*, perché rientrano nella classe dei verbi intransitivi detti *inaccusativi* (cfr. capitolo 2). L'esclusione di questi verbi è frutto di una seconda fase di elaborazione della stringa, al fine di evitare di includere nei risultati le frasi inaccusative, che ricorrevano nella prima versione.

aggettivo, sia esso possessivo, qualificativo o di altro tipo. Pertanto, è bene ricordare che i risultati ottenuti dall'utilizzo di questa stringa di ricerca dipendono necessariamente dalla formulazione della stringa stessa come già anticipato nella sezione 4.6.1 in riferimento ad altri casi⁶⁰.

7d) Soggetto in posizione rematica **rispetto a verbi composti con l'ausiliare essere** (esempio: *Sono state moltissime le reazioni*): valgono le stesse considerazioni fatte per la stringa precedente, a differenza della quale si è imposto che il verbo sia composto dalle voci del lemma del verbo *essere*, seguite dal participio passato di qualunque verbo, e abbia come soggetto un qualunque sostantivo o pronome:

```
[word="\."] [lemma="essere"] [postag="V.*" &
feats=".*mod=p.*"] [postag="R.*|A.*|D.*"] {0,2}
[postag="P.*|S.*" & deprel="subj"]
```

Sebbene nella descrizione del fenomeno si sia fatto correttamente riferimento al participio passato (es. è *arrivata Anna*), è necessario precisare che, al momento della realizzazione di questa ricerca, nel corpus PAISÀ i participi non risultano taggati come “passati” o “presenti”, bensì solo come participi: la *feature* `tmp=r` prevista nella guida al *tagset* per il passato (*past tense*), e dunque applicabile anche al participio passato, in realtà non risulta attribuita ad alcun elemento linguistico. Ad esempio, il participio *andato* risulta taggato come: `V/Vnum=s|mod=p|gen=m`, ossia come un verbo avente numero singolare, modo participio e genere maschile. Tale considerazione vale per tutti i casi successivi in cui verrà menzionato un participio.

Per avere un riscontro circa l'utilizzo del *tag* per i tempi passati è stata, inoltre, fatta una verifica inserendo nella maschera di ricerca dei verbi al passato prossimo, passato remoto e imperfetto (esempio: *è andato; andai; andava*). Ne risulta che: non esiste un *tag* specifico che identifica il passato prossimo, poiché l'ausiliare e il participio sono sempre stati taggati separatamente (Esempio: `è = V/VA/num=s|per=3|mod=i|ten=p`; `andato = V/V num=s; mod=p; gen=m`); i verbi al passato remoto sono taggati con `ten=s`, anche se questa voce non è riportata nell'elenco della guida al *tagset*, mentre,

⁶⁰ Ad un successivo riesame della stringa prodotta, si è valutato che possa eventualmente essere implementata nel seguente modo:

```
[word="\."] [postag="V.*" & word="[A-Z].+" &
lemma!="essere|occorrere|servire|avere|esistere" & deprel="ROOT" &
feats=".*mod=i.*"] [postag="R.*|D.*"]?[postag="A.*"]{0,}
[postag="S.*|P.*" & deprel="subj"]
```

Questa intuizione, giunta al termine del presente studio, potrà essere il punto di partenza per ulteriori sviluppi della ricerca sul tema.

contrariamente a quanto riportato nella guida, i verbi all'imperfetto sono taggati con `ten=i` (e non `tmp=i`). Data quest'ultima discrepanza tra il *tag* corretto dell'imperfetto reperito nel corpus (`ten=i`) e quello erroneamente riportato nella guida (`tmp=i`), è stato fatto un ultimo tentativo di ricerca del participio passato, modificando il *tag* `tmp=r`, previsto nella guida per il tempo passato, in `ten=r`. Tuttavia, quest'ultimo non è stato reperito tra i risultati, confermando l'osservazione originaria sull'assenza di distinzione tra participi presenti e passati.

7e) Soggetto in posizione rematica rispetto a verbi composti con l'ausiliare avere, senza l'oggetto diretto tra participio e soggetto (esempio: *ha cancellato lei stessa la pagina*): valgono le stesse considerazioni fatte per la stringa precedente, a differenza della quale si è imposto che il verbo sia composto dalle voci del lemma del verbo *avere*, seguite dal participio passato di qualunque verbo. Inoltre, si è imposto che immediatamente dopo il verbo non vi sia la parola *che*. Questo accorgimento è frutto di una seconda fase di elaborazione della stringa, al fine di evitare di includere nei risultati i casi con soggetto ellittico, in cui o il verbo è immediatamente seguito dalla congiunzione subordinante *che* (esempio: *ho pensato che*), oppure vi è un complemento oggetto tra il verbo e la congiunzione *che* (esempio: *ho visto Marco che*). A fronte di questi casi, reperiti in una prima fase di ricerca con una versione differente della stringa, si è ipotizzato che, nella fase di *part-of-speech tagging*, in cui alle parole dei testi che compongono il corpus PAISÀ sono stati attribuiti i *tag* relativi alle parti del discorso, il lemmatizzatore abbia assegnato la funzione di soggetto (`deprel="subj"`) a qualunque sostantivo o pronome collocato subito dopo il verbo. In alcuni casi tale procedura si è rivelata corretta, ma nel caso delle frasi con soggetto ellittico la categoria di soggetto è stata assegnata ad elementi della frase che in realtà erano complementi oggetti (ad esempio, in *ho visto Marco* la funzione di soggetto è stata assegnata a *Marco*). Pertanto, è stata progettata la seguente stringa:

```
[word="\."] [lemma="avere"] [postag="V.*" &
feats=".*mod=p.*"] [postag="R.*|A.*|D.*"] {0,2}
[postag="P.*|S.*" & deprel="subj" & word!="che"]
```

7f) Soggetto in posizione rematica rispetto a verbi composti con l'ausiliare avere, con l'oggetto diretto tra participio e soggetto (esempio: *ha suscitato sdegno l'articolo che racconta...*): valgono le stesse considerazioni fatte per la stringa precedente, a differenza della quale si è imposto che tra il verbo composto e il soggetto vi sia un gruppo nominale con la funzione di oggetto (`deprel="obj"`):


```
[word="\."] [lemma="avere"] [postag="V.*" &
feats=".*mod=p.*"] [postag="R.*|A.*|D.*"]{0,2}
[postag="S.*" & deprel="obj"] [postag="R.*|A.*|D.*"]{0,2}
[postag="P.*|S.*" & deprel="subj" & word!="che"]
```

7g) Soggetto in posizione rematica rispetto a verbi passivi: valgono le stesse considerazioni fatte per le stringhe precedenti, a differenza delle quali si è imposto di ricercare le voci dei lemmi dei verbi *essere*, *venire*, *andare* (che sono i tre ausiliari possibili per la costruzione della forma passiva, cfr. Grandi 2011), seguite opzionalmente dal participio dei soli verbi *essere* e *andare*⁶¹ e poi da qualunque verbo al participio. In chiusura della stringa vi è come sempre un gruppo nominale in funzione di soggetto, per la cui descrizione si rinvia alle stringhe precedenti. Tutto ciò permette di reperire casi come: *fu istituito il Commissario distrettuale* (tempo non composto dell’ausiliare *essere*); *vennero svolte numerose prove* (tempo non composto dell’ausiliare *venire*); *va garantita la centralità* (tempo non composto dell’ausiliare *andare*); *sono state invitate le autorità* (tempo composto dell’ausiliare *essere*); *è andato distrutto il tempio* (tempo composto dell’ausiliare *andare*).

```
[word="\."]
[lemma="essere|venire|andare"] [lemma="essere|andare" &
feats=".*mod=p.*"]? [postag="V.*" & feats=".*mod=p.*"]
[postag="R.*|A.*|D.*"] {0,2} [postag="P.*|S.*"&
deprel="subj"]
```

7h) Soggetto in posizione rematica in frase subordinata (esempio: *non cambia nulla, se resta l'unilateralismo*): è stato imposto che la stringa inizi con una congiunzione subordinante con lettera non maiuscola; successivamente vi è un verbo che non sia al modo infinito, seguito da altri elementi linguistici (ossia, eventuali altre parole che compongono la frase) in quantità variabile da 1 a 2. A seguire possono essere presenti in quantità compresa tra 0 e 2 un articolo determinativo o indeterminativo e/o un aggettivo o, in alternativa all’articolo, un determinante e, infine, con il ruolo di soggetto può esservi un sostantivo di qualunque tipo o un pronome di qualunque tipo:

⁶¹ L’uso del verbo *venire* come ausiliare (esempio: viene distrutto) è limitato solo ai tempi semplici del verbo *venire* stesso (esempio: viene/veniva/venga/venisse distrutto) e non ai suoi tempi composti (è ~~venuto~~ distrutto) (Sabatini 2004: 8-9).

```
[postag="CS" & word!="[A-Z].*"] [postag="V.*"&
feats!=".*mod=f.*"] {1,2} [postag="R.*|A.*|D.*"] {0,2}
[postag="P.*|S.*"& deprel="subj"]
```

8) impiego della frase scissa: il fenomeno della frase scissa può realizzarsi secondo due tipologie: 1) verbo *essere* + soggetto + *che* + verbo coniugato (esempio: *è lui che parla*); 2) verbo *essere* + soggetto + *a* + verbo all'infinito (esempio: *è lui a parlare*). Sebbene in entrambe le versioni della lista dei tratti dell'italiano dell'uso medio Sabatini (1985: 163, 1990: 98) vi sia solo la tipologia della frase scissa costituita dal verbo *essere* + [altro elemento grammaticale] + *che* e manchi, invece, la tipologia basata su *a* + verbo all'infinito, nell'impostazione della ricerca su corpus si è scelto di tenere conto di entrambe le tipologie e dei vari costituenti della frase che possono essere inseriti nel focus della scissa, ossia: il gruppo nominale, il gruppo nominale preceduto da preposizione (sintagma preposizionale), l'avverbio e il verbo nella sua forma nominale, ossia all'infinito, nel caso della costruzione della frase scissa con *che* e il solo gruppo nominale nel caso della costruzione della frase scissa con *a* + infinito. Pertanto sono state progettate due stringhe di ricerca:

8a) verbo *essere* + gruppo nominale + *che* + verbo coniugato: è stato imposto che la stringa inizi con un segno di punteggiatura e/o una congiunzione per evitare che il verbo *essere* sia preceduto da un altro gruppo nominale di cui costituirebbe la copula. Esempi: *Questa è la donna che ha telefonato*, ove il verbo *essere* funge da copula vs la frase scissa: *Poiché/Ma/Ed è la donna che ha telefonato* (e non qualcun altro). A seguire vi è un verbo che rientra nelle voci del lemma del verbo *essere*, dopo il quale possono essere presenti in quantità compresa tra 0 e 2 un articolo determinativo o indeterminativo e/o un aggettivo o, in alternativa all'articolo, un determinante; vi è poi qualunque tipo di pronome o sostantivo, seguito da *che*. Esempi: *è Anna che parla*; *è Firenze che ospita la mostra* (verbo *essere* + qualsiasi tipo di sostantivo, sia esso nome proprio o nome comune); *sono loro che parlano* (verbo *essere* + pronome personale); *è questo/quello che parla* (verbo *essere* + pronome dimostrativo); *è mio padre che parla* (verbo *essere* + aggettivo possessivo + sostantivo); *è un/l'amico che parla* (verbo *essere* + articolo determinativo/indeterminativo + sostantivo); *è il/un mio amico che parla* (verbo *essere* + articolo determinativo/indeterminativo + aggettivo possessivo + sostantivo); *è questo/quel mio amico che parla* (verbo *essere* + aggettivo dimostrativo + aggettivo possessivo + sostantivo). Segue la stringa:

```
[postag="F.*|C.*"] [lemma="essere"]  
[postag="R.*|A.*|D.*"] {0,2} [postag="P.*|S.*"] [word="che"]  
[postag="V.*"]
```

8b) verbo *essere* + verbo all'infinito + *che* + verbo coniugato: è stato imposto che la stringa inizi con un segno di punteggiatura o una congiunzione; a seguire vi è un verbo che rientra nelle voci del lemma del verbo *essere*, dopo del quale vi è un verbo al modo infinito, seguito da *che* e da qualunque verbo coniugato. Esempio: *è decidere che è fondamentale*. Segue la stringa:

```
[postag="F.*|C.*"] [lemma="essere"] [postag="V.*"&  
feats=".*mod=f.*"] [word="che"] [postag="V.*"]
```

8c) verbo *essere* + avverbio + *che* + verbo coniugato: valgono le stesse considerazioni fatte per la stringa precedente, a differenza della quale il costituente della frase che è stato dislocato è l'avverbio. Esempi: *è velocemente che dobbiamo andare*; *è qui che dovete stare*. Segue la stringa:

```
[postag="F.*|C.*"] [lemma="essere"] [postag="B.*"]  
[word="che"] [postag="V.*"]
```

8d) verbo *essere* + gruppo nominale preceduto da preposizione + *che* + verbo coniugato: in questo caso il verbo che rientra nelle voci del lemma del verbo *essere* è seguito da qualunque tipo di preposizione, dopo della quale possono essere presenti in quantità compresa tra 0 e 2 un aggettivo o un determinante; infine, vi è qualunque tipo di pronome o sostantivo seguito da *che* e da un verbo coniugato. Di seguito verrà fornito un esempio per ciascuna delle casistiche reperibili tramite questa stringa:

1. verbo *essere* seguito da una preposizione, un sostantivo, *che* e un verbo: *è a casa che dobbiamo entrare*;
2. verbo *essere* seguito da una preposizione, un aggettivo, un sostantivo, *che* e un verbo: *è nella mia casa che dobbiamo entrare* (aggettivo possessivo); *è nella grande casa che dobbiamo entrare* (aggettivo qualificativo);
3. verbo *essere* seguito da una preposizione, un determinante, un sostantivo, *che* e un verbo: *è in questa casa che dobbiamo entrare*; *è in quella casa che dobbiamo entrare*;
4. verbo *essere* seguito da una preposizione, un determinante, un aggettivo, un sostantivo, *che* e un verbo: *è in questa grande casa che dobbiamo entrare*.

Segue la stringa:

```
[postag="F.*|C.*"] [lemma="essere"] [postag="E.*"] [postag="A.*|  
D.*"] {0,2} [postag="P.*|S.*"] [word="che"] [postag="V.*"]
```

8e) verbo *essere* + gruppo nominale + a + verbo al modo infinito: è stato imposto che la stringa inizi con un segno di punteggiatura e/o una congiunzione coordinante; a seguire vi è un verbo che rientra nelle voci del lemma del verbo *essere*, dopo il quale possono essere presenti in quantità compresa tra 0 e 2 un articolo determinativo o indeterminativo e/o un aggettivo o, in alternativa all'articolo, un determinante; poi vi è qualunque tipo di pronome o sostantivo seguito dalla preposizione *a* e da un verbo al modo infinito. Esempi: *è Anna a parlare*; *è Firenze a ospitare la mostra* (verbo *essere* + qualsiasi tipo di sostantivo, sia esso nome proprio o nome comune); *sono loro a parlare* (verbo *essere* + pronome personale); *è questo/quello a parlare* (verbo *essere* + pronome dimostrativo); *è mio padre a parlare* (verbo *essere* + aggettivo possessivo + sostantivo); *è un/l'amico a parlare* (verbo *essere* + articolo determinativo/indeterminativo + sostantivo); *è il/un mio amico a parlare* (verbo *essere* + articolo determinativo/indeterminativo + aggettivo possessivo + sostantivo); *è questo/quel mio amico a parlare* (verbo *essere* + aggettivo dimostrativo + aggettivo possessivo + sostantivo). Segue la stringa:

```
[postag="F.*|C.*"] [lemma="essere"] [postag="R.*|A.*|D.*"] {0,2}
[postag="P.*|S.*"] [word="a"] [postag="V.*" &
feats=".*mod=f.*"]
```

9) impiego del verbo in forma pronominale (per indicare partecipazione affettiva) (esempi: *mi mangio la pizza*, *mi sono bevuto il caffè*). Per poter ricercare nel corpus questo fenomeno grammaticale è stato necessario suddividere la progettazione delle stringhe di ricerca in due macroaree, ossia l'impiego del verbo in forma pronominale con i tempi non composti (esempio: *mi bevo/ berrò/ bevvi* (ecc.) *il caffè*) e composti (esempio: *mi sono bevuto/ ero bevuto/ mi sarei bevuto* (ecc..) *il caffè*) dei principali “verbi con uso intensivo, (o verbi di affetto)” (Ježek 2011⁶²), quali: bersi (*mi bevo una birra*), farsi (*mi faccio una mangiata*), guardarsi (*mi guardo un film*), leggersi (*mi leggo un romanzo*), mangiarsi (*mi mangio un panino*) e vedersi (*mi vedo la partita*). Pertanto, sono state progettate 12 stringhe di ricerca: 6 stringhe corrispondenti alle 6 forme pronominali riflessive (*mi, ti, si, ci, vi, si*) in presenza di tempi non composti dei verbi (esempi: *mi bevo/berrò/bevvi* (ecc.) *il caffè*) e altre 6 stringhe per i casi dei tempi composti dei verbi (esempi: *mi sono bevuto/ero bevuto/ mi sarei bevuto* (ecc.) *il caffè*). Di seguito si riportano con finalità di esempio due stringhe, rinviando all'appendice 2 per una visualizzazione completa del fenomeno rispetto a tutte le forme pronominali riflessive.

⁶² Versione html, pertanto priva di pagine.

Stringa per la prima persona singolare con i tempi non composti dei verbi:

```
[word="( ?i)mi" ] [lemma="mangiare|leggere|bere|fare|guardare|vedere"& feats=".*num=s.*" & feats=".*per=1.*" ] [postag="R.*|A.*|D.*" ] {0,2} [postag="P.*|S.*" &deprel="obj" ]
```

è stato imposto che la stringa inizi con il pronome personale (in questo caso *mi*) scritto con lettera maiuscola o minuscola; successivamente vi sono le voci di lemmi dei verbi *mangiare, leggere, bere, fare, guardare, vedere* coniugate alla prima persona singolare. A seguire è stata imposta la presenza del gruppo nominale: possono essere presenti in quantità compresa tra 0 e 2 un articolo determinativo o indeterminativo e/o un aggettivo o, in alternativa all'articolo, un determinante e, infine, con il ruolo di oggetto ("obj") può esservi un sostantivo di qualunque tipo o un pronome di qualunque tipo, quest'ultimo inserito per reperire anche casi come *mi mangio questo; mi bevo tutto*, eccetera.

Stringa per la prima persona singolare con i tempi composti dei verbi:

```
[word="( ?i)mi" ] [lemma="essere"& feats=".*num=s.*" & feats=".*per=1.*" ] [lemma="mangiare|leggere|bere|fare|guardare|vedere"& feats=".*mod=p.*" ] [postag="R.*|A.*|D.*" ] {0,2} [postag="P.*|S.*" &deprel="obj" ]
```

Valgono le stesse considerazioni fatte per la stringa precedente, a differenza della quale si è imposto che il verbo sia composto dalle voci del lemma del verbo *essere* alla prima persona singolare, seguite dal participio passato dei verbi sopra nominati, e che abbia come oggetto il gruppo nominale precedentemente descritto. La scelta di imporre in entrambe le stringhe che dopo il verbo vi sia il gruppo nominale con valore di oggetto è nata dall'esigenze di restringere il campo della ricerca: in una precedente versione delle stringhe, che terminavano con il solo verbo (a seconda dei casi, tempo semplice o composto) la ricerca restituiva molte occorrenze che non rientravano nell'ambito dell'uso del verbo in forma pronominale per indicare partecipazione affettiva, come ad esempio: *Mi faccio (del) male / mi sono fatto (del) male, non mi faccio niente/ non mi sono fatto niente, mi sono fatto/visto + infinito (mi sono fatto accompagnare; mi sono visto rimproverare), mi faccio/vedo + infinito (mi faccio prendere; mi vedo arrivare), mi sono visto + participio (mi sono visto costretto), mi sono guardato da + articolo + infinito (mi sono guardato dal disfare), mi guardo da + articolo + infinito (mi guardo dal disfare), mi sono guardato*

(tutt')attorno/intorno, eccetera. Pertanto, allo scopo di ridurre il “rumore” tra i risultati, si è preferito restringere la ricerca.

10) impiego della concordanza a senso tra il soggetto collettivo singolare seguito da un'espressione partitiva e il predicato plurale (esempio: *un milione di elettori non hanno votato*): per questo fenomeno è stata progettata una stringa che prevede un sostantivo singolare a cui seguono la preposizione *di*, un sostantivo plurale e, infine, un verbo di cui si considerano solo le persone plurali del lemma e il cui modo sia diverso dal participio. L'esclusione del participio permette di eliminare dalla ricerca tutte le occorrenze simili a *un centinaio di prodotti venduti*, che non rientrano nella casistica della concordanza a senso tra un soggetto singolare e un predicato plurale, poiché in casi di questo tipo il participio semplicemente funge da aggettivo. Ciò permette di ottenere casi come: *un centinaio di persone hanno partecipato*. Segue la stringa:

```
[postag="S.*" & feats=".*num=s.*"] [word="di"] [postag="S.*" & feats=".*num=p.*"] [postag="V.*" & feats=".*num=p.*" & feats!=".*mod=p.*"]
```

11) impiego del *ci* attualizzante con *avere* (*non c'ho tempo; non c'ho voglia*): è stato necessario elaborare 4 stringhe di ricerca al fine di includere tutte le possibili realizzazioni del fenomeno:

11a) “C’ + verbo *avere* + un sostantivo” per reperire casi come *c'ho tempo/voglia*. Segue la stringa:

```
[word="( ?i) c' " ] [lemma="avere"] [postag="S.*"]
```

11b) “Ci + verbo *avere* + un sostantivo” per reperire casi come *ci ho tempo/voglia*. Segue la stringa:

```
[word="( ?i) ci " ] [lemma="avere"] [postag="S.*"]
```

11c) “C’ + verbo *avere* + preposizione *da* + verbo” per reperire casi come *c'ho da fare*. Segue la stringa:

```
[word="( ?i) c' " ] [lemma="avere"] [word="da"] [postag="V.*"]
```

11d) “Ci + verbo *avere* + preposizione *da* + verbo” per reperire casi come *ci ho da fare*. Segue la stringa:

```
[word="( ?i) ci " ] [lemma="avere"] [word="da"] [postag="V.*"]
```

12) impiego del *che* polivalente: per poter inquadrare questo fenomeno grammaticale sono state utilizzate le stringhe progettate da Budassi (2017) con l'obiettivo di restringere il campo di tutti gli usi di *che* ed avvicinarsi il più possibile all'uso del *che* polivalente come forma generica dei pronomi relativi o delle congiunzioni subordinanti. In questa ricerca

sono state utilizzate le stringhe progettate da Budassi, dopo aver apportato alcune modifiche necessarie per uniformarle al *tagset* in uso per la consultazione del corpus PAISÀ⁶³. Nella seguente descrizione si riutilizzerà la terminologia impiegata da Budassi (2017) nella definizione delle diverse tipologie di *che* polivalente:

12a) *Che* esplicativo/consecutivo; *che* causale: come espresso da Budassi (2017:160), “i ‘che’ polivalenti di tipo esplicativo consecutivo (*Vieni che ti pettino*) e causale (*Vai a dormire che ne hai bisogno*) esprimono un legame tra due frasi”. La stringa si basa sull’esclusione dei verbi che introducono una frase dichiarativa o una completiva:

```
[postag="V.*" &
lemma!="ricordare|dimostrare|raccontare|parere|sembrare|sentire|dire|confidare|convenire|sperare|vedere|volere|sapere|scrivere|rispondere|scoprire|ribadire|promettere|mettere|insistere|dichiarare|osservare|sostenere|immaginare|temere|ritenere|premettere|augurare|pensare|pretendere|calcolare|lasciare|ripetere|scommettere|riconoscere|significare|vedere|prevedere|constatare|supporre|importare|capire|trovare|decidere|precisare|aggiungere|ammettere|notare|credere|chiedere|aspettare|considerare|bisognare|occorrere|avvenire|bastare|essere"] [word="che"] [postag="V.*"]
```

La stringa consente, indubbiamente, di restringere il campo d’indagine, ma si è rivelata poco produttiva, perché tra i risultati reperiti nel corpus si trovano soprattutto le frasi oggettive e le relative standard. Tra gli esempi reperiti figurano: *vieni che ti spalmo la crema* (esplicativo-consecutivo); *potevano fare tutte le guerre che volevano che a lui gli facevano un baffo* (causale).

12b) *Che* consecutivo-presentativo (con avverbi di dubbio): l’esempio inizialmente fornito per questa categoria da Budassi (2017), a sua volta tratto da Sornicola (1981), è il seguente: *Io sono una donna tranquilla che sto in casa, lavoro*. Casi come questo si collocano nel *continuum* esistente tra le due categorie di *che* polivalente: il *che* introduttore generico di frase subordinata, giustappositivo, e il *che* introduttore di frasi relative restrittive (Budassi 2017: 155-156). La progettazione della stringa di ricerca per la tipologia “consecutivo-presentativo” si è svolta attraverso varie fasi di raffinazione (cfr. ivi: 162-164), che conducono alla seguente riflessione:

⁶³ Per le differenze tra le versioni originali e quelle impiegate in questa ricerca si rinvia a Budassi (2017: 147-172).

[...] la percentuale di polivalenza del ‘che’ rintracciabile con la [versione precedente della stringa] sale vertiginosamente se, nello slot immediatamente successivo a [tag="CHE"], vengono ricercati alcuni tipi di avverbi, la maggior parte dei quali appartiene al gruppo degli avverbi di dubbio. (ivi: 163)

Budassi è, pertanto, approdato ad una stringa che ha come elemento iniziale un segno di punteggiatura, per delimitare la fine della frase precedente; segue la parola *che*, seguita da alcuni avverbi di dubbio, da eventuali altri elementi linguistici nel numero di 0 o 1 e, infine, da un verbo. La stringa è stata riutilizzata in questa ricerca, apportando delle modifiche necessarie per uniformarla al *tagset* in uso per la consultazione del corpus PAISÀ (esempio: [word="che"] vs [tag="CHE"] dell'originale) ed impostando che la parola *che* possa iniziare con lettera maiuscola o minuscola (ossia, (?i)) a seconda che sia ad inizio frase o meno:

```
[word="\.|\;|\!|\?|\, " ] [word="( ?i) che " ]  
[word="se|forse|chissà|eventualmente|probabilmente|poi|magari"] [] {0,1} [postag="v.*"]
```

Ciò ha consentito di reperire nel corpus PAISÀ casi come: *Io mi rendo conto di una cosa. Che se devo fare questa vita per altri dieci anni, una vita da trottola, girando tutto il giorno da un posto all' altro, rischio di sfinirme.* (sic.)

12c) *Che* pseudorelativo: l'esempio inizialmente fornito per questa categoria da Budassi (2017: 155), a sua volta tratto da Berruto (2012), è il seguente: *Li vedo che scendono.*

[Questa frase] è molto utile per individuare la struttura all'interno della quale potrebbe essere rintracciato il che polivalente pseudorelativo. Essa potrebbe essere la seguente: Pn (pronome) – V1– che – V2, dove V2 è concordato con Pn. Se anche il ‘che’ sembra svolgere funzione di soggetto relativo di V2, il suo valore è primariamente giustappositivo (proprio in valore di ciò è detto pseudorelativo). (ivi, p. 164)

La stringa elaborata da Budassi prevede un pronome seguito dalle voci del lemma dei verbi *vedere* e *sentire*, seguite da *che* e da un verbo ed è stata presa come modello per l'elaborazione di una nuova stringa che include i principali verbi di percezione e che è strutturata come segue: un pronome seguito dalle voci del lemma dei verbi *vedere*, *sentire*, *guardare*, *scorgere*, *notare*, *osservare*, *mirare*, *ascoltare*, *udire*, *intendere*, *percepire*, seguite da *che* e da un verbo. Ossia:


```
[postag="P.*"]
[lemma="vedere|sentire|guardare|scorgere|notare|osservare|
mirare|ascoltare|udire|intendere|percepire"] [word="che"] [p
ostag="V.*"]
```

Tuttavia, come già anticipato nella sezione 4.6.1, tra i risultati ottenuti dall'uso di questa stringa vi erano molte occorrenze del *si* impersonale (esempio: *si vede che hanno qualche domanda*), che chiaramente non rientrano nel fenomeno indagato. La struttura indicata da Budassi (*ibidem*), infatti, prevede la presenza di un pronome in funzione di oggetto, seguito dal verbo principale (V1), *che* e da un secondo verbo (V2), che concorda con il pronome. È il caso del già citato *Li vedo che scendono*, ove il pronome *Li* concorda con *scendono* (V2). Da tale descrizione risulta evidente che i casi di *si* impersonale (esempio: *si vede che hanno qualche domanda*) non rientrino all'interno di questa struttura sintattica, poiché il pronome *si* concorda con verbo principale (Esempio: *si vede che hanno qualche domanda*, il pronome *si* concorda con *vede*).

Pertanto, si è scelto di escludere il pronome *si* dai pronomi ricercabili. Ciò ha premesso di eliminare le occorrenze del *si* impersonale e di reperire occorrenze sul modello di *Li vedo che scendono*, trovando un compromesso tra la perdita di occorrenze corrette e l'esclusione di casi non inerenti, come già anticipato nella sezione 4.6.1, dedicata all'uso di CQL. Segue la stringa modificata:

```
[postag="P.*"&word!="si"]
[lemma="vedere|sentire|guardare|scorgere|notare|osservare|
mirare|ascoltare|udire|intendere|percepire"] [word="che"] [p
ostag="V.*"]
```

13) impiego dell'indicativo al posto del congiuntivo in alcune subordinate (dopo i verbi di opinione, dopo una dichiarativa negativa (*non credo che hai torto*), in un'interrogativa indiretta (*gli chiesi se poteva aiutarmi*) e nelle ipotetiche (*se me lo dicevi, ci pensavo io*) (cfr. Mazzoleni 1992). La progettazione delle stringhe è avvenuta sulla base delle seguenti casistiche:

13a) dopo i verbi di opinione/sentimento: per questo fenomeno è stata progettata una stringa che prevede le voci dei lemmi dei principali verbi che esprimono opinione o sentimento siano seguite da *che*, da eventuali altri elementi linguistici in quantità variabile da 0 a 5, purché diversi dal punto (`[word!="\."]`) e, infine, da un verbo al modo indicativo (esempi: *penso che è giusto; mi dispiace che questa cosa è successa*). L'esclusione del punto dagli eventuali elementi linguistici presenti dopo *che* è un

espediente adottato per evitare che, come era accaduto in una prima fase di progettazione della stringa, la ricerca restituisca anche casi in cui il verbo all'indicativo si trovi dopo il punto, ossia al di fuori della frase retta da verbi di opinione/sentimento (esempio: *Penso che questo sia giusto. Insisto.*, ove *penso* è il verbo di opinione + *che* + 4 elementi linguistici, tra cui un verbo al congiuntivo e il punto, che verrebbero percepiti dal software come possibili elementi da includere, e infine un verbo al modo indicativo, che regge un'altra frase.) Segue la stringa:

```
[lemma="pensare|credere|ritenere|sembrare|parere|supporre|dubitare|immaginare|temere|sperare|dispiacersi|stupirsi|augurarsi|aspettarsi|volere"] [word="che"] [word!="\."]{0,5} [postag="V.*" & feats=".*mod=i.*"]
```

13b) dopo i verbi di opinione/sentimento preceduti da negazione: valgono le stesse considerazioni fatte per la stringa precedente, a differenza della quale l'elemento iniziale è la negazione *non* (esempio: *Non penso che è giusto*). Segue la stringa:

```
[word="non"]
[lemma="pensare|credere|ritenere|sembrare|parere|supporre|dubitare|immaginare|temere|sperare|dispiacersi|stupirsi|augurarsi|aspettarsi|volere|sapere"] [word="che"] [word!="\."]{0,5} [postag="V.*" & feats=".*mod=i.*"]
```

13c) dopo la costruzione “verbo essere + aggettivo + che”: questo fenomeno include sia i casi di costruzioni impersonali (esempio: *è necessario che*), sia casi di espressioni di sentimento (esempio: *sono contenta che*). Pertanto, è stata progettata una stringa in cui le voci del lemma del verbo *essere* sono seguite da un aggettivo, da *che*, da eventuali altri elementi linguistici in quantità variabile da 0 a 5, purché diversi dal punto ([word!="\."]) e, infine, seguite da un verbo al modo indicativo:

```
[lemma="essere"] [cpostag="A"] [word="che"] [word!="\."]{0,5}
[postag="V.*" & feats=".*mod=i.*"]
```

13d) dopo la costruzione “verbo avere + paura/timore + che”: valgono le stesse considerazioni fatte per la stringa precedente, a differenza della quale l'elemento iniziale è il lemma del verbo *avere*, seguito dalle parole *paura* o *timore* (esempio: *ho paura che non ce la fai*). Segue la stringa:

```
[lemma="avere"] [word="paura|timore"] [word="che"] [word!="\."]{0,5} [postag="V.*" & feats=".*mod=i.*"]
```

13e) nell'interrogativa indiretta: per questo fenomeno è stata impostata una stringa in cui i lemmi dei verbi *chiedere* e *domandare* sono seguiti da *se*, da eventuali altri elementi linguistici in quantità variabile da 0 a 5, purché diversi dal punto (`[word!="\."]`) e, infine, seguiti da un verbo al modo indicativo:

```
[lemma="chiedere|domandare"] [word="se"] [word!="\."]{0,5} [postag="V.*" & feats=".*mod=i.*"]
```

13f) nelle ipotetiche: per questo fenomeno, che esprime l'ipotesi irreali, è stata impostata una stringa in cui si richiede che la congiunzione subordinante *se*, scritta con la lettera maiuscola o minuscola, possa essere seguita da eventuali altri elementi linguistici in quantità variabile da 0 a 5, purché diversi da *ne* e *stesso/stessa/stessi/stesse* (in modo da evitare costruzioni come *se ne* e *se stesso/stessa/stessi/stesse*, che erano comunque emerse tra i risultati di una precedente versione della stringa). Seguono poi un verbo al modo indicativo e al tempo imperfetto (*i*), eventuali altri elementi linguistici in quantità variabile da 0 a 5, purché diversi dal punto, e infine un altro verbo con le stesse caratteristiche sopra nominate. La stringa permette di individuare esempi assimilabili a *se lo sapevo*, *te lo dicevo*, in cui un'ipotesi non più realizzabile è espressa tramite l'indicativo imperfetto e non tramite la combinazione di congiuntivo trapassato e condizionale passato (*se l'avessi saputo*, *te l'avrei detto*):

```
[postag="CS" & word="( ?i) se" ] [word!="ne|stess(o|a|i|e)"] {0,5} [postag="V.*" & feats=".*mod=i.*" & feats=".*ten=[i].*" ] [word!="\." ] {0,5} [postag="V.*" & feats=".*mod=i.*" & feats=".*ten=[i].*" ]
```

Le stringhe finora descritte sono state utilizzate per l'interrogazione del *corpus* PAISÀ, che verrà descritta nel capitolo 5. Di seguito verrà data una breve descrizione della metodologia che è stata adottata nel salvataggio delle occorrenze ottenute dall'interrogazione del *corpus* al fine di ottenere dati affidabili.

4.7 Metodo per il salvataggio e il trattamento dei risultati dell'interrogazione del corpus

Le stringhe di ricerca precedentemente descritte sono state utilizzate per interrogare il *corpus* PAISÀ. Allo scopo di ottenere dati di frequenza affidabili, nella fase di salvataggio dei risultati dell'interrogazione del corpus è stata svolta una quantificazione basata su un

campione scelto in modo casuale e poi ripulito manualmente. Nello specifico, dopo aver ottenuto le occorrenze dei fenomeni ricercati, si è proceduto al salvataggio in formato .txt dei risultati e alla creazione, per ciascun fenomeno, di tre campioni da 200 occorrenze l'uno, per un totale di 600 occorrenze da analizzare per fenomeno⁶⁴. La scelta di fissare il limite a 600 risponde principalmente alla necessità di disporre di un numero gestibile di occorrenze da visualizzare, dato l'elevato numero di stringhe di ricerca elaborate, nonché di avere un totale di partenza che consentisse di dividere i risultati in tre campioni uguali⁶⁵. Al fine di ottenere dei campioni casuali, ossia che non seguissero l'ordine delle occorrenze fornito dal *corpus*, prima del salvataggio dei risultati è stata selezionata la funzione *shuffle*, che consente di rimescolare in ordine casuale i risultati ottenuti, evitando in tal modo anche che si verificassero eventuali raggruppamenti dei doppioni.

Da ciascun campione sono stati eliminati i doppioni e i casi non corrispondenti al fenomeno oggetto, tra cui vi sono sia casi contenenti errori di *tagging*, sia casi in cui il *tagging* è corretto, ma è la stringa di ricerca a trovare anche risultati non inerenti al caso in esame. Per citare alcuni esempi, alcuni errori di *tagging* sono stati rilevati nel caso del soggetto in posizione rematica rispetto a verbi composti con ausiliare *avere*, con l'oggetto diretto tra participio e soggetto (esempio: *ha suscitato sdegno l'articolo che racconta...*, ove *l'articolo* è il soggetto e *sdegno* è l'oggetto diretto posto tra il participio e il soggetto): tra i risultati della ricerca sono presenti anche casi in cui è stato taggato erroneamente come soggetto qualunque elemento linguistico di carattere nominale successivo al verbo (esempio: *ha fatto i compiti la bambina*, ove *i compiti* è erroneamente taggato come soggetto). Invece, nel caso del verbo in forma pronominale (esempio: *mi mangio il panino*) alla terza persona singolare e plurale sono stati reperiti numerosissimi casi di strutture impersonali (esempi: *si mangia bene qui*; *si mangiano i calamari*), che chiaramente non rientrano nel caso in questione.

Dopo aver dato una breve descrizione della metodologia adottata nel salvataggio delle occorrenze ottenute dall'interrogazione del *corpus*, nella sezione 4.8 verrà illustrato il metodo adottato per la progettazione e sperimentazione delle unità didattiche.

⁶⁴ Allo scopo di mettere da parte un ampio campionario da cui eventualmente attingere ulteriori occorrenze, nel caso in cui le prime 600 non avessero fornito risultati sufficienti, nella fase di salvataggio dei dati dal *corpus* sono state scaricate in formato .txt le prime 2000 occorrenze del totale fornito dal *corpus*; da queste sono state estratte le prime 600, che sono state suddivise nei tre campioni da 200 a cui si fa riferimento nel testo.

⁶⁵ Nei casi in cui il totale delle occorrenze reperite è risultato inferiore a 600 e maggiore di 60, il totale è stato comunque diviso in tre campioni identici o che presentassero uno scarto di 1-2 unità di differenza; nei casi in cui il totale delle occorrenze ottenute è risultato invece inferiore a 60, i risultati sono stati trattati come un campione unico.

4.8 Metodo per la progettazione e sperimentazione delle unità didattiche

Richiamando gli obiettivi principali di tale progetto di ricerca, enunciati in apertura del capitolo, la selezione del corpus PAISÀ tra le varie risorse attualmente disponibili è stata svolta con l'obiettivo di reperire dei testi contenenti tali fenomeni grammaticali, con i quali progettare delle unità didattiche per l'apprendimento dell'italiano L2.

A tale scopo, le stringhe di ricerca precedentemente illustrate sono state utilizzate per ricercare nel corpus i fenomeni grammaticali oggetto di studio, scaricare dei campioni dei risultati ottenuti ed esaminarli con il duplice obiettivo sia di ottenere degli esempi dell'uso reale di tali strutture in testi autentici, sia di calcolare la ricorrenza dei fenomeni ricercati. Il procedimento di interrogazione del corpus, esame dei risultati e calcolo della frequenza sarà dettagliatamente descritto nel capitolo 5; nelle sezioni seguenti, invece, si illustreranno le attività finalizzate alla produzione delle unità didattiche, ossia: la selezione dei testi, la progettazione delle unità didattiche e la selezione di una piattaforma *e-learning* sulla quale collocare le suddette unità didattiche.

4.8.1. Selezione dei testi

Dopo aver individuato quali tra i fenomeni grammaticali elencati da Sabatini (1990) siano i più frequenti nel corpus (cfr. capitolo 5), ne sono stati selezionati due sui quali incentrare le unità didattiche: l'uso dell'indicativo al posto del congiuntivo e l'uso del verbo in forma pronominale riflessiva in senso affettivo (sulle motivazioni della scelta di tali fenomeni, si rinvia ai capitoli 5 e 6).

Al fine di selezionare i testi da utilizzare nella progettazione delle unità didattiche, innanzitutto è stata svolta una rilettura delle occorrenze scaricate dal corpus PAISÀ per i due fenomeni grammaticali. Poiché il corpus permette di visualizzare il contesto testuale degli elementi linguistici ricercati, per ciascuna occorrenza del fenomeno oggetto di ricerca è stata attivata la visualizzazione di alcune righe di testo precedenti e successive alla posizione nella frase dell'elemento ricercato.

Tale procedura ha permesso di identificare in ciascuno dei campioni di occorrenze i casi ritenuti più interessanti sia dal punto di vista grammaticale, poiché il loro utilizzo nell'unità didattica avrebbe permesso di approfondire aspetti particolari dell'uso di un determinato fenomeno linguistico, sia dal punto di vista tematico, poiché nella selezione dei testi si è tenuto conto anche del loro contenuto, avendo cura di selezionare temi generalmente trattati nella manualistica per stranieri rispetto ai livelli intermedi B1-B2 (cfr.

capitolo 6) e che, pertanto, potessero risultare familiari agli apprendenti e, soprattutto, adeguati alle loro conoscenze e competenze linguistiche. Per quanto riguarda questi due aspetti, la riflessione è stata guidata dalla consultazione dei seguenti testi: 1) il QCER (Consiglio d'Europa 2018, 2002); 2) i sillabi della lingua italiana attualmente disponibili (cfr. capitolo 6); 3) i contenuti tematici trattati nei 32 manuali di italianoL2/LS citati nello studio di Ślapek (2017) sulla manualistica per stranieri attualmente in commercio.

La consultazione delle risorse citate ai punti 1 e 2 ha consentito di avere un quadro completo degli argomenti grammaticali da inserire nelle unità didattiche (1) e della loro collocazione all'interno dei vari livelli previsti al QCER (2): ad esempio alcuni argomenti grammaticali sono caratterizzanti di un determinato livello e possono essere ripresi come approfondimento nei livelli successivi; la manualistica attualmente in commercio ha permesso di determinare quali tematiche vengano generalmente trattate nei livelli intermedi.

Una volta identificati i casi ritenuti più interessanti dal punto di vista grammaticale e tematico per ciascuno di essi sono state visitate le rispettive pagine web originarie, così da poter leggere i testi integralmente e in modo più agevole. Ciò è stato reso possibile dal fatto che per ciascuno dei testi che compongono il corpus PAISÀ sono indicate tra le informazioni generali anche le pagine web da cui i testi sono stati scaricati.

Infine, una volta avuta una panoramica del contesto più ampio e/o dell'intero testo in cui era utilizzato un dato fenomeno grammaticale, per ciascun caso, in base alle specificità grammaticali o tematiche del testo in questione, si è valutato se per la progettazione delle unità fosse più opportuno utilizzare i testi in forma integrale o parziale (ad esempio, sotto forma di brani da leggere), o se utilizzare solo le singole frasi contenenti il fenomeno in questione (ad esempio, per creare degli esercizi di completamento o delle liste, ecc.). Quest'ultima opzione è stata scelta nel caso di testi risultati inadatti al livello linguistico dell'unità (ad esempio, dal punto di vista del registro linguistico, del tema trattato, dei riferimenti culturali presenti ecc.).

4.8.2. Valutazione di alcune piattaforme di *e-learning*

Parallelamente all'elaborazione delle unità didattiche si è svolta la ricerca di una piattaforma dedicata all'apprendimento in formato elettronico in cui collocare i materiali progettati. L'*e-learning* è un insieme di metodologie e strategie didattiche che vengono impiegate per la progettazione di un ambiente di apprendimento basato sull'uso del web e

della multimedialità. Una sua possibile definizione è quella elaborata dall'osservatorio Anee (Associazione nazionale dell'editoria elettronica, dal 2001 Associazione dei Servizi e Contenuti Multimediali):

L'e-learning è una metodologia di insegnamento e apprendimento che coinvolge sia il prodotto sia il processo formativo. Per prodotto formativo si intende ogni tipologia di materiale o contenuto messo a disposizione in formato digitale attraverso supporti informatici o di rete. Per processo formativo si intende invece la gestione dell'intero iter didattico che coinvolge gli aspetti di erogazione, fruizione, interazione, valutazione. In questa dimensione il vero valore aggiunto dell'e-learning emerge nei servizi di assistenza e tutorship, nelle modalità di interazione sincrona e asincrona, di condivisione e collaborazione a livello di Community. Peculiarità dell'e-learning è l'alta flessibilità garantita al discente dalla reperibilità sempre e ovunque dei contenuti formativi, che gli permette l'autogestione e l'autodeterminazione del proprio apprendimento. (Liscia 2004: 12)

Si tratta, dunque, di una metodologia che propone un'innovazione nella concezione e progettazione dei contenuti formativi, nella loro organizzazione e fruizione da parte degli utenti. Tra le risorse utilizzate per l'*e-learning* vi sono le piattaforme specificamente progettate per gestire attività didattiche online, che si basano su due sistemi: LMS (*Learning Management System*) e LCMS (*Learning Content Management System*), ove il primo gestisce la distribuzione dei corsi online, l'iscrizione degli apprendenti e il tracciamento delle loro attività online, mentre il secondo è dedicato alla gestione dei contenuti. I due sistemi spesso cooperano per la realizzazione di un ambiente di apprendimento online.

I criteri che hanno guidato la selezione della risorsa *e-learning* da utilizzare per la presente ricerca sono:

- a. la struttura della piattaforma (struttura e grafica dell'interfaccia, tipologie di attività didattiche progettabili, la possibilità di caricare/scaricare contenuti multimediali, quali audio/ video, immagini, creare collegamenti ad altri siti web ecc.);
- b. l'accessibilità (sia dal punto di vista tecnico, ossia la possibilità di accedere efficacemente al sito web sia da tecnologia fissa che da dispositivi mobili, quali *smartphone* e *tablet*, sia dal punto di vista economico, ossia se si tratta di risorse *open source* o a pagamento);
- c. il tipo di formazione erogata: in autoapprendimento (basata sull'offerta di contenuti e di un supporto minimale da parte di un tutor), assistita (che prevede sia momenti di

studio individuale, sia momenti di interazione con un tutor) o collaborativa (impostata sull'interdipendenza tra i membri del gruppo-classe che si realizza attraverso attività basate sulla condivisione dei contenuti e collaborazione in cui il docente ricopre il ruolo di moderatore della comunità di apprendimento);

- d. la possibilità di realizzare le attività didattiche direttamente sulla piattaforma (invece di caricarne i singoli file) e il grado di personalizzazione con cui tale progettazione può avvenire.

Innanzitutto è stata svolta un'indagine finalizzata alla raccolta di informazioni sulle piattaforme di e-learning attualmente disponibili. Sulla base delle classiche fornite da *Capterra*⁶⁶, *Academy of mine*⁶⁷, *Wikipedia*⁶⁸ risulta che le piattaforme più diffuse tra gli enti di formazione al momento della scrittura di questo capitolo siano: *A-tutor*, *Blackboard*, *Chamilo*, *Claroline*, *Docebo*, *Dokeos*, *Edmodo*, *E-Front*, *Ilias*, *Moodle*, *Sakai* e *Schoology*⁶⁹. Ciascuna risorsa è stata valutata sulla base dei criteri sopra menzionati, che hanno portato all'immediata esclusione di alcune di esse. Per quanto riguarda, ad esempio, l'accessibilità open source vs pagamento, *Blackboard* e *Docebo*, entrambe ampiamente diffuse e utilizzabili sia da tecnologia fissa che mobile, presentano costi che si aggirano intorno alle migliaia di euro l'anno; mentre invece *Dokeos* ed *E-Front* presentano versioni base con funzionalità molto ridotte rispetto alle versioni a pagamento.

Inoltre, alcune di esse sono concepite come una "libreria", ossia non consentono la progettazione di un'intera unità didattica online, bensì si basano principalmente sul caricamento da parte del docente di file contenenti le attività, i quali devono essere poi scaricati dagli studenti e ricaricati per la consegna al docente, con un evidente aggravio di tempo e perdita di immediatezza nell'uso della piattaforma. È il caso, ad esempio, della piattaforma *Moodle*, sebbene essa consenta entro centri limiti di svolgere attività direttamente online (lettura/produzione di testi in formato blog o wiki, utilizzo di file multimediali ecc.). Si è rinunciato, inoltre, all'uso delle piattaforme *Edmodo* e *Ilias* per ragioni analoghe. *Edmodo* consente di caricare file, organizzarli in cartelle e biblioteche e

⁶⁶ Capterra, LMS Software www.capterra.com/learning-management-system-software/#top-20 (consultato il 04/01/2019).

⁶⁷ Academy Of Mine, *Best Free Learning Management Systems (LMS)* www.academyofmine.com/12-best-free-learning-management-systems-lms (consultato il 04/01/2019).

⁶⁸ Wikipedia, *Learning Management System* https://it.wikipedia.org/wiki/Learning_management_system (consultato il 04/01/2019).

⁶⁹ *A-tutor* (<https://atutor.ca>), *Blackboard* (www.blackboard.com/it-it), *Chamilo* (<https://chamilo.org/it>), *Claroline* (<https://claroline.net>), *Docebo* (www.docebo.com), *Dokeos* (www.dokeos.com/it) *Edmodo* (<https://new.edmodo.com/?go2url=%2Fhome>), *E-Front* (www.efrontlearning.com), *Ilias* (www.ilias.de), *Moodle* (<http://moodle.org>), *Sakai* (www.sakailms.org) e *Schoology* (www.schoology.com) (consultati il 05/01/2019).

condividerli con Google Drive e rispetto a *Moodle* presenta il vantaggio di poter svolgere la correzione dei compiti direttamente online con una funzione simile a quella dei commenti nei documenti Word, tuttavia non consente la creazione di quiz direttamente sulla piattaforma. *Ilias*, invece, è concepita come un LMS aperto: offre un ambiente per l'organizzazione e strutturazione in corsi e classi dei contenuti, tuttavia questi ultimi sono liberamente condivisi con chiunque e non con i soli utenti del sistema.

Pur non rientrando nelle classifiche fornite dai siti sopra citati, era stata valutata anche la piattaforma italiana *Fidenia*⁷⁰ che ha un funzionamento e una grafica molto simili a *Facebook* e il cui uso, pertanto, sarebbe stato facilmente intuibile dagli apprendenti: si basa sulla pubblicazione da parte del docente di immagini, testi e collegamenti ipertestuali sotto forma di *post* commentabili dagli apprendenti, ma analogamente ad *Edmodo* non consente di creare quiz direttamente sulla piattaforma.

Al di là di tali specificità, ciò che è emerso in generale dalla documentazione disponibile online sulle risorse sopra nominate è che esse siano state concepite per la creazione e gestione di interi corsi online, piuttosto che per la creazione di un set di attività didattiche raggruppabili all'interno di un'unità didattica. Poiché, come emergerà nel capitolo 6, per le finalità di questa ricerca è stata valutata la produzione di due sole unità didattiche focalizzate su due dei fenomeni elencati da Sabatini, l'utilizzo di piattaforme che presentano una certa complessità nella gestione dei contenuti sarebbe risultato sproporzionato rispetto all'obiettivo in questione.

Di conseguenza, l'attenzione è stata rivolta verso strumenti che consentano la produzione di fogli elettronici interattivi. Consultando il sito *E-Learning industry*⁷¹ sono state valutate varie risorse di questo tipo sulla base dei criteri sopra citati ed, infine, è stata selezionata la risorsa *Wizer.me*⁷².

Definito come “in assoluto uno dei migliori strumenti per creare diverse tipologie di quiz ed esercizi interattivi online”⁷³, il fattore principale che ne ha determinato la selezione è che viene dato ampio spazio alla personalizzazione da parte del docente, sia dal punto di

⁷⁰ www.fidenia.com (consultato il 05/01/2019).

⁷¹ “It is the largest online community of eLearning professionals in the industry, and was created first and foremost as a knowledge-sharing platform to help eLearning professionals and instructional designers connect in a safe online community where they can stay up to date with the latest industry news and technologies, and find projects or jobs.” (<https://elearningindustry.com/>, consultato il 10/01/2019).

⁷² www.wizer.me (consultato il 12/01/2019).

⁷³ www.robertosconocchini.it/ (consultato il 10/01/2019) sito che raccoglie informazioni sulle risorse didattiche digitali, nato come iniziativa informale di un ricercatore non affiliato.

vista didattico, sia dal punto di vista grafico, nella progettazione delle attività all'interno della risorsa sotto forma di fogli di lavoro interattivi.

Nello specifico, si tratta di una piattaforma con registrazione gratuita che offre una vasta gamma di modelli di attività (ad esempio: inserimento di brani da leggere, creazione di domande aperte o chiuse, tabelle, esercizi di riempimento, esercizi di collegamento immagine-parola o collegamento tra parole, oltre alla possibilità di inserire collegamenti ipertestuali, come link a video o pagine web ecc.). Le consegne di ciascuna attività possono, poi, essere impostate non solo in forma scritta, ma anche sotto forma di file audio, favorendo lo svolgimento delle attività anche da parte di apprendenti con disturbi specifici di apprendimento (DSA) e, in tal modo, l'inclusività all'interno del gruppo-classe.

Inoltre, offre numerose opzioni per sfondi, caratteri del testo, colori, nonché la possibilità di stabilire manualmente il punteggio da attribuire a ciascun quesito e la possibilità di attivare la funzione di correzione automatica delle attività, previo inserimento da parte del docente delle soluzioni dei diversi quesiti in fase di progettazione. Quest'ultimo aspetto ha costituito un ulteriore punto a favore, poiché permette sia di svolgere le attività in autoapprendimento, sia di abbreviare i tempi di correzione in aula. I docenti hanno il report dettagliato di tutte le prove degli studenti cliccando su *Answers*.

La piattaforma è accessibile sia tramite desktop, sia tramite cellulare. L'accesso da parte degli apprendenti avviene tramite la condivisione da parte del docente del link del foglio di lavoro creato: una volta copiato il link nel browser, si apre una maschera per il login con profilo studente e, dopo aver effettuato la registrazione, il foglio elettronico con le attività si apre automaticamente.

Oltre che per la progettazione dei fogli elettronici, la piattaforma può essere utilizzata anche per reperire materiali progettati da altri utenti, riutilizzarli ed eventualmente personalizzarli in base ad esigenze diverse. L'accesso ai materiali da parte dell'intera community degli utenti di *Wizer.me* avviene di default. In alternativa, è possibile sottoscrivere un abbonamento annuale che consente la creazione di un account riservato: in quest'ultimo caso è l'utente che decide se condividere o meno i materiali all'interno della community. Nel caso della presente ricerca, si è optato per la sottoscrizione dell'abbonamento, al fine di poter preservare la riservatezza dei materiali didattici sperimentali.

4.8.3. Progettazione delle unità didattiche

La glottodidattica necessita di modelli operativi che permettano di realizzare le sue componenti teoriche (l'approccio) e pratiche (il metodo):

[l]'approccio costituisce la filosofia di fondo di ogni proposta glottodidattica. L'approccio valuta e seleziona dati e impianti epistemologici dalle varie teorie e dalle varie scienze di riferimento, e li riorganizza secondo i parametri propri della glottodidattica, individuando le mete e gli obiettivi dell'insegnamento linguistico. Un approccio genera uno o più metodi che ne realizzano l'applicazione nelle varie situazioni. Nella storia della glottodidattica alcuni approcci sono stati definiti "metodi" (Comunicativo, Formalistico, Metodo diretto, Metodo naturale, Reading method, Silent Way, Strutturalistico, Suggestopedia)" (Balboni 1999: 5).

Tra i vari modelli operativi esistenti (cfr. Diadori, Semplici, Troncarelli 2020) in questa trattazione si farà riferimento all'unità didattica, facendo successivamente una breve distinzione tra due delle varie tipologie di unità attualmente esistenti.

Nella seconda metà del Novecento viene gradualmente superato il modello operativo della lezione frontale nell'insegnamento delle lingue, con l'emergere di nuove teorie sull'apprendimento delle lingue. Dagli anni Settanta prende il via una revisione dei modelli di insegnamento delle lingue moderne, ispirata in particolare alla psicologia della *Gestalt* (Koffka 1935, Köhler 1929), una corrente psicologica che descrive la percezione umana secondo la sequenza "globalità → analisi → sintesi" che va dal generale al particolare e che viene utilizzata anche per avviare una nuova concezione della didattica.

Poiché la percezione umana segue questa sequenza, emerge in quegli anni l'idea che anche il contatto da parte degli apprendenti con i contenuti di una disciplina di studio possa essere proposto in maniera più efficace utilizzando un percorso che dalla globalità passi all'analisi e infine alla sintesi.

Di fondamentale importanza in questo contesto fu l'elaborazione da parte di Krashen della *Second Language Acquisition Theory* (1982), basata su sei ipotesi inerenti ai processi attraverso i quali si impara una lingua. Di seguito si farà riferimento brevemente solo alle ipotesi strettamente legate alla modalità di presentazione dei contenuti linguistici agli apprendenti.

Innanzitutto Krashen identifica due processi fondamentali attraverso i quali si impara una lingua: l'apprendimento e l'acquisizione. L'apprendimento è un processo consapevole e razionale, generalmente tipico del percorso scolastico, che può portare a competenze

linguistiche pressoché perfette, ma non durature, poiché si basa sulla memoria a breve termine. L'acquisizione, invece, è un processo inconscio, che porta ad interiorizzare la lingua in maniera duratura e alla graduale capacità di comprendere e produrre frasi riutilizzando le nozioni assorbite.

Si basa su un percorso graduale che presenta molte tappe intermedie, che costituiscono l'interlingua (Selinker 1972), ossia un sistema linguistico provvisorio, in continua evoluzione, relativo alla lingua oggetto di studio, che l'apprendente ricostruisce gradualmente, man mano che nuove regole e strutture si aggiungono alla sua competenza.

Il processo inconscio dell'acquisizione, che sfrutta l'intuizione e la memorizzazione naturale, è paragonabile alla maniera in cui il bambino impara la propria lingua madre. (ipotesi di distinzione tra acquisizione e apprendimento). Da ciò si deduce che l'insegnante tramite la didattica debba puntare a favorire l'acquisizione duratura da parte degli apprendenti e il raggiungimento di tale obiettivo chiaramente è influenzato dalla modalità di presentazione dei contenuti.

Al fine di realizzare l'acquisizione linguistica, è necessario che le regole di una lingua siano acquisite in modo spontaneo, seguendo il cosiddetto "ordine naturale di acquisizione", che rappresenta l'ordine sequenziale con cui il cervello processa i contenuti linguistici, procedendo dalle strutture più semplici alle strutture più complesse (ipotesi dell'ordine naturale).

Krashen (1982, 1985) sintetizza l'ordine naturale attraverso la formula "i+1", che descrive l'input linguistico presentato con gradualità: rispetto all'input linguistico già noto (*i*), il nuovo input fornito deve avere un grado di difficoltà leggermente maggiore (*+1*). Solo in questo modo l'input linguistico risulta comprensibile, poiché viene presentato in un momento in cui l'apprendente è in grado di coglierne il significato e l'utilità (ipotesi dell'input comprensibile).

Infine l'input verrà successivamente immagazzinato nella memoria a lungo termine solo se è stato fornito in una situazione priva di ansia e paura di essere giudicato, in cui le difese psicologiche (il cosiddetto "filtro affettivo") siano abbassate (ipotesi del filtro affettivo).

L'ordine naturale è studiato e descritto dalla linguistica acquisizionale, una disciplina che si è inizialmente sviluppata nell'area anglosassone e successivamente nei vari paesi europei che hanno vissuto il fenomeno dall'immigrazione, tra cui la Germania, la Francia e in seguito anche l'Italia, come un nuovo settore di ricerca focalizzato sull'apprendente e sullo sviluppo della sua competenza linguistico-comunicativa, soprattutto nelle L2.

Studiare il funzionamento della facoltà di linguaggio e alcuni suoi principi coinvolti nel processo di costruzione della competenza linguistica ha dei risvolti nella didattica di L2, poiché consente nella progettazione di materiali didattici e nella prassi didattica di predisporre delle strategie che facilitino l'acquisizione linguistica. La prospettiva assunta in questa ricerca coinvolge sia la linguistica acquisizionale, sia la glottodidattica, ricadendo all'interno della cosiddetta didattica acquisizionale (Vedovelli 2003, Vedovelli, Villarini, 2003).

Dando brevemente delle definizioni di tali ambiti, si può affermare che la linguistica acquisizionale “studia processi e modelli relativi all'acquisizione di una lingua non materna” (Chini, 2005: 9), focalizzando la sua attenzione sull'apprendente e sul contesto comunicativo che gli fornisce gli input linguistici, mentre la glottodidattica è la scienza dell'educazione linguistica (Balboni, 1999) che “studia l'apprendimento e l'apprendente dalla prospettiva dell'insegnamento” (Vedovelli, 2003: 179). Dall'interesse per le ricadute didattiche degli esiti della ricerca sulla L2 ha avuto origine la didattica acquisizionale, “un modello teorico di didattica linguistica che si rapporti alle prospettive acquisizionali, rispetti le fasi di sviluppo della L2, non violi le sequenze implicazionali” (Vedovelli, 2003: 178. Cfr. anche Rastelli 2009, 2010)⁷⁴.

A questo proposito, è alla linguistica acquisizionale che si deve l'identificazione delle sopra menzionate sequenze implicazionali, note anche come sequenze di acquisizione o acquisizionali, ossia le tappe che costituiscono l'ordine naturale e che risultano comuni a tutti gli apprendenti di una determinata lingua, qualsiasi sia la loro lingua d'origine e indicano quali contenuti grammaticali siano propedeutici all'acquisizione di altri:

La linguistica acquisizionale fornisce delle indicazioni utili tramite appunto le sequenze acquisizionali, riscontrate per diversi ambiti specie morfosintattici dell'italiano L2 (come di altre L2). Esse attestano che soggetti con L1 diversa, studiati longitudinalmente per un periodo più o meno lungo, acquisiscono nello stesso ordine (stadio 1 > stadio 2 > stadio n) alcune regole, nozioni e sottosistemi di L2 (ad esempio, strutture negative, interrogative, morfologia verbale), pur percorrendo la sequenza con velocità diverse, in dipendenza da vari fattori (la

⁷⁴ Si vedano a questo proposito i seguenti riferimenti per la didattica acquisizionale: per la lingua italiana si segnalano il volume sull'apprendimento e la didattica dell'italiano a studenti universitari sinofoni (Rastelli 2010 a c. di), Banfi (2003), Chini (2011, 2012a, 2012b, 2012c, 2016), lo studio di Della Putta, Visigalli (2010) sulla didattica del sintagma nominale italiano ad anglofoni, gli atti del convegno CIS del 2006 (Grassi *et al.*, 2008), gli atti dei convegni e le pubblicazioni del Centro di Italiano per Stranieri CIS dell'Università di Bergamo, dell'Associazione Italiana di Linguistica Applicata (AItLA) e dalla rivista on-line Italiano LinguaDue. Sul piano internazionale si segnalano Ellis (1997), Doughty (2003), Hinkel (2005), Housen, Pierrard (2005), Véronique (2000, 2005) e Trévisiol-Okamura, Komur-Thillooy (2011).

maggiore o minore marcatezza della struttura; la maggiore o minore distanza genetica e strutturale dalla L1 dell'apprendente; fattori extralinguistici come età, motivazione, esposizione all'input, attitudine alle lingue, ecc.).

Un aspetto di tali stadi, interessante sia in chiave teorica che didattica, è la loro valenza implicazionale: il raggiungimento dello stadio 3 implica l'acquisizione delle strutture degli stadi 1 e 2, il che ha un chiaro risvolto diagnostico e prognostico, fornendo un punto di orientamento anche per la programmazione didattica, in considerazione della cosiddetta "insegnabilità". (Chini 2016: 5)

Per quanto riguarda la lingua italiana, le sequenze acquisizionali finora scoperte riguardano la morfologia verbale e nominale, i pronomi clitici, la negazione, la connessione interproposizionale, la subordinazione, la modalità, gli ordini marcati e le strategie anaforiche (si vedano i capp. 3-6 di Giacalone Ramat 2003 e una sintesi in Chini 2005: 94-105). Sono state identificate a partire dagli anni '80 nell'ambito del "Progetto di Pavia"⁷⁵, il più longevo ed esteso progetto di ricerca sull'italiano L2, che ha studiato e descritto le fasi dell'acquisizione dell'italiano L2⁷⁶. Si cita a titolo di esempio la sequenza dell'acquisizione di tempi e modi verbali dell'italiano, ove il segno > indica che ciò che si trova alla sua sinistra è appreso prima di ciò che si trova alla sua destra: *presente (e infinito) > (ausiliare) participio passato > imperfetto > futuro > condizionale > congiuntivo* (Banfi, Bernini 2003: 90).

Le riflessioni che sono derivate in ambito didattico dalla scoperta delle sequenze acquisizionali sono le seguenti: è necessario insegnare le strutture linguistiche nell'ordine in cui si manifestano naturalmente nell'interlingua degli apprendenti, senza aspettarsi che le strutture che si trovano nelle tappe più avanzate della sequenza vengano usate correttamente prima che siano state apprese le strutture precedenti. È ciò a cui fa riferimento la *teachability hypothesis* (ipotesi della insegnabilità) formulata da Pienemann

⁷⁵ Nato a metà degli anni '80, il progetto è stato coordinato dapprima da Anna Giacalone Ramat e dal 2003 da Giuliano Bernini. Attivo in una prima fase a Pavia, poi all'Università di Bergamo, ha coinvolto vari atenei (Pavia, Bergamo, Milano Bicocca, Udine, Vercelli, Verona, Siena Stranieri, Torino, poi Verona e Modena-Reggio), con l'obiettivo principale di individuare, attraverso lo studio delle produzioni - principalmente orali - degli apprendenti i principi che li guidano nella scoperta del rapporto fra forma (linguistica) e funzione nella lingua di arrivo.

⁷⁶ Si citano a questo proposito i seguenti ambiti di ricerca e alcuni studi a titolo di esempio. Per l'italiano L2: le principali differenze tra le sequenze di acquisizione dell'italiano L1 e quelle dell'italiano L2 (Calleri *et al.* 2003: 221-228), la morfologia del nome L2 (Berretta 1987, Lo Duca 1990), il sistema del tempo, modo e aspetto verbale (Banfi, Bernini 2003), la sintassi della frase: (Calleri *et al.* 2003: 236-245) e la sintassi del periodo: (Ferraris 1999:53-62). Per l'italiano L1: la morfologia del nome: (Chini 1995: 129-151), il sistema di tempo, modo e aspetto verbale (Banfi, Bernini 2003, Lo Duca, Solarino 2006), la sintassi della frase: (Andorno 2003: 127-140, Andorno *et al.* 2003:116-127), la sintassi del periodo: (Ferraris 1999, Andorno 2003); (Ferraris 1999:34-42).

(1986), secondo la quale ha senso insegnare solo ciò che un apprendente può imparare in un certo momento del suo percorso acquisizionale, poiché “l’insegnamento può favorire l’acquisizione linguistica solo se l’interlingua è vicina al punto in cui la struttura da insegnare viene acquisita nella situazione naturale” (Pienemann, 1986: 313). Successivamente Pienemann (1998, 2005) fornisce un inquadramento teorico per tale ipotesi tramite la Teoria della Processabilità, con la quale spiega che ad ogni fase dell’acquisizione l’apprendente può produrre e comprendere solo le strutture linguistiche che il suo “processore” linguistico (una sorta di processore mentale della lingua) riesce ad elaborare in quel momento⁷⁷.

Gli studi nell’ambito della linguistica acquisizionale andarono ad innestarsi nel filone di ricerca in ambito glottodidattico, che, dagli anni Settanta del Novecento in poi portò al modello di unità didattica elaborato da Freddi (1970, 1994, 1999) e (Titone 1976), poi ripreso e approfondito da Danesi (1988, 1998) da un punto di vista neurolinguistico, mettendo in luce l’importanza di strutturare le attività sulla base dell’ordine in cui si attivano i due emisferi cerebrali (Danesi 1998: 152 e sgg.) e, infine, da Balboni (1994, 2002, 2008), con un’evoluzione nella sua organizzazione interna. Nel modello di Freddi e Danesi alle tre fasi originarie della psicologia della *Gestalt* se ne aggiungono altre tre, andando a costituire il seguente percorso di fasi di apprendimento, collegabili a processi psicodidattici, che corrispondono a: motivazione → globalità → analisi → sintesi → riflessione → controllo/verifica. Tale modello gestaltico è stato applicato in questa ricerca nella progettazione delle unità didattiche

Pertanto, in entrambe le unità didattiche progettate, le attività si svolgono seguendo la sequenza di fasi sopra descritta: *Contatto con il tema*, con attività di stimolo della motivazione e mirate ad esprimere le conoscenze pregresse sul tema che verrà trattato; *Comprensione*, ossia la fase della globalità, in cui viene chiesto agli studenti di leggere uno o più testi e rispondere a delle domande di comprensione testuale; nella fase di *Analisi* si chiede di identificare nel/i testo/i alcuni elementi lessicali e sintattici; *Applicazione pratica*, ossia la fase della sintesi, in cui si chiede agli studenti di sistematizzare le strutture sintattiche e lessicali precedentemente individuate e reimpiegarle; *Riflessione*, in cui si chiede di elaborare o completare una definizione del loro uso. Infine, vengono proposte alcune attività di verifica dell’acquisizione dei contenuti grammaticali e pragmatici al

⁷⁷ Tra i richiami teorici di tale teoria vi è il modello psicolinguistico di Levelt (1989) e la Grammatica Lessico-Funzionale di Bresnan (2001) e ha trovato conferma in alcune sperimentazioni didattiche sull’italiano L2 (Bettoni, Di Biase 2008, Whittle 2015, Di Biase, Bettoni, 2015).

centro dell'unità didattica. Nel modello di unità didattica elaborata da Freddi la successione delle diverse fasi dell'unità consente dunque, di realizzare

un armonico gioco di rimandi tra i due emisferi: dalla fase di motivazione, che integra le emozioni e la curiosità del cervello destro con l'analisi dei bisogni di quello sinistro, si passa ad una fase di globalità (modalità destra), cui segue l'analisi della lingua (modalità sinistra) per approdare alla sintesi conclusiva. (Balboni 2002: 28)

Per quanto riguarda la durata dello svolgimento delle unità didattiche, Freddi (1994: 111) propone come valutazione temporale per ciascuna unità "4-6 ore di lezione circa a seconda degli obiettivi posti dall'unità stessa e della consistenza dei materiali da manipolare"; la durata aumenta a 6-8 ore nel caso di unità pensate per la scuola secondaria (Freddi 1999: 148) o a 6-10 ore o oltre nel caso di unità basate su testi letterari o tecnico-specialistici (Balboni 2002: 103).

Lo stesso Balboni nel 2002 ha rivisto l'idea di unità didattica, proponendone un'innovazione attraverso il modello operativo della rete di unità di apprendimento:

Un'unità d'*apprendimento* può durare pochi minuti o anche un'ora: è l'unità di misura secondo la quale lo studente percepisce il suo apprendimento: "oggi a scuola abbiamo fatto..., ho imparato a..."; un'unità *didattica* è invece una tranche linguistico-comunicativa più complessa, realizzata mettendo insieme eventi, atti, espressioni, strutture linguistiche legati da un contesto situazionale. (Balboni 2002: 105-106)

In questo caso il modello operativo si basa su un'unità didattica da 6-10 ore che al suo interno contiene una rete di unità di apprendimento di varia durata, che può essere svolta in maniera flessibile, senza alcuna sequenza obbligata delle attività in termini di ordine temporale o gerarchico (l'unica sequenza rispettata nella sua progettazione è quella acquisizionale "i+1", che presenta le strutture linguistiche secondo l'ordine naturale di acquisizione, a cui si è fatto riferimento in precedenza).

Nel caso del presente progetto di ricerca, si è optato, invece per la produzione di due brevi unità didattiche sul modello di Freddi, aventi però una durata compresa tra 1 ora e 30 e 2 ore. Tale scelta è dovuta al fatto che la sperimentazione delle attività didattiche sull'uso della varietà neostandard si è svolta nell'ambito di accordi siglati con alcuni atenei esteri che hanno concesso la disponibilità all'interno dei propri corsi curriculari per massimo 2 o 3 incontri della durata di 1 ora e 30 ciascuno, fino ad un massimo di 3-4 ore totali.

Pertanto, è bene precisare che, nonostante la piena consapevolezza dell'evoluzione che ha investito il modello dell'unità didattica (cfr. i già citati Diadori, Palermo e Troncarelli 2015: 285-312) - dall'originaria matrice gestaltica fino alla moderna unità di apprendimento - la scelta di progettare due brevi unità didattiche improntate sul modello originario gestaltico è dovuta al contesto specifico di sperimentazione, che non disponeva di una continuità temporale tale da permettere lo svolgimento di una rete di unità di apprendimento.

Di conseguenza, le brevi unità prodotte possono essere considerate come delle forme ibride tra l'unità didattica tradizionale, di cui rispettano la struttura ispirata alla bimodalità emisferica secondo i già citati studi di Danesi, e le unità di apprendimento, di cui colgono la struttura reticolare, dato che offrono al docente varie attività con diversi livelli di complessità, da realizzare nell'ambito di un'unica sessione di lavoro di 1 ora e 30 / 2 ore e non di un percorso superiore alle 3 -4 ore.

Dopo aver fornito i necessari riferimenti teorici e metodologici, per maggiori dettagli sulla progettazione delle singole unità si rinvia al capitolo 6, mentre la sezione successiva sarà dedicata alla descrizione di un questionario progettato per la raccolta delle opinioni dei partecipanti sulle attività svolte.

4.9. Questionario sulle opinioni degli studenti

Per raccogliere dati sulla popolazione coinvolta e sulla percezione personale relativa all'utilità delle attività svolte, all'acquisizione dei contenuti linguistici e all'adeguatezza della piattaforma di *e-learning* rispetto agli scopi della didattica, è stato elaborato uno strumento d'osservazione da somministrare ai partecipanti al termine di ciascuna lezione.

Tra i vari strumenti che consentono l'osservazione diretta della didattica delle lingue tramite una registrazione di dati strutturata, si è scelto di progettare uno strumento che integrasse la tipologia del questionario con quella della *checklist* (si vedano gli studi di Nunan 1989, Coonan 2000, Pozzo 2008). Entrambi gli strumenti sono tra quelli maggiormente utilizzati nell'ambito della ricerca-azione per la raccolta quantitativa dei dati (Coonan 2000:25). In particolare, la *checklist* si basa su una selezione di voci prestabilite, si tratta dunque di uno strumento chiuso, volto a focalizzare l'attenzione su aspetti specifici individuati dal ricercatore. Come espresso da Coonan (ivi: 37):

Una checklist può essere usata per investigare praticamente qualsiasi aspetto del processo d'insegnamento e d'apprendimento. In sostanza non è altro che un elenco di elementi che sono da segnare, scegliere, categorizzare oppure mettere in ordine. Può fungere sia come scheda di osservazione 'in diretta' (cfr. COLT) sia come strumento di riflessione 'a distanza' dove l'uso dello strumento non è legato direttamente ad un qualche aspetto osservato in tempo reale. [...] La checklist può anche avere incorporata una scala valutativa se viene richiesto di valutare l'item che si è osservato o sul quale si sta riflettendo.

Pertanto, si è deciso di strutturare lo strumento come segue:

- a) Domande aperte di tipo conoscitivo: lingua materna; altre lingue straniere conosciute oltre all'italiano; da quanto tempo è in corso l'apprendimento dell'italiano; eventuali conoscenze pregresse sul tema linguistico-grammaticale dell'unità ed eventuale presenza di fenomeni linguistici analoghi nella lingua madre e/o in altre lingue straniere conosciute;
- b) 3 liste di controllo (*checklist*), in cui si richiede di attribuire alle voci dell'elenco un valore su una scala numerica da 0 a 3. La scala di valori è stata ulteriormente esplicitata da *emoticon*, in modo da adattare la scheda ai diversi stili di apprendimento (analitici o olistici) degli apprendenti:
 - a. la prima contiene l'elenco delle attività svolte nell'unità e si chiede agli apprendenti di valutarle in base alla loro utilità percepita;
 - b. la seconda contiene quattro affermazioni sulla percezione dell'acquisizione del tema grammaticale dell'unità, ordinate dal livello massimo al livello minimo della soddisfazione personale (Es.; dopo queste attività ho un'idea più chiara dell'uso dei riflessivi).
 - c. la terza contiene quattro affermazioni sulla soddisfazione degli studenti rispetto all'uso della piattaforma *e-learning*.

Sia la seconda che la terza lista offrono, inoltre, spazi per aggiungere apporti personali, affinché il campo di osservazione non sia ristretto alle sole voci prestabilite, come messo in evidenza da Coonan (*ibidem*):

[...] le checklist (come tutte le schede di osservazione) consistono di una serie di categorie pre-designate per la registrazione delle osservazioni per cui sono capaci di dare informazioni solo riguardo le categorie previste. In questo modo si corre il rischio di non 'vedere' aspetti interessanti proprio perché la scheda ci orienta in una precisa direzione che esclude moltissime cose. Ciò non toglie che se si è ben individuato un ambito preciso di indagine la checklist o

scheda di osservazione ci permette di raccogliere delle informazioni a riguardo in maniera precisa e sistematica.

Per quanto riguarda, invece, la formulazione dei contenuti dello strumento di osservazione si è basata sul concetto neurobiologico di *Stimulus Appraisal* (Schumann 1999, Schumann *et al.* 2004), ossia di valutazione dell'input. Tale concetto, applicato alla glottodidattica (cfr. Daloiso 2009; 2015) permette di descrivere la motivazione all'apprendimento come un processo continuo del cervello di valutazione e confronto degli input ricevuti dal mondo esterno con le aspettative, i bisogni e i desideri. Se in base a tale confronto l'input verrà giudicato positivamente, verrà inviato dalla memoria di lavoro ai centri deputati alla memoria a lungo termine. Tale processo di valutazione dell'input è stato definito *stimulus evaluation check* Scherer (1984, 2001, 2009) e si basa sui seguenti criteri:

1. Motivazione e *novità* dell'input: il soggetto valuta la novità dell'input, la sua prevedibilità e la sua discrepanza rispetto ad aspettative e conoscenze pregresse;
2. Motivazione e *piacevolezza intrinseca* dell'input: viene valutato il senso di piacere o di dis-piacere provocato dall'input;
3. Pertinenza dell'input rispetto ai bisogni e agli obiettivi: si valuta se l'input è adeguato al raggiungimento dei propri obiettivi formativi, culturali, sociali ecc.;
4. Realizzabilità sulla base della comprensibilità dell'input e della sua adeguatezza rispetto alle capacità del soggetto;
5. Sicurezza psico-sociale: il soggetto valuta se la situazione (nello specifico, la situazione comunicativa, nel caso della glottodidattica) può rafforzare la propria immagine sociale o comprometterla; si tratta della paura della "perdita della faccia" (Balboni 2002) che può innescare il cosiddetto filtro affettivo (Krashen 1982), ossia una difesa psicologica che il cervello innalza quando un soggetto ha paura di sbagliare e teme di compromettere la propria immagine.

È su questi criteri che si è basata la progettazione del questionario da somministrare ai partecipanti alla sperimentazione, nella consapevolezza che, come affermato da Daloiso (2009, 2015), nell'insegnamento e apprendimento delle lingue è necessario che l'input susciti emozioni positive, evitando situazioni che generino paura e ansia.

Come per le unità didattiche, anche per la progettazione e somministrazione del questionario è stata utilizzata una piattaforma online, nello specifico *Socrative*⁷⁸, il cui riferimento è stato reperito spesso in associazione alle piattaforme di *e-learning* prese in

⁷⁸ <https://socrative.com> (13/01/2019).

esame per la progettazione delle unità didattiche. *Socrative* consente di svolgere verifiche in classe scegliendo tra varie tipologie di quiz.

I fattori principali che ne hanno determinato l'adozione per la presente ricerca sono i seguenti. Innanzitutto, la possibilità di scegliere tra varie tipologie di quiz e di personalizzarne la struttura e la visualizzazione a seconda delle esigenze. Inoltre, vi è la possibilità per il docente di ricevere i risultati in tempo reale sul suo dispositivo durante lo svolgimento del test e non solo al termine, il che permette di monitorare l'andamento del quiz e di individuare a colpo d'occhio eventuali blocchi da parte dei partecipanti, funzione particolarmente utile nel caso di partecipanti italofoni non nativi che svolgono il test in lingua italiana. Una volta ottenuti i dati finali, la piattaforma offre una serie di funzioni che agevolano il lavoro di analisi dei dati, poiché consente di analizzare i report online dei dati ottenuti su un foglio di lavoro di Google o di impostare l'invio via e-mail in formato Ms-Excel, di effettuare il calcolo automatico delle percentuali delle risposte date nella classe e di realizzare automaticamente istogrammi. Infine, consente sia una visualizzazione a livello globale degli esiti dell'intera classe, sia di avere una visione dettagliata delle risposte dei singoli partecipanti, con la possibilità di nascondere i nomi per preservarne l'anonimato.

In appendice (cfr. Appendice 3) si riportano gli strumenti di osservazione utilizzati per valutare rispettivamente l'unità didattica sui verbi riflessivi e quella sulla scelta tra congiuntivo e indicativo. Per le informazioni sugli esiti del questionario si rinvia al capitolo 6, mentre il capitolo successivo sarà dedicato alla descrizione dei risultati ottenuti dall'interrogazione del *corpus*.

CAPITOLO V

FREQUENZA DEI FENOMENI GRAMMATICALI INDAGATI

Dopo aver descritto nel capitolo precedente l'elaborazione delle stringhe di ricerca in linguaggio CQL (cfr. sezione 4.6.2), in questo capitolo verranno esposti i risultati ottenuti dall'interrogazione del corpus. Dopo una breve descrizione dell'interrogazione del corpus e del salvataggio dei risultati ottenuti, si illustreranno i procedimenti adottati per il calcolo della frequenza all'interno del corpus dei fenomeni grammaticali oggetto di studio, al fine di determinare quali di essi siano i più ricorrenti.

Tale aspetto è di particolare importanza per la presente ricerca, poiché, come verrà descritto, la frequenza dei fenomeni grammaticali nel corpus è uno dei criteri adottati per la selezione dei fenomeni grammaticali sui quali costruire dei materiali didattici per l'italiano L2. Essa permette di individuare quali tra i tratti linguistici che fino a 30 anni fa, all'epoca della pubblicazione degli studi di Sabatini sul tema, erano ancora percepiti come in contrasto rispetto alla norma e in via di graduale inserimento al suo interno, siano invece ora maggiormente utilizzati nella produzione scritta di media formalità o informalità dagli italofoeni nativi. Ciò è possibile in quanto questi fenomeni vengono percepiti non più (o comunque meno) in contrasto con la norma e quindi considerati più facilmente accettabili in un registro non particolarmente "basso". La loro conoscenza è necessaria agli apprendenti di lingua italiana L2/LS, al fine di raggiungere una competenza linguistica vicina all'uso reale della lingua.

5.1. Frequenza dei fenomeni indagati

Allo scopo di determinare la diffusione dei tratti linguistici indagati, per ciascuno di essi sono state svolte le seguenti operazioni, che verranno di seguito illustrate prendendo l'esempio del fenomeno grammaticale del doppio partitivo:

1. innanzitutto, per ciascun campione è stato calcolato il *rapporto tra il numero di occorrenze salvate* dopo l'eliminazione di eventuali casi non inerenti e doppioni (ossia le occorrenze effettivamente corrispondenti al fenomeno grammaticale indagato) *e le 200 occorrenze iniziali del campione*. Tale procedura permette di verificare quanto sia efficace la stringa di ricerca, al fine di avere una stima più veritiera e affidabile della

frequenza dei fenomeni indagati. Nello specifico, nel caso del doppio partitivo, per ciascun campione sono stati calcolati i seguenti rapporti: campione 1: $161/200= 0,805$; campione 2: $167/200= 0,835$; campione 3: $174/200= 0,87$;

2. *media campione*: è stata calcolata la media dei tre rapporti ottenuti, determinando così la frequenza media del fenomeno per campione, ad esempio: $(0,805 + 0,835 + 0,87)/3= 0,84$;
3. *deviazione standard campione*: la deviazione standard fornisce una misura di come siano distribuiti i dati nel campione rispetto alla media, ossia permette di sapere se la media costituisce una rappresentazione affidabile dei dati. Quanto più il valore della deviazione standard è vicino a 0, tanto più i dati si avvicinano alla media, e quindi si può affermare che c'è poca dispersione nel campione. Nel caso del doppio partitivo, la media campione è pari a 0,84 e la deviazione standard campione è pari a 0,03, il che lascia intendere l'omogeneità dei risultati all'interno del campione. Si è scelto di calcolare la deviazione standard campione e non l'errore standard della media per pura semplicità di calcolo.

Al fine di rendere paragonabili i valori ottenuti per ciascuno dei fenomeni grammaticali indagati, sono state, inoltre, calcolate:

1. *frequenza stimata*: è data dal prodotto tra il totale delle occorrenze inizialmente ottenute per il fenomeno indagato tramite l'interrogazione del corpus PAISÀ e la media campione, ossia nel caso del doppio partitivo: $659*0,84= 553,56$;
2. *frequenza stimata per 1 milione di parole*: è ottenuta attraverso due passaggi: si calcola il prodotto tra la frequenza stimata e 1.000.000 e il risultato ottenuto si divide per il numero totale di *token* del corpus PAISÀ, ossia: $frequenza\ stimata*1.000.000/265.062.636^1$. Esempio: $553,56* 1.000.000 / 265.062.636$.

Ciò è stato svolto nella consapevolezza che la scelta di avere tre campioni casuali su cui calcolare dapprima la frequenza stimata (al netto del rumore) e, successivamente, la normalizzazione per milione di parole non rappresentino procedure standard. Infatti, generalmente per quantificare i casi inerenti al fenomeno ricercato rispetto al cosiddetto "rumore" le stime di frequenza si basano su un campione unico. Tuttavia, si è ritenuto che avere tre campioni casuali consentisse di disporre di una base di dati più eterogenea e, pertanto, di poter formulare delle affermazioni sulla frequenza dei fenomeni basate – si auspica – su una maggiore fondatezza.

¹ Fonte: https://corpora.dipintra.it/unibo/run.cgi/corp_info?corpname=paisa (link accessibile solo da rete UNIBO) Cfr. voce *token counts*.

Si è scelto, infine, di applicare la procedura di normalizzazione anche ai valori della deviazione standard campione, in modo che fosse in proporzione alla frequenza, calcolando, pertanto:

1. *deviazione standard campione stimata*: è data dal prodotto tra il totale delle occorrenze inizialmente ottenute da PAISÀ e la deviazione standard campione. Esempio: $659 * 0,03 = 19,77$
2. *deviazione standard stimata per 1 milione di parole*: è data dal prodotto tra deviazione standard campione stimata e 1.000.000, diviso per il numero totale di token di PAISÀ, ossia: deviazione standard campione stimata * 1.000.000 / 265.062.636. Esempio: $19,77 * 1.000.000 / 26.5062.636$;

I valori della frequenza stimata per 1 milione di parole (cfr. punto 2) e della deviazione standard stimata per 1 milione di parole (cfr. punto 4) di tutti i fenomeni grammaticali indagati sono stati utili per confrontare tra loro i fenomeni e ordinarli in una lista dal fenomeno meno frequente a quello più frequente all'interno del corpus (cfr. seguenti tabelle 5.1, 5.2 e 5.3).

I 26 casi esaminati, che includono 13 tratti della lista di Sabatini (1990) e i relativi sottofenomeni di alcuni di essi, sono stati suddivisi in terzili in base alla loro frequenza nel corpus PAISÀ (ove con *frequenza* si intende la sopra menzionata frequenza stimata per milione di parole). I tre terzili comprendono gli 8 casi a frequenza bassa, i 9 casi a frequenza media e i 9 casi a frequenza più alta, riportati nelle seguenti tabelle riassuntive.

Tabella 5.1: Fenomeni a bassa frequenza

Fenomeni a bassa frequenza ($0 < fq < 3$)	Frequenza stimata per 1M di parole	Deviazione standard stimata per 1M di parole	N.ro di stringhe di ricerca
1. Soggetto in posizione rematica rispetto al predicato: <i>niente</i> + sostantivo pl.	2,21	0,10	1
2. Dislocazione a sinistra: doppio partitivo	2,08	0,08	1
3. Impiego dell'indicativo al posto del congiuntivo: indicativo nelle ipotetiche	1,59	0,45	1
4. Dislocazione a sinistra: ripetizione del pronome in forma atona e tonica	0,41	0,29	8
5. <i>Ci</i> attualizzante	0,38	0,33	3

6. Impiego dell'indicativo al posto del congiuntivo: <i>avere paura/timore che</i>	0,03	0,01	1
7. Soggetto in posizione rematica rispetto al predicato (in frasi subordinate)	< 0,01	< 0,01	1
8. Impiego dell'indicativo al posto del congiuntivo: verbi di opinione/sentimento	< 0,01	< 0,01	1

Tabella 5.2: Fenomeni a media frequenza

Fenomeni a media frequenza ($3 < fq < 20$)	Frequenza stimata per 1M di parole	Deviazione standard stimata per 1M di parole	N.ro di stringhe di ricerca
1. Pronome <i>gli</i> vs <i>le/loro</i>	16,50	2,48	1
2. Soggetto in posizione rematica rispetto al predicato (in frasi principali)	15,96	5,79	5
3. Frase scissa con <i>a</i>	12,77	0,55	1
4. <i>Che</i> polivalente	9,46	7,95 ²	3
5. Verbo in forma pronominale (es. <i>io mi bevo</i>)	6,48	6,61	12
6. Impiego dell'indicativo al posto del congiuntivo: indicativo nell'interrogativa indiretta	5,35	0,24	1
7. Concordanza a senso: soggetto collettivo sing.+ <i>di</i> partitivo + predicato pl. (es. <i>un milione di persone hanno visto</i>)	3,55	0,07	1
8. Impiego dell'indicativo al posto del congiuntivo: <i>non</i> + verbi di opinione/sentimento	3,33	0,31	1
9. <i>Cosa</i> interrogativo vs <i>che cosa</i> nell'interrogativa indiretta	3,01	0,11	1

² I fenomeni che presentano un valore superiore a zero della deviazione standard saranno discussi nei paragrafi seguenti.

Tabella 5.3: Fenomeni ad alta frequenza

Fenomeni ad alta frequenza (fq > 20)	Frequenza stimata per 1M di parole	Deviazione standard stimata per 1M di parole	N.ro di stringhe di ricerca
1. <i>Ma</i> ad inizio frase	344	2,64	1
2. <i>E</i> ad inizio frase	289,48	1,67	1
3. Il partitivo preceduto da preposizione	67,06	1,16	1
4. <i>Per cui</i> con valore conclusivo	52,28	14,28	1
5. <i>Comunque</i> ad inizio frase	32,31	0	1
6. Impiego dell'indicativo al posto del congiuntivo: verbo <i>essere</i> + aggettivo + <i>che</i>	26,73	2,29	1
7. <i>Cosa</i> interrogativo vs <i>che cosa</i> nell'interrogativa diretta	26,01	0,21	1
8. Frase scissa con <i>che</i>	24,33	12,39	4
9. <i>Allora</i> ad inizio frase	20,62	0	1

Per ciascun terzile è stato realizzato un istogramma (cfr. grafici 5.1, 5.2 e 5.3) in cui l'altezza delle colonne blu corrisponde ai valori sopra citati della frequenza stimata per 1 milione di parole, e la lunghezza delle linee sottili nere sovrapposte alle colonne indica l'ampiezza della deviazione standard stimata (v. Tabelle 5.1-5.3).

Grafico 5.1: Fenomeni a bassa frequenza

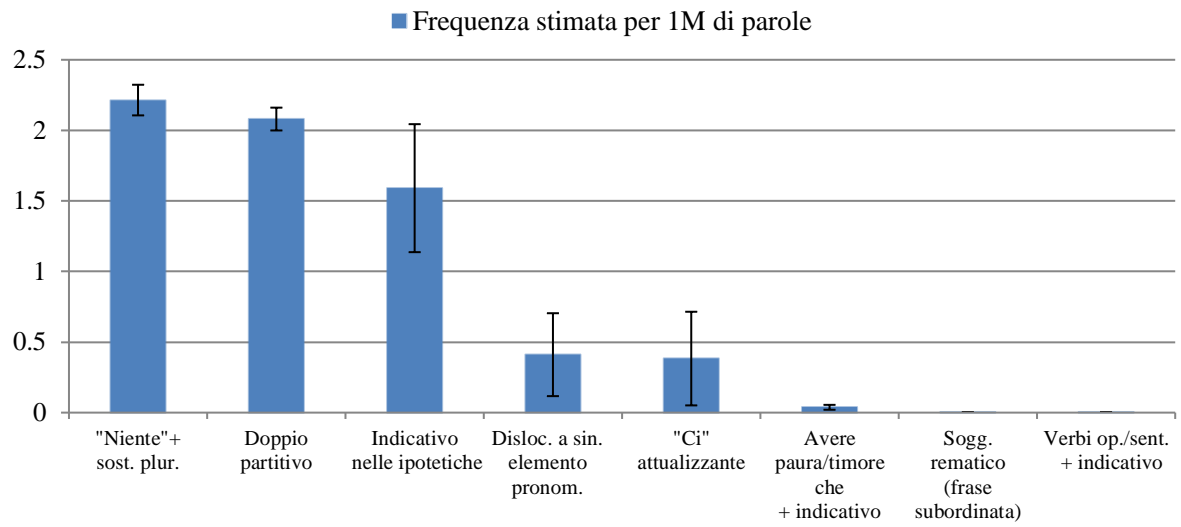


Grafico 5.2: Fenomeni a media frequenza

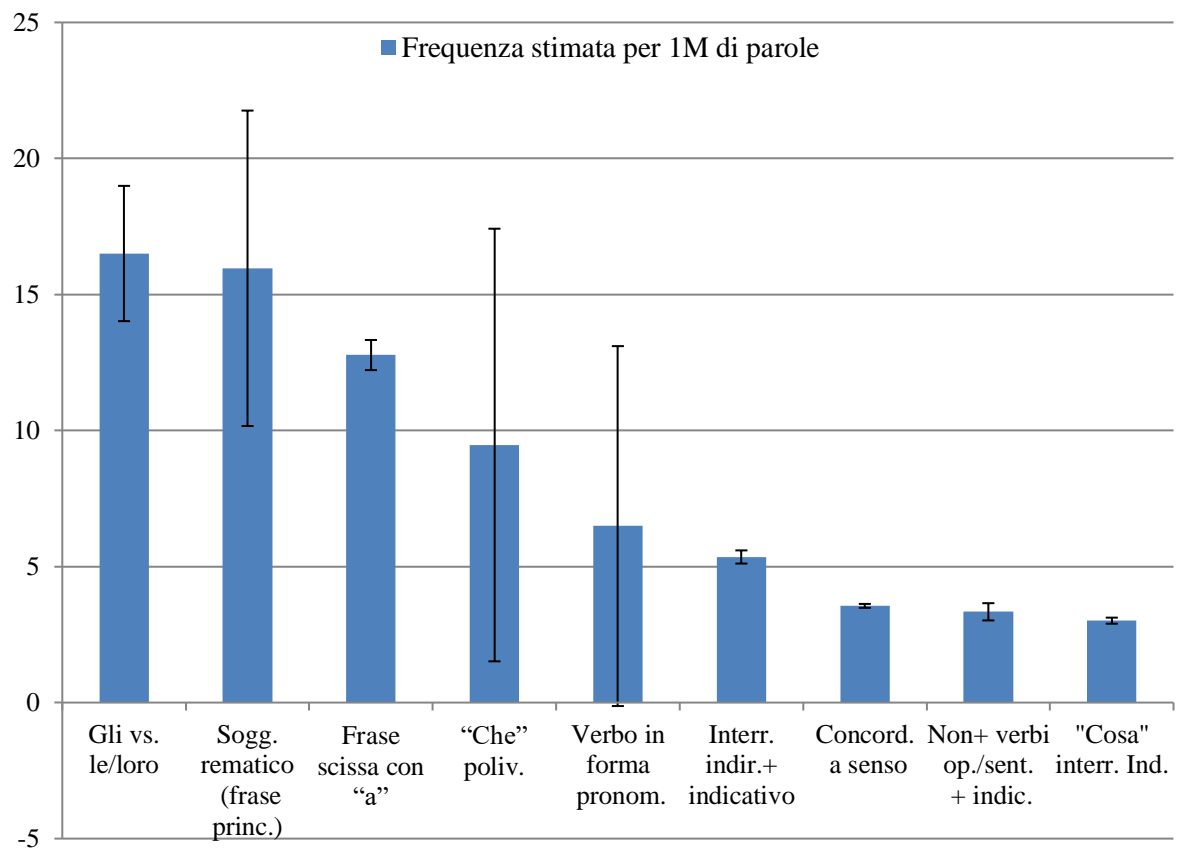
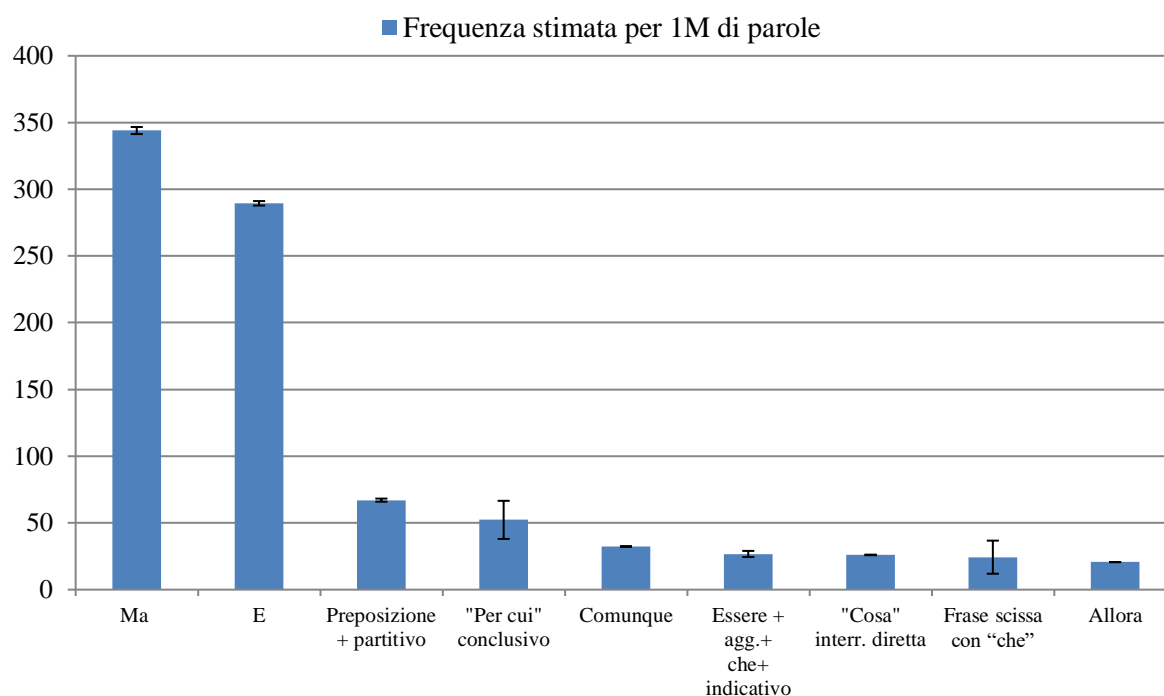


Grafico 5.3: Fenomeni ad alta frequenza



Osservando la tabella 5.3 e i grafici 5.1, 5.2 e 5.3 emerge immediatamente che alcuni fenomeni presentano un'elevata deviazione standard rispetto alla frequenza. Al fine di analizzare questi dati e determinare da cosa dipenda una simile variazione dell'errore, per ciascun terzile saranno analizzati due casi che presentano molta variazione e due casi con poca variazione e per ciascun fenomeno si porrà l'attenzione sui seguenti aspetti: il numero di stringhe di ricerca prodotte, il numero di campioni ricavati da ciascuna stringa, il totale iniziale di occorrenze da cui è composto ciascun campione e il totale delle occorrenze salvate per ciascun campione. Nello specifico, saranno analizzate le seguenti coppie di fenomeni, che sono stati abbinati sulla base delle analogie dei rispettivi dati della frequenza e della deviazione standard (elevata o scarsa variazione):

1. *Primo terzile, fenomeni a bassa frequenza:*

1.1. Fenomeni che presentano un'alta variazione: l'uso dell'indicativo nelle frasi ipotetiche vs il *ci* attualizzante;

1.2. Fenomeni che presentano poca variazione: soggetto in posizione rematica (*niente* seguito da un sostantivo plurale) vs dislocazione a sinistra (uso del doppio partitivo);

2. *Secondo terzile, fenomeni a media frequenza:*

- 2.1. Fenomeni che presentano un'alta variazione: il soggetto pronominale *vs* il *che* polivalente;
- 2.2. Fenomeni che presentano poca variazione: la frase scissa retta da *a* *vs* l'uso dell'indicativo nella frase interrogativa indiretta;
3. *Terzo terzile, fenomeni ad alta frequenza:*
 - 3.1. Fenomeni che presentano un'alta variazione: l'uso di *per cui* con valore conclusivo *vs* la frase scissa retta da *che*;
 - 3.2. Fenomeni che presentano poca variazione: il partitivo preceduto da preposizione *vs* l'uso dell'indicativo al posto del congiuntivo nella costruzione “verbo *essere* + aggettivo + *che* + indicativo”.

5.2. Fattori di variazione

Questa analisi è stata necessaria per poter affermare che la frequenza è un dato affidabile. Sulla base dei risultati ottenuti, i casi indagati possono essere suddivisi in base alle cause della variazione:

- 1) variazione non casuale, alla cui base vi sono fattori ben identificabili, di cui si citano a titolo di esempio: il numero di stringhe di ricerca prodotte per il singolo fenomeno, il numero di campioni ricavati da ciascuna stringa (3 campioni *vs* un campione unico), il totale iniziale di ciascun campione (200 unità o meno), il fatto che per alcuni fenomeni le stringhe abbiano permesso di reperire un numero elevato di occorrenze e in altri casi meno, la differenza tra il totale iniziale di occorrenze per ciascun campione e il totale delle occorrenze salvate al termine della ripulitura del campione dai casi non inerenti (ad esempio, 176 su 200), eccetera;
- 2) variazione casuale, ove non emerge una causa evidente per cui per alcuni fenomeni la frequenza sia alta e per altri bassa. In generale, ci si aspetta che la variazione sia casuale nei fenomeni a bassa frequenza per pure ragioni probabilistiche.

Per ciascuno dei fattori di variazione identificati verranno elencati a seguire alcuni esempi specifici dei fenomeni a cui si applicano.

1) Variazione non casuale:

1a) Numero di stringhe di ricerca prodotte per il singolo fenomeno: seguono i casi in cui la variazione è determinata dalla quantità di stringhe prodotte per inquadrare il fenomeno grammaticale.

Nel caso del **verbo in forma pronominale per indicare partecipazione affettiva**, non essendo possibile progettare un'unica stringa che coprisse tutte le forme pronominali relative all'intera coniugazione, sono state realizzate 12 stringhe di ricerca, al fine di reperire occorrenze per tutte le 6 forme pronominali riflessive in presenza di verbi dai tempi composti (esempio: *mi sono mangiato la pizza*) e non composti (esempio: *mi mangio la pizza*). A questo proposito, la scelta di costituire un unico campione finale si giustifica alla luce del fatto che le 12 stringhe di ricerca rappresentano, come già spiegato, un unico fenomeno, ossia l'intera coniugazione, nei tempi composti e non composti, di verbi usati in forma riflessiva per sottolineare la partecipazione affettiva, per i quali non è possibile progettare un'unica stringa che copra tutte le forme pronominali usate per l'intera coniugazione.

Nel caso della **frase scissa nella sua variante retta da *che*** sono state realizzate 4 stringhe di ricerca allo scopo di reperire occorrenze per tutte le sue diverse realizzazioni: 1) verbo *essere* + gruppo nominale + *che* + verbo coniugato; 2) verbo *essere* + verbo all'infinito + *che* + verbo coniugato; 3) verbo *essere* + avverbio + *che* + verbo coniugato; 4) verbo *essere* + gruppo nominale preceduto da preposizione + *che* + verbo coniugato.

Infine, si citano i casi del ***che* polivalente** e del ***ci* attualizzante**, per ciascuno dei quali sono state progettate 3 stringhe per reperire i vari casi dell'uso della particella *ci* in combinazione con il verbo *avere* seguito da verbo o sostantivi e delle tre tipologie di *che* polivalente individuate da Budassi (2017).

1b) Produttività della stringa e differenza tra occorrenze iniziali scaricate e occorrenze finali salvate: a seguire verranno elencati i tratti grammaticali per i quali la variabilità della frequenza dipende sia dalla produttività delle stringhe di ricerca, ossia dal fatto che in alcuni casi le stringhe abbiano permesso di reperire un numero elevato di occorrenze e in altri casi meno, sia dalla differenza esistente tra il totale iniziale di occorrenze per ciascun campione e il totale delle occorrenze salvate al termine della ripulitura del campione dai casi non inerenti.

Nel caso del già citato ***ci* attualizzante**, a causa del numero ridotto di risultati ottenuti dall'interrogazione del corpus tramite le 3 stringhe, sono stati realizzati campioni nettamente inferiori alle 200 unità: per la prima stringa, inerente ai casi *c'* + verbo *avere* + sostantivo (esempio: *c'ho voglia*) sono stati realizzati 2 campioni da 21 occorrenze e 1 da 22; nel caso della seconda stringa, inerente ai casi *ci* + verbo *avere* + sostantivo (esempio: *ci ha voglia*) sono stati realizzati 2 campioni da 40 e 1 da 41 e, infine, per la terza stringa, inerente ai casi *c'* + verbo *avere* + *da* + verbo (esempio: *c'ho da fare*): è stato realizzato un

unico microcampione da 6 unità. Osservando i campioni al termine della ripulitura dai casi non inerenti, emergono delle notevoli differenze: mentre nel caso della prima stringa il totale delle occorrenze salvate differisce di poche unità rispetto alle occorrenze iniziali (tre campioni: 19/21, 19/21, 21/22), nel caso della seconda stringa quasi la totalità dei risultati ottenuti non è inerente al fenomeno in questione (tre campioni: 1/40, 1/40, 3/41). Vi è, poi, il caso della terza stringa, la meno produttiva in assoluto, in cui 5 occorrenze su 6 del microcampione unico corrispondono al caso in esame.

Riassumendo, nel momento in cui i risultati ottenuti da tutte le stringhe del *ci* attualizzante sono stati selezionati tutti insieme per il calcolo della media e della deviazione standard, essi sono andati a costituire campione unico finale relativo all'intero fenomeno in questione, la cui variabilità dipende dai seguenti fattori: il già citato numero disomogeneo di campioni per ciascuna stringa (3, 3, campione unico), la variabilità dei totali iniziali delle occorrenze per ciascun campione, compresi tra 41 e 6, e la variabilità dei totali delle occorrenze salvate, comprese tra 1 a 21.

Nel caso del **verbo in forma pronominale per indicare partecipazione affettiva**, per quanto riguarda la produttività delle stringhe progettate per reperire i tempi non composti dei verbi (esempi: *mi mangio, ti farai, ci berremmo*, ecc.), solo nel caso della terza persona singolare e plurale è stato possibile reperire nel corpus 600 occorrenze da suddividere in tre campioni da 200 unità ciascuno. Tuttavia, entrando nel dettaglio, si può notare che, a fronte di un numero elevato di occorrenze reperite, in realtà le occorrenze inerenti al tratto grammaticale in questione e, dunque, salvate per ciascun campione, sono esigue: terza persona singolare: 2/200, 3/200, 3/200; terza persona plurale: 0/200, 4/200, 3/200. Come emerso dalla lettura delle occorrenze reperite, quasi la totalità di queste erano in realtà casi di *si* passivante (esempio: *ora si fanno delle ipotesi*) o impersonale (esempio: *Si fa ricorso ad un provvedimento; vi si mangiano i piatti tipici carnevaleschi*) e non forme riflessive utilizzate per indicare partecipazione affettiva (esempio: *quanti di noi si fanno il cous-cous invece che il riso? vs quanti di noi fanno il cous-cous*) e pertanto sono state scartate.

Nel caso della seconda persona plurale sono state reperite solo 17 occorrenze, che hanno costituito un campione unico del quale sono state salvate 14 occorrenze. In tutti gli altri casi delle stringhe dei tempi non composti dei verbi si nota come, a seconda dei casi, le stringhe abbiano permesso di reperire un numero maggiore o minore di occorrenze e che vi siano anche differenze tra il totale iniziale di occorrenze per ciascun campione e il totale delle occorrenze salvate al termine della ripulitura del campione dai casi non inerenti:

prima persona singolare: 13/56, 24/56, 24/56; seconda persona singolare: 11/21, 11/21, 6/20; prima persona plurale: 5/24, 6/24, 9/24.

Per quanto riguarda, invece, le stringhe progettate per reperire i tempi composti dei verbi, nel caso della prima e della seconda persona singolare e plurale il numero ristretto di occorrenze ottenute dal corpus ha comportato la creazione di un campione unico per ciascuna stringa, ciascuno di essi caratterizzato da un numero ridotto di occorrenze salvate: prima persona singolare: 13/49; seconda persona singolare: 4/23; prima persona plurale 5/37 e, infine, seconda persona plurale: 3/10. Per quanto riguarda la differenza tra la quantità di occorrenze reperite e quelle salvate per la terza persona singolare e plurale, valgono le stesse riflessioni fatte precedentemente per i tempi non composti dei verbi.

Anche in questo caso è necessario ricordare che, nel momento in cui i risultati ottenuti da tutte le stringhe sono stati selezionati collettivamente per il calcolo della media e della deviazione standard, essi sono andati a costituire campione unico finale relativo all'intero fenomeno del verbo in forma pronominale per indicare partecipazione affettiva, la cui variabilità dipende dai seguenti fattori: il già citato numero elevato di stringhe, un numero disomogeneo di campioni per ciascuna stringa (3 *vs* campione unico) e la differenza tra il numero elevato di occorrenze reperite e le occorrenze salvate per ciascun campione, con queste ultime che variano da 0 a 24.

Come già anticipato nel capitolo 4 (cfr. sezione 4.6.2), nel caso del ***che* polivalente** le tre stringhe sono risultate poco produttive soprattutto rispetto alle occorrenze effettive salvate. La stringa inerente al *che* esplicativo/consecutivo e al *che* causale ha consentito indubbiamente di restringere il campo d'indagine, ma si è rivelata decisamente poco produttiva, perché tra i risultati reperiti nel corpus si trovano soprattutto le frasi oggettive e le relative standard e su tre stringhe da 200 occorrenze sono state reperite solo 4 occorrenze del caso in questione.

La stringa relativa al *che* pseudorelativo, è l'unica per la quale, a causa del numero ridotto di risultati ottenuti dal corpus, non sono stati realizzati campioni da 200, bensì 2 campioni da 55 e uno da 57 unità, caratterizzati anche in questo caso da scarse occorrenze del caso in questione: 2/55, 6/55 e 7/57. La stringa relativa al *che* consecutivo-presentativo (con avverbi di dubbio) è risultata lievemente più produttiva, con valori pari a 23/200, 11/200 e 12/200.

1c) Differenza tra occorrenze iniziali scaricate e occorrenze finali salvate: seguono i casi in cui la variabilità dipende esclusivamente dalla differenza tra il totale iniziale di

occorrenze per ciascun campione e il totale delle occorrenze salvate al termine della ripulitura del campione dai casi non inerenti eccetera.

Per l'uso dell'**indicativo nelle frasi ipotetiche** è stata realizzata una sola stringa di ricerca, i cui risultati sono stati suddivisi in 3 campioni da 200. I totali delle occorrenze salvate da ciascun campione differiscono notevolmente dai totali iniziali, nello specifico: 17/200, 30/200 e 22/200.

Per l'uso di **per cui con valore conclusivo** è stata realizzata una sola stringa di ricerca, i cui risultati sono stati suddivisi in 3 campioni da 200. I totali delle occorrenze salvate da ciascun campione da 200 variano da un minimo di 92, ossia meno della metà del campione, ad un massimo di 157. Dunque, oltre alla differenza tra il totale delle occorrenze iniziali (3 campioni da 200) e i totali delle occorrenze finali salvate, si ipotizza che la variabilità possa dipendere anche dall'eterogeneità dei totali delle occorrenze finali salvate, che differiscono tra loro di 20-30 unità, nello specifico: 157/200, 114/200 e 92/200.

Per il fenomeno della **frase scissa nella sua variante retta da che** sono state realizzate 4 stringhe di ricerca, come già precedentemente descritto. Nel caso del verbo *essere* + infinito + *che* il numero esiguo di risultati ottenuti ha comportato la creazione di un campione unico da 7 unità. Per quanto riguarda, invece, le altre tre stringhe, sono stati realizzati 3 campioni da 200, ciascuna caratterizzata da maggiore o minore produttività: prima stringa: 31/200, 34/200 e 35/200; terza stringa: 120/200, 116/200 e 107/200; quarta stringa: 88/200, 90/200 e 83/200.

Come già osservato in alcuni casi descritti precedentemente, nel momento in cui i risultati di tutte le stringhe sono stati selezionati insieme per il calcolo della media e della deviazione standard, essi sono andati a costituire un unico macrocampione caratterizzato da una notevole variabilità, con valori relativi alle occorrenze salvate che variano da 1 a 120. A questo proposito, la scelta di costituire un unico campione finale si giustifica alla luce del fatto che le quattro stringhe di ricerca rappresentano 4 formulazioni sintattiche dello stesso fenomeno grammaticale, per il quale non è stato possibile progettare un'unica stringa che permettesse di reperire tutte le diverse formulazioni.

1d) Numero di campioni ricavati da ciascuna stringa: si citano in questa categoria i fenomeni, per i quali, il numero estremamente ridotto di occorrenze ricavate dal corpus ha comportato la creazione di un campione unico e non di 3 campioni: la tipologia verbo *essere* + infinito + *che* (esempio: *è studiare che è importante*) della frase scissa nella sua variante retta da *che*, per la quale è stato ottenuto un microcampione unico da 7 unità, la tipologia *c'* + verbo *avere* + *da* + verbo (esempio: *c'ho da fare*) con un microcampione da

6 unità e, infine, per i verbi in forma pronominale “affettiva” solo nel seconda persona plurale dei tempi non composti dei verbi (esempio: *vi mangiate un panino*) sono state reperite solo 17 occorrenze, che hanno costituito un campione unico.

2) Variazione casuale: è stato reperito un unico caso di variazione casuale al di fuori delle coppie analizzate, ossia l’uso di *ma* ad inizio frase, per il quale è stata prodotta una sola stringa, caratterizzata da un’alta produttività: le occorrenze salvate sono pari a 200/200, 197/200 e 199/200.

Queste sono le possibili cause della variazione riscontrate. Per contro, la media è stabile nei casi caratterizzati da una sola stringa di ricerca, i cui risultati sono distribuiti su tre campioni da 200 unità e in cui la differenza tra il totale iniziale di occorrenze per ciascun campione e il totale delle occorrenze salvate al termine della ripulitura dei casi non inerenti è molto ridotta. Tra i fenomeni che presentano poca variazione si citano a titolo di esempio i due casi presi in esame per ciascun terzile:

1. *Primo terzile, fenomeni a bassa frequenza:*

1.1. Soggetto in posizione rematica (*niente* seguito da un sostantivo plurale): una sola stringa di ricerca con tre campioni da 200, i cui totali dei risultati salvati variano da 160 a 176;

1.2. Dislocazione a sinistra (uso del doppio partitivo): una sola stringa di ricerca con tre campioni da 200, i cui totali dei risultati salvati variano da 161 a 174;

2. *Secondo terzile, fenomeni a media frequenza:*

2.1. la frase scissa retta da *a*: una sola stringa di ricerca con tre campioni da 200, i cui totali dei risultati salvati variano da 166 a 180;

2.2. l’uso dell’indicativo nella frase interrogativa indiretta: una sola stringa di ricerca con tre campioni da 200, i cui risultati salvati variano da 159 a 174;

3. *Terzo terzile, fenomeni ad alta frequenza:*

3.1. il partitivo preceduto da preposizione semplice: una sola stringa di ricerca con tre campioni da 200, i cui totali dei risultati salvati variano da 165 a 170;

3.2. l’uso dell’indicativo al posto del congiuntivo nella costruzione “verbo *essere* + aggettivo + *che* + indicativo” una sola stringa di ricerca con tre campioni da 200, i cui totali dei risultati salvati variano da 109 a 128.

5.3. Conclusioni sulla frequenza dei fenomeni indagati

Dall'analisi per ciascun fenomeno del numero di stringhe prodotte, del numero di campioni ottenuti per ciascuna stringa (3 campioni *vs* un campione unico), del totale iniziale di ciascun campione (200 unità o meno) e del totale delle occorrenze salvate per ciascun campione (ad esempio, 176 su 200) è emerso che presentano un valore elevato della deviazione standard stimata per 1 milione di parole diviso per il numero totale di token di PAISÀ in particolare quei fenomeni linguistici per i quali è stata realizzata più di una stringa di ricerca (ad esempio, 12 stringhe per il verbo in forma pronominale per indicare partecipazione affettiva, 4 per la frase scissa nella sua variante retta da *che*, 3 per il *che* polivalente e l'uso del *ci* attualizzante). In questi casi, i valori relativi alla media campione, deviazione standard campione ecc. sono il risultato di calcoli che hanno coinvolto collettivamente tutti i campioni ottenuti da tutte le stringhe di ricerca (a seconda dei casi 12, 4 o 3, come è stato descritto) elaborate per quel fenomeno, andando a costituire un unico campione caratterizzato da una notevole eterogeneità interna.

Alla luce di quanto esposto finora, si è valutato di fare un'analisi più approfondita su alcuni dei fenomeni che presentano un valore elevato della deviazione standard stimata per 1 milione di parole diviso per il numero totale di token di PAISÀ, scegliendone in maniera rappresentativa uno per ciascun terzile:

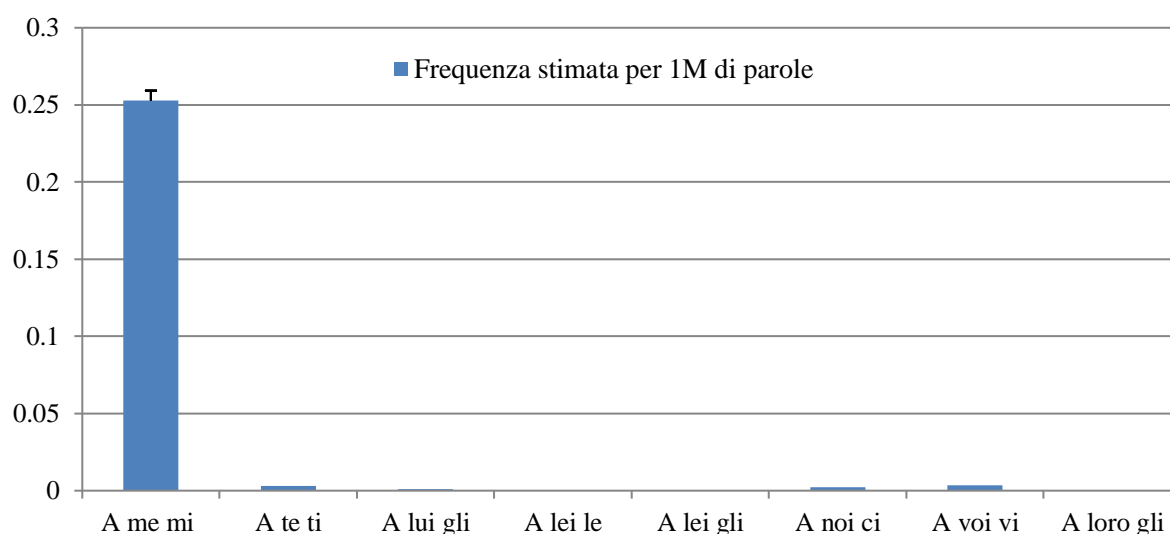
1. Dislocazione a sinistra dell'elemento pronominale (fenomeno a bassa frequenza);
2. Verbo in forma pronominale per indicare partecipazione affettiva (fenomeno a media frequenza);
3. Frase scissa retta da *che* (alta frequenza).

A differenza della fase di lavoro appena descritta, in cui le operazioni di calcolo utili ad un'analisi quantitativa³ erano state svolte selezionando globalmente tutti i dati relativi ai vari sottofenomeni grammaticali che rientrassero in un determinato fenomeno grammaticale (ad esempio, selezionando globalmente tutti i dati di frequenza relativi alle 8 stringhe di ricerca corrispondenti agli 8 sottofenomeni grammaticali della dislocazione a sinistra dell'elemento pronominale), in questa nuova fase di analisi tali operazioni di calcolo sono state svolte separatamente per ciascuno dei sottofenomeni grammaticali dei tre fenomeni in questione. Nello specifico:

³ Ossia: rapporto tra il numero di occorrenze salvate dopo l'eliminazione di doppioni e casi non inerenti e le 200 occorrenze iniziali del campione, media campione, deviazione standard campione, frequenza stimata, frequenza stimata per 1 milione di parole, deviazione standard campione stimata, deviazione standard stimata per 1 milione di parole diviso per il numero totale di token di PAISÀ.

1. **Fenomeno a bassa frequenza:** dislocazione a sinistra dell'elemento pronominale; sottofenomeni: *a me mi; a te ti; a lui gli; a lei le; a lei gli; a noi ci; a voi vi; a loro gli*.
2. **Fenomeno a media frequenza:** verbo in forma pronominale per indicare partecipazione affettiva. 12 stringhe di ricerca: 6 stringhe corrispondenti alle 6 forme pronominali riflessive (*mi, ti, si, ci, vi, si*) in presenza di tempi non composti dei verbi (esempi: *mi bevo/berrò/bevvi... ecc. il caffè*) e altre 6 stringhe per i casi dei tempi composti dei verbi (esempi: *mi sono bevuto/ero bevuto/ mi sarei bevuto... ecc. il caffè*).
3. **Fenomeno ad alta frequenza:** frase scissa retta da *che*:⁴
 - 3.1. verbo *essere* + gruppo nominale + *che* + verbo coniugato;
 - 3.2. verbo *essere* + avverbio + *che* + verbo coniugato;
 - 3.3. verbo *essere* + gruppo nominale preceduto da preposizione + *che* + verbo coniugato.

Grafico 5.4: Dislocazione a sinistra dell'elemento pronominale



⁴ La stringa relativa al sottofenomeno “verbo *essere* + verbo all’infinito + *che* + verbo coniugato” (esempi: *è restare che dobbiamo fare; è votare che è importante* ecc.) è stata esclusa dall’analisi, poiché su 7 occorrenze scaricate dal corpus solo una corrisponde al fenomeno in esame.

Grafico 5.5: Verbo in forma pronominale

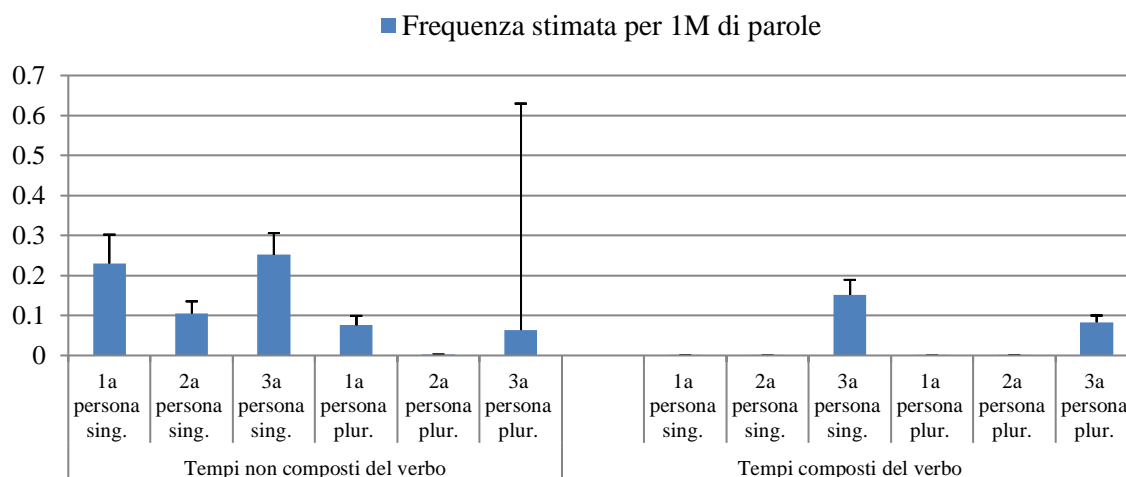
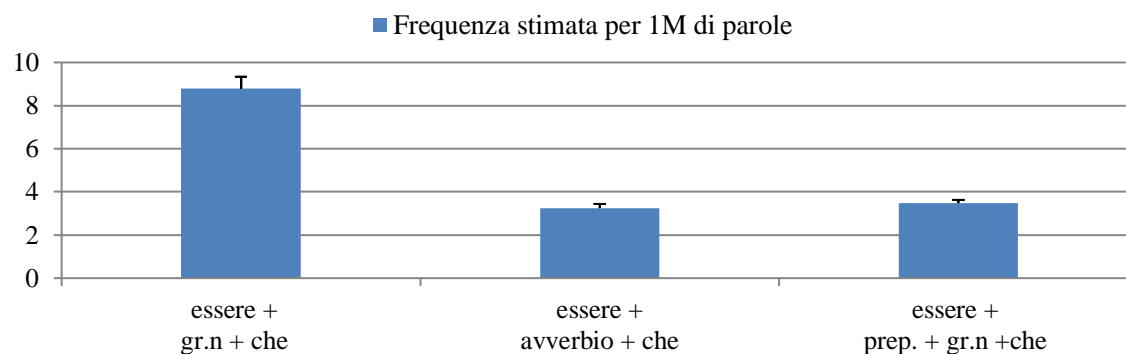


Grafico 5.6: Frase scissa retta da “che”



Di seguito nelle tabelle 5.4, 5.5 e 5.6 si riportano i dati ottenuti dall’approfondimento dell’analisi sui fenomeni in questione:

Tabella 5.4: Approfondimento di un fenomeno a bassa frequenza ($0 < fq < 3$)

Dislocazione a sinistra: ripetizione del pronome in forma atona e tonica. Segue in corsivo il dettaglio dei sottofenomeni:	Frequenza stimata per 1M di parole ⁵	Deviazione standard stimata per 1M di parole	N.ro di campioni
<i>A me mi</i>	0,25	< 0,01	3
<i>A te ti</i>	< 0,01	-	1
<i>A lui gli</i>	< 0,01	-	1
<i>A lei le</i>	< 0,01	-	1
<i>A lei gli</i>	0	-	1
<i>A noi ci</i>	< 0,01	-	1
<i>A voi vi</i>	< 0,01	-	1
<i>A loro gli</i>	< 0,01	-	1

Tabella 5.5: Approfondimento di un fenomeno a media frequenza ($3 < fq < 20$)

Verbo in forma pronominale (es. <i>io mi bevo un caffè</i>). Segue in corsivo il dettaglio dei sottofenomeni: <i>Tempi non composti del verbo</i>	Frequenza stimata per 1M di parole ⁶	Deviazione standard stimata per 1M di parole	N.ro di campioni
<i>1a persona sing.</i>	0,23	0,07	3
<i>2a persona sing.</i>	0,10	0,03	3
<i>3a persona sing.</i>	0,25	0,05	3
<i>1a persona plur.</i>	0,07	0,02	3
<i>2a persona plur.</i>	0	-	1
<i>3a persona sing.</i>	0,06	0,56	3
<i>Tempi composti del verbo</i>			
<i>1a persona sing.</i>	< 0,01	-	1
<i>2a persona sing.</i>	< 0,01	-	1
<i>3a persona sing.</i>	0,15	0,03	3

⁵ NB: Il numero estremamente ristretto di occorrenze reperite per alcuni sottofenomeni grammaticali, quali *a te ti*, *a lui gli*, *a lei le*, *a lei gli*, *a noi ci*, *a voi vi* e *a loro gli* ha comportato la creazione di un campione unico per ciascuno di essi (cfr. sez. 5.1), rendendo, quindi, impossibile il calcolo della media e relativa deviazione standard. Per questi sottofenomeni è stato dapprima calcolato il rapporto tra il numero di occorrenze salvate dopo l'eliminazione di eventuali casi non inerenti e dopponi e le occorrenze iniziali del campione; successivamente questo valore è stato normalizzato per 1 milione di parole (ossia, moltiplicato per 1.000.000 e diviso 265.062.636).

⁶ Per i sottofenomeni del verbo in forma pronominale alla 2a persona plurale dei tempi non composti del verbo e alla 1a e 2a persona singolare e alla 1a e 2a persona plurale dei tempi composti del verbo valgono le stesse considerazioni fatte nella nota precedente.

<i>1a persona plur.</i>	< 0,01	-	1
<i>2a persona plur.</i>	< 0,01	-	1
<i>3a persona plur.</i>	0,08	0,01	3

Tabella 5.6: Approfondimento ad alta frequenza (fq > 20)

Frase scissa con <i>che</i> Segue in corsivo il dettaglio dei sottofenomeni:	Frequenza stimata per 1M di parole	Deviazione standard stimata per 1M di parole	N.ro di campioni
<i>Verbo "essere" + gruppo nominale + che</i>	8,78	0,54	3
<i>Verbo "essere" + avverbio + che</i>	3,25	0,18	3
<i>Verbo "essere" + preposizione + gruppo nominale + che</i>	3,48	0,14	3

Questo approfondimento ha permesso di identificare le cause della variazione all'interno dei campioni dei fenomeni grammaticali presi in esame e ha mostrato come la variazione si sia notevolmente ridotta dopo che le operazioni di calcolo utili ad un'analisi quantitativa sono state svolte separatamente per ciascuno dei sottofenomeni grammaticali dei fenomeni indagati e non più selezionando globalmente tutti i dati relativi ai sottofenomeni grammaticali.

Osservando le tabelle 5.1-5.3 si può notare che vi sono ancora due soli fenomeni che presentano un valore della deviazione standard elevato rispetto alla frequenza, di cui finora non erano stati presi in esame i fattori di variabilità: l'uso del pronome *gli vs le/loro* e il soggetto in posizione rematica rispetto al predicato nelle frasi principali. Nel primo caso si ipotizza che la variabilità all'interno del campione dipenda dalla scarsissima produttività della stringa di ricerca, per la quale quasi la totalità delle occorrenze reperite nel corpus è stata scartata dalle occorrenze da salvare, nello specifico: 9/200, 7/200 e 7/200. Invece, il soggetto in posizione rematica rispetto al predicato nelle frasi principali presenta tutti i fattori finora presi in considerazione per la variabilità non casuale interna al campione: la ricerca di questo fenomeno grammaticale si basa, infatti, su 5 stringhe, finalizzate a reperire le varie tipologie di predicato coinvolto in questo fenomeno sintattico (ossia: 1) verbo non composto; 2) verbo composto con ausiliare *essere*; 3) verbo composto con ausiliare *avere* con l'oggetto tra il participio e il soggetto; 4) verbo composto con ausiliare *avere* (senza oggetto); 5) verbo passivo). I risultati della stringa dedicata al verbo composto con ausiliare *avere* con l'oggetto tra il participio e il soggetto costituiscono un campione unico (15 occorrenze su 38), mentre per le restanti 4 stringhe di ricerca sono stati creati 3

campioni. Dal punto di vista della produttività delle stringhe, per 3 di queste i totali delle occorrenze salvate sono inferiori ai totali iniziali scaricati dal corpus (verbi non composti: 52-60-58 su 200; verbi composti con ausiliare *essere*: 68-68-65 su 200; verbi passivi: 53-51-52 su 200); in particolare per la stringa dei verbi con ausiliare *avere* senza l'oggetto tra participio e verbo sono state reperite meno occorrenze, che non hanno permesso di realizzare 3 campioni da 200, bensì da 2 campioni da 137 e uno da 136, caratterizzati da una scarsa produttività (nello specifico: 17/137, 22/137 e 8/136).

A questo punto è necessario spiegare il motivo per cui per questo fenomeno, così come per il *che* polivalente si è deciso di non ricalcolare separatamente i risultati di ciascuna stringa, come è stato fatto, invece, per la frase scissa retta da *che*, il soggetto pronominale e la dislocazione a sinistra dell'elemento pronominale, ma di conteggiarli collettivamente in un unico macrocampione eterogeneo.

Innanzitutto, dal punto di vista grammaticale, non vi erano ragioni per separare le occorrenze: sia per il *che* polivalente che per il soggetto in posizione rematica rispetto al predicato nelle frasi principali le diverse stringhe di ricerca prodotte non identificano dei veri e propri sottofenomeni diversi tra loro, né dei cambiamenti morfologici all'interno di un fenomeno di per sé unitario (ad esempio, l'intera flessione dei verbi riflessivi con soggetto pronominale "affettivo": *mi mangio, ti mangi* ecc.).

Inoltre, data la complessità di tali fenomeni che investono la sintassi della frase, si è appurato nella fase di progettazione delle stringhe in linguaggio CQL che una singola stringa non avrebbe permesso di ottenere le occorrenze di tutti gli ambiti d'uso dei due fenomeni in questione, dunque era inevitabile la progettazione di più di una stringa, i cui risultati andassero conteggiati collettivamente per le ragioni sopra spiegate.

La riflessione finale, derivante da quanto esposto finora in questo capitolo, è che i valori della frequenza e della deviazione standard dipendono massicciamente dalla formulazione delle stringhe di ricerca, la quale a sua volta è parzialmente determinata dalle opzioni a disposizione nel linguaggio CQL. Ne consegue che alcune stringhe di ricerca sono più precise, mentre altre lo sono meno. Come già illustrato nel capitolo 4 (cfr. sezioni 4.6.1 e 4.6.2, dedicate al linguaggio CQL e alla progettazione delle stringhe), alcuni tratti grammaticali sono semplicemente più facilmente ricercabili tramite linguaggio CQL rispetto ad altri, perché legati a fenomeni sintattici e morfologici basati su delle sequenze fisse di elementi grammaticali che co-occorrono. Esemplicando, un conto è ricercare la ricorrenza di *a me mi*, o di un verbo di opinione seguito sempre da un verbo al congiuntivo, ben altro è riuscire ad inquadrare un fenomeno come il *che* polivalente, che risulta più

“evanescente” dal punto di vista sintattico e distribuzionale, dato che non ci sono categorie morfosintattiche specifiche con cui co-occorre.

A questo proposito vale la riflessione fatta da Budassi (2017: 157) circa la reperibilità del *che* polivalente, applicabile anche ad altri fenomeni grammaticali la cui reperibilità risulta complessa: “La ricerca del ‘che’ polivalente in un corpus taggato unicamente per *part of speech* è difficoltosa. Il ‘che’ polivalente non co-occorre (in maniera univoca) con una qualunque altra parte del discorso con un alto livello di probabilità”. Pertanto risultati significativi possono essere ottenuti “non solo ricercando quanto normalmente è presente nel contesto di un ‘che’ polivalente, ma anche escludendo tutto ciò che è probabilmente presente nel contesto di un ‘che’ di altra natura” (*ibidem*). Al termine di queste riflessioni sulla reperibilità e sulla frequenza di tali fenomeni nel corpus, la sezione successiva sarà dedicata alla selezione dei fenomeni grammaticali sui quali costruire le unità didattiche.

5.4. I fenomeni grammaticali selezionati

Nella selezione dei fenomeni grammaticali sui quali incentrare le unità didattiche si è tenuto conto non solo della loro frequenza all’interno del corpus, descritta nelle sezioni precedenti, ma anche della rilevanza sul piano glottodidattico.

Per quanto riguarda la frequenza, come già anticipato nell’introduzione del capitolo, è importante considerarla tra i criteri di scelta per le seguenti ragioni. Innanzitutto consente di individuare quali tra i tratti linguistici che all’epoca della pubblicazione degli studi di Sabatini erano ancora percepiti come in contrasto con la norma grammaticale si siano gradualmente inseriti nell’uso linguistico di media formalità/informalità. Tracciare questa panoramica dei fenomeni grammaticali più utilizzati dagli italofoeni nativi permette di avere la percezione di quali di essi è più probabile che i lettori apprendenti di lingua italiana L2/LS trovino nei testi scritti e, quindi, su quali di essi sia più rilevante progettare dei materiali didattici.

Tuttavia, la sola frequenza non è sufficiente come criterio di selezione, poiché nella produzione dei materiali didattici è necessario tenere conto anche di molti altri fattori, ad esempio dei processi cerebrali alla base dell’acquisizione linguistica, quali ad esempio la bimodalità emisferica, il cosiddetto “ordine naturale” e le sequenze di acquisizione (cfr. capitolo 4), da cui non si può prescindere nella presentazione dei contenuti linguistici agli apprendenti.

Ad esempio, se nell'ambito di questa ricerca si prendesse in considerazione il solo criterio della frequenza nel selezionare i fenomeni grammaticali per le unità didattiche, emergerebbe che la frase scissa retta da *che* (esempio: *è lui che l'ha detto*) è risultata il secondo fenomeno più frequente nella consultazione del corpus. Nel caso in cui si volesse progettare dei materiali didattici su questo tema, sarebbe indispensabile tenere conto del fatto che l'uso di questa tipologia della frase scissa presuppone delle conoscenze di morfosintassi (pronomi relativo *che*, costruzione della frase, ordine marcato e non marcato dei costituenti ecc.) che non possono essere date per scontate prima di introdurre l'argomento. Infine, oltre all'ordine naturale di acquisizione è necessario tenere conto anche della rilevanza sul piano glottodidattico dei diversi fenomeni indagati.

Per quanto riguarda quest'ultimo aspetto, infatti, è stato necessario determinare quali sono i tratti del neostandard che non siano già stati esaminati nella manualistica per stranieri attualmente in commercio e per i quali, quindi, potesse essere rilevante e utile creare dei materiali didattici.

A questo proposito merita una segnalazione il manuale di lingua italiana per stranieri di Zamora, Alessandro e Hernández González (2010), realizzato con l'obiettivo di far fronte alla scarsità di materiali didattici che includano i tratti dell'italiano colloquiale e di cui si dà una descrizione nello studio di Zamora e Alessandro (2011)⁷. Il manuale ha il pregio di offrire un punto di vista innovativo rispetto allo standard linguistico proposto e la sua consultazione è risultata utile sul piano metodologico. Tuttavia, l'opera si focalizza in maniera specifica sul registro colloquiale, che come descritto da Berruto (1987: 139-140), costituisce il nucleo principale dell'italiano substandard e condivide alcune caratteristiche del neostandard, ma non coincide con quest'ultimo, dato che solo alcune sue forme e strutture meno marcate in termini di informalità sono passate e continuano a entrare a far parte dell'italiano dell'uso medio. Dunque, pur apprezzandone l'impostazione e gli obiettivi, è stato escluso dall'analisi a cui si farà riferimento nei paragrafi successivi, perché il suo obiettivo differisce da quello della presente ricerca, incentrata sul neostandard.

Pertanto, la riflessione sulla manualistica per stranieri attualmente in commercio è stata guidata dalla consultazione di uno studio condotto da Słapek (2017), che evidenzia

⁷ In questo studio si pone l'accento più volte sul fatto che esistano pochi materiali mirati all'insegnamento e apprendimento colloquiale in classe: "Si miramos a la didáctica del italiano como lengua extranjera (I/LE) o lengua segunda (I/L2), observamos que son escasos los materiales que reflejan esta pluralidad y, especialmente, los que se plantean y abordan de forma sistemática la enseñanza y el aprendizaje del coloquial en el aula" (versione online, priva di numeri di pagina).

come su 32 pubblicazioni delle maggiori case editrici del settore della didattica dell'italiano per stranieri, quali AlmaEdizioni, CDL Edizioni, Edilingua, Eli, Guerra Edizioni, Hoepli, Le Monnier e Loescher, solo 2 affrontano alcuni tratti dell'italiano dell'uso medio: "Nuovo Magari C1- C2" della casa editrice Alma Edizioni, che tratta la frase scissa nel volume dedicato al livello C1, e "Affresco italiano C2" della casa editrice Le Monnier, che affronta la dislocazione a sinistra e a destra, la frase scissa, l'uso del *c'è* presentativo, l'uso di *lui, lei e loro* in funzione di soggetto, l'uso di *gli* con i valori di *a loro* e *a lei*, l'uso di *cosa* come pronome interrogativo in concorrenza con *che* e *che cosa* (dati tratti da Słapek 2017).

Una volta esclusi questi tratti già presenti nei due manuali sopra menzionati, si è proceduto alla selezione tra i fenomeni grammaticali rimanenti nella lista di Sabatini (1990) valutando caso per caso la rilevanza a livello didattico della produzione di materiali didattici incentrati su di essi.

Esaminando i fenomeni ad alta e media frequenza nel corpus, si è fatta strada l'idea di creare due unità didattiche che potessero collocarsi a cavallo tra più livelli del Quadro Comune Europeo di Riferimento per la conoscenza delle lingue (QCER) (Consiglio d'Europa 2018, 2002) e che permettessero di coprirne delle grandi aree tematiche.

Per tale ragione, l'attenzione è stata rivolta all'uso dell'indicativo al posto del congiuntivo (che include sottofenomeni ad alta, media e bassa frequenza nel corpus) e all'uso del verbo in forma pronominale riflessiva in senso affettivo (es. io *mi* mangio un panino; fenomeno a media frequenza nel corpus).

In entrambi i casi la scelta si è basata sia sulla frequenza del fenomeno, sia sulla potenziale produttività dell'argomento ai fini di un'unità didattica: circa quest'ultimo punto, come illustrato dai principali sillabi di riferimento della lingua italiana (cfr. capitolo 6, sez. 6.2), la trattazione dell'uso standard dei pronomi (e più nello specifico dei pronomi riflessivi) e dell'uso del congiuntivo si dispiega su più livelli del QCE. Ne consegue che gli argomenti vengano affrontati più volte nel percorso di apprendimento, sia nella fase di primo approccio con il tema, sia negli approfondimenti dei livelli successivi e che, quindi, dei temi così articolati possano rappresentare un oggetto di interesse, nonché una fonte di dubbi sull'uso della lingua, su più livelli, al progredire della competenza linguistica. Si ritiene, pertanto, che la produzione di materiali didattici su questi temi abbia una maggiore rilevanza rispetto alla produzione relativa ad un singolo argomento specifico collocato in un solo livello.

A questo proposito, ad esempio, Zamora e Alessandro (2011) in riferimento al già citato manuale sull'italiano colloquiale affermano che l'insegnamento/apprendimento del colloquiale debba essere svolto in modo graduale ma progressivo, unito allo studio dei tratti dello standard e del neostandard e legato alle situazioni o ai contesti comunicativi che li richiedono⁸.

Inoltre, si è considerato in prospettiva futura che proporre argomenti distribuibili su più livelli linguistici avrebbe permesso alle unità didattiche prodotte di integrarsi più agevolmente nella programmazione didattica dei dipartimenti coinvolti nella sperimentazione, a differenza di altri argomenti (è il caso, ad esempio, della frase scissa e delle dislocazioni) che, come già accennato, necessitano di conoscenze preliminari di sintassi e ordine dei costituenti che sarebbero state difficilmente appurabili prima della sperimentazione e del contatto con i gruppi-classe (contrariamente a quanto avverrebbe se per la sperimentazione delle unità si avesse a disposizione un intero corso di lingua). Altri fenomeni ad alta/media frequenza nel corpus, quali l'uso di *allora/comunque/ma* ad inizio frase e l'uso dell'interrogativo *cosa* al posto di *che cosa*, sono stati scartati perché giudicati poco produttivi dal punto di vista della realizzazione di un'intera unità didattica.

Per queste ragioni sono stati selezionati i tratti linguistici sopra menzionati. La descrizione della struttura delle unità didattiche prodotte e degli esiti della loro sperimentazione sarà oggetto del capitolo 6.

⁸ “Desde nuestro punto de vista, estimamos que la enseñanza-aprendizaje del coloquial debe ser llevada a cabo de forma paulatina pero progresiva, combinada con el estudio de los rasgos del estándar y neoestándar y vinculada con las situaciones o contextos comunicativos en las que se requieran y se exijan” (versione online, priva di numeri di pagina).

CAPITOLO VI

PROGETTAZIONE E SPERIMENTAZIONE DELLE UNITÀ DIDATTICHE

Nei capitoli precedenti è stato illustrato l'approccio metodologico adottato per l'identificazione *corpus-based* dei fenomeni e per la scelta sia dei testi da includere nelle unità didattiche, sia della piattaforma di *e-learning* sulla quale collocare le attività (capitolo 4). Inoltre, tra i fenomeni quantitativamente più rilevanti sono stati identificati quelli su cui progettare le unità didattiche (capitolo 5).

Sulla base di quanto precedentemente esposto, in questo capitolo verranno brevemente richiamati i due tratti linguistici scelti per le unità, ossia l'uso del verbo in forma pronominale riflessiva in senso affettivo e l'uso dell'indicativo *vs* congiuntivo (cfr. cap. 2).

Successivamente, verranno descritte le attività inserite nelle unità didattiche (sez. 6.2) e i loro esiti nelle sperimentazioni svolte presso i dipartimenti di italianistica di alcuni atenei europei, a cui si farà riferimento nella sezione 6.3.

Infine, alla sezione 6.4 verranno presentati gli esiti del questionario per la raccolta delle opinioni degli studenti sulle attività svolte.

6.1. I tratti linguistici selezionati

Come esposto nel capitolo 5, nella selezione dei fenomeni grammaticali sui quali incentrare le unità didattiche si è tenuto conto non solo della loro frequenza all'interno del corpus, ma anche della rilevanza sul piano glottodidattico, ossia per quali tratti linguistici potesse essere rilevante e utile creare un approfondimento sull'uso neostandard nell'ambito della didattica dell'italiano per stranieri.

Pertanto, i tratti linguistici selezionati attraverso queste valutazioni sono l'uso del verbo in forma pronominale riflessiva in senso affettivo (es. *io mi mangio un panino*), a cui in seguito per brevità si farà riferimento come verbi riflessivi "affettivi" e l'impiego dell'indicativo al posto del congiuntivo, la cui descrizione è stata affrontata nel capitolo 2 e di cui ora si riprenderanno solo alcuni punti salienti.

Per quanto riguarda il primo tratto, la "costruzione riflessiva apparente o di affetto" (Sabatini 1985: 167) è utilizzata per indicare la partecipazione o il coinvolgimento emotivo di una persona rispetto a un'azione o a una circostanza indicata dal predicato. Il suo uso è frequente con i verbi *mangiare* e *bere* (e loro sinonimi) e con altri verbi che esprimono

azioni o atteggiamenti che hanno effetti sulla persona del soggetto, anche con riferimento a parti del corpo o a indumenti che riguardano la sua persona. Si citano alcuni esempi tratti da Sabatini (*ibidem*) e Cignetti 2010: *Luca si è mangiato mezza torta; verso le 11 mi bevo un caffè; Valerio si vede un film in televisione; Aldo si gode la vacanza; stamattina mi sono fatto una splendida passeggiata; farsi quattro risate* e casi simili.

La costruzione “riflessiva affettiva” è composta, dunque, da un verbo a cui si accompagna un pronome atono che “serve a enfatizzare la partecipazione emotiva del parlante, come effetto del riferimento deittico al soggetto enunciativo” (Cignetti 2010).

Per quanto riguarda, invece, l’alternanza del congiuntivo con l’indicativo, Sabatini (1985, 1990) la segnala nei seguenti cinque casi: dopo i verbi di opinione (*credo che hai ragione*), dopo una dichiarativa negativa (*non credo che hai torto*); in un’interrogativa indiretta (*gli chiesi se poteva aiutarmi*); nel periodo ipotetico dell’irrealtà (*se me lo dicevi, ci pensavo io*); nelle relative restrittive (*sei l’unico che parla arabo*).

Come già anticipato nel capitolo 2, questa ricerca si riallaccia a quanto espresso da Prandi (2002), secondo il quale il congiuntivo ha un suo valore solo nei casi in cui il suo uso si basa su una scelta libera del parlante, che può alternarlo all’indicativo per cambiare il registro stilistico della frase. Dunque, l’opposizione indicativo-congiuntivo non andrebbe più ricondotta alla sola opposizione tra codifica della realtà e dell’irrealtà, secondo la posizione tradizionale (si veda la già citata interpretazione semantica, con l’indicativo come “modo della certezza” e congiuntivo come “modo dell’incertezza” cfr. Schneider 1999), bensì bisognerebbe tenere conto anche dei casi in cui la sua scelta è un’opzione del parlante e, di conseguenza, il suo valore si esprime nel contenuto del messaggio (cfr. i già citati Lepschy e Lepschy 1988: 204 e Renzi 2012: 52-53), secondo i quali la scelta tra congiuntivo e indicativo è di tipo stilistico).

Dopo aver richiamato brevemente i tratti salienti dei fenomeni grammaticali sui quali sono state incentrate le unità, nella sezione seguente si illustrerà la struttura delle unità didattiche prodotte e gli esiti della loro sperimentazione presso il Dipartimento di Italianistica dell’Università di Breslavia (Polonia) dal 27 maggio al 31 luglio 2019 e presso i seguenti dipartimenti spagnoli: Dipartimento di Studi Romanzi, Francesi, Italiani e Traduzione dell’Università Complutense di Madrid dal 13 al 30 settembre 2019; Dipartimento di Filologia Francese, Romanza, Italiana e Araba dell’Università di Murcia

per tutto il mese di ottobre 2019 e, infine, il Dipartimento di Traduzione e Interpretazione dell'Università di Malaga dal 13 al 24 gennaio 2020¹.

6.2. Descrizione delle unità

Di seguito verranno descritte le due unità didattiche, per i cui fondamenti teorici si rinvia al capitolo 4. Il punto di partenza nella progettazione è stata l'individuazione all'interno del Quadro Comune Europeo di Riferimento per la conoscenza delle lingue (QCER) del livello linguistico rispetto al quale progettare le unità didattiche, con i relativi obiettivi comunicativi e pragmatici. Il QCER è un sistema descrittivo elaborato dal Consiglio d'Europa (2018, 2002) nell'ambito del progetto *Language Learning for European Citizenship*, adottato dal 2001 sia per valutare, ed eventualmente certificare, le abilità conseguite dagli apprendenti di una lingua, sia per programmare il livello di un insegnamento linguistico. Il QCER identifica tre ampie fasce di competenza a loro volta suddivise in due livelli ciascuna, per un totale di sei livelli complessivi:

A – livello base

- A1 – Livello di Contatto o Livello Base
- A2 – Livello di Sopravvivenza o Livello Elementare

B – livello autonomia

- B1 – Livello Soglia o Livello Intermedio Basso
- B2 – Livello di Progresso o Livello Intermedio Alto

C – livello padronanza

- C1 – Livello di Efficacia o Livello Avanzato
- C2 – Livello di Padronanza o Livello Competenza

Il QCER ha lo scopo di definire la competenza linguistico-comunicativa per i diversi livelli in riferimento alle abilità di ricezione e di produzione scritte e orali e alla conoscenza delle strutture linguistiche, del lessico, dei tratti della testualità, della pragmatica e della sociolinguistica; tuttavia, esso non fornisce indicazioni sulle strutture grammaticali richieste per i diversi livelli. Tali informazioni possono, invece, essere ricavate consultando i sillabi di riferimento della lingua italiana, ossia dei testi che propongono degli elenchi delle strutture grammaticali richieste per ciascun livello di apprendimento linguistico del QCER. I sillabi sono compilati principalmente da enti

¹ L'ultima fase di sperimentazione, prevista per il periodo di maggio-giugno 2020 presso l'Università della Slesia (Sosnowiec, Polonia), è stata cancellata in seguito alle misure di contenimento dell'epidemia di Covid-19.

certificatori e centri di formazione linguistica e sono “parte dell’attività curricolare che si riferisce alla specificazione e alla sequenziazione dei contenuti di insegnamento fatta in termini di conoscenze e/o capacità” (Ciliberti, 1994: 99-100). Nello specifico, il syllabo di certificazione “elenca le competenze che il candidato deve possedere come acquisite, e saper utilizzare in modo produttivo, per poter essere considerato a un certo livello di competenza” (Arcangeli *et al.*, 2014: 14).

Pertanto, al fine di identificare innanzitutto in quali livelli vengano collocati l’uso standard dei verbi riflessivi e del congiuntivo e, in secondo luogo, se vi siano delle indicazioni sull’uso dei verbi riflessivi in forma “affettiva” e sulla scelta tra indicativo e congiuntivo in contesti informali, sono stati consultati sillabi mirati alla programmazione di un corso di lingua o alla certificazione per la lingua italiana, di seguito elencati seguendo l’ordine della loro data di pubblicazione: Minciarelli, Comodi (2005), Lo Duca (2006), Barni *et al.* (2009), Benucci (2007 a c. di), Spinelli, Parizzi (2010 a c. di), Arcangeli *et al.* (2014), AIL: Accademia Italiana di Lingua (2015) e infine Cardillo, Vecchio (2015, 2016, 2017, 2018: a c. di).

La tabella 6.1 permette di avere un quadro generale dei livelli del QCER nei quali gli autori dei sillabi consultati inseriscono i tratti grammaticali in oggetto. Sotto ciascun livello sono riportate delle abbreviazioni corrispondenti ai vari sillabi, seguite da un numero tra parentesi, corrispondente alla ricorrenza della trattazione di tale argomento: dalla sua prima introduzione (1) ai successivi approfondimenti (esempi: (2) e (3)). Segue una legenda delle abbreviazioni per ciascun sillabo: **MC**: Minciarelli, Comodi (2005); **LD**: Lo Duca (2006); **Ba**: Barni *et al.* (2009); **Be**: Benucci (2007 a c. di); **SP**: Spinelli, Parizzi (2010 a c. di); **ADA**: Arcangeli *et al.* (2014), **AIL**: Accademia Italiana di Lingua (2015) e infine **PLIDA**: Cardillo, Vecchio (2015a, 2015b, 2016, 2017, 2018 a c. di).

Tabella 6.1: Livelli del QCER in cui gli autori dei sillabi inseriscono i tratti grammaticali in oggetto

	A1	A2	B1	B2	C1	C2
Pronomi riflessivi (uso standard)	SP (1) Be (1) ADA (1) AIL (1)	MC (1) LD (1) Be (2) SP (2) ADA (2) AIL (2)	MC (2) Ba (1) Be (3) ADA (3)	SP (2) Ba (2) PLIDA (1)		
Pronomi riflessivi “affettivi”			MC (1) Be (1)	LD (2)		
Congiuntivo (uso standard)		MC (1)	MC (2) LD (1) SP (1) Be (1) PLIDA (1) ADA (1) AIL (1)	LD (2) SP (2) Ba (1) Be (2) PLIDA (2) ADA (2) AIL (2)		
Indicativo vs congiuntivo nell’informalità			LD (1) SP (1)		MC (1) Be (1)	

Dalla tabella 6.1 emerge che 4 sillabi su 8 suggeriscono una prima introduzione nel livello A1 del tema dei pronomi riflessivi, che viene sviluppato appieno nel livello A2, con successivi approfondimenti nei livelli B1-B2. Le uniche eccezioni sono i casi di Barni (2009) e del PLIDA, che menzionano i riflessivi rispettivamente solo al livello B1 e B2. Per quanto riguarda i pronomi riflessivi “affettivi”, vengono menzionati solo in tre sillabi: Minciarelli, Comodi (2005) e Benucci (2007 a c. di) inseriscono questo argomento al livello B1, mentre Lo Duca (2006) lo inserisce al livello B2, come prosecuzione dell’uso standard dei riflessivi trattato al livello A2.

Per quanto riguarda il congiuntivo, ad eccezione di Minciarelli e Comodi (2005), che propongono una prima introduzione delle strutture del congiuntivo nel livello A2, gli altri sillabi propongono l’introduzione delle forme del presente e del passato nel livello B1, per poi proseguire nel livello B2 con il congiuntivo imperfetto e trapassato, la concordanza dei tempi e il periodo ipotetico. In alcuni casi, tali ulteriori sviluppi vengono inseriti nei livelli C1 o C2; tuttavia, in questa sede sono state omesse queste informazioni dettagliate, poiché lo scopo di tale confronto è determinare il livello in cui le strutture vengono introdotte per la prima volta.

Per quanto riguarda, invece, la scelta dell’indicativo al posto del congiuntivo in contesti informali, il tema viene proposto soltanto da quattro sillabi, che si dividono in

maniera equa tra l'introduzione di questo argomento al livello B1 (Lo Duca 2006, Spinelli, Parizzi 2010) e al livello C1 (Minciarelli, Comodi 2005, Benucci 2007 a c. di).

Pertanto, sulla base di tale confronto, le unità sono state progettate per essere adatte globalmente al livello B1. Nel caso del congiuntivo la trattazione è stata limitata ai tempi del presente e del passato, introducendo l'uso neostandard dell'indicativo al posto del congiuntivo come integrazione delle norme dell'uso standard. Invece, nel caso dei verbi riflessivi "affettivi" sono stati inseriti alcuni collegamenti con il livello A2, proponendo attività in cui si richiama la coniugazione dei verbi e la posizione del pronome clitico nella frase, che possono fungere da primo approccio per il livello A2 o da rinforzo o recupero per il livello B1. Sarà dunque possibile scegliere se svolgere tutte le attività proposte o se creare dei percorsi personalizzati adattabili al livello della classe. Anche i contenuti tematici delle unità sono stati scelti tenendo conto delle indicazioni per i livelli B1-B2 del QCER e dei vari sillabi.

Per facilitare i processi cognitivi degli apprendenti, è necessario elaborare delle tecniche che tengano conto dei meccanismi alla base dell'apprendimento delle lingue (cfr. Daloiso 2009; 2015) e, che in tal modo, favoriscano l'acquisizione linguistica. A questo proposito, si cita qui di seguito il percorso proposto da Balboni (2002:137-138):

- a) *individuazione*: nei testi che costituiscono input gli allievi sottolineano, cerchiano, ricopiano ecc. le «regole» che costituiscono l'obiettivo di quell'unità di apprendimento;
- b) *creazione di ipotesi*: sul funzionamento di quel dato meccanismo: ipotesi suggerite dal docente o guidate da uno schema del manuale e discusse eventualmente con i compagni;
- c) *fissazione*, sia attraverso attività di natura comportamentistica, quindi abbastanza ripetitive, sia per mezzo di giochi che, pur condividendo la natura del *pattern drills*, non ne abbiano le debolezze;
- d) *reimpiego* via via più creativo, dapprima con esercizi di applicazione e poi in attività comunicative;
- e) *riflessione* esplicita guidata dal docente: e mirata alla creazione di una rappresentazione mentale e della sua eventuale «trasposizione» in uno schema grafico, una tabella ecc.

Queste indicazioni sono state seguite nella progettazione delle unità didattiche, secondo lo schema già citato nella sezione 4.8.3 del capitolo 4. Di seguito verranno descritte entrambe le unità in maggior dettaglio.

6.2.1. Unità didattica sui verbi riflessivi “affettivi”

Come anticipato nella sezione precedente, le attività sono state progettate per poter essere adatte sia ad un primo approccio con il tema grammaticale dell’uso dei riflessivi, nel caso del livello A2 - B1, sia come attività di ripasso per i livelli superiori. Come già anticipato, i testi utilizzati, intesi sia come interi brani, sia come singole frasi, sono tratti dal corpus PAISÀ, mentre le immagini utilizzate come illustrazione delle attività sono state scaricate da siti che mettono a disposizione immagini prive di copyright, gratuite e liberamente utilizzabili anche a fini commerciali².

Richiamando le basi teoriche descritte nel capitolo 4, le unità sono state strutturate secondo le seguenti fasi di apprendimento, che permettono agli apprendenti di approcciarsi con le strutture linguistiche in maniera induttiva: motivazione, globalità, analisi, sintesi, riflessione e verifica. Le unità in forma completa sono disponibili in appendice (Appendice 4), mentre nella descrizione delle attività verrà fornita un’immagine esplicativa.

Come già anticipato nel capitolo 4, sono state collocate sulla piattaforma *Wizer.me*, che consente un ampio margine di personalizzazione nella progettazione delle attività, inclusa l’attribuzione dei punteggi. A questo proposito, è stato stabilito di attribuire un punteggio pari ad 1 per ciascun *item* corretto, 0,5 per ciascun *item* corretto dal punto di vista del contenuto, ma contenente errori di ortografia (esempio: *andasi* vs *andassi*) e 0 per ciascun *item* del tutto scorretto (esempio: *andrei* vs *andassi*). Alla base di questa scelta c’è l’intento di dare preminenza alla competenza comunicativa e pragmatica rispetto alla pura correttezza formale: riprendendo l’esempio di cui sopra, se alla richiesta di usare un congiuntivo l’apprendente inserisce la risposta *andasi*, a tale risposta viene attribuito mezzo punto rispetto al totale, con l’intento di valorizzare comunque la capacità dell’apprendente di riconoscere la struttura grammaticale utile in quel contesto.

Motivazione



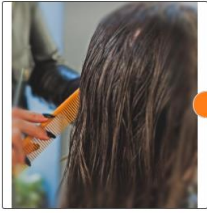



La prima pagina, intitolata *Mi preparo per uscire*, ha l’obiettivo di avvicinare gli apprendenti al tema dell’unità: l’immagine della donna con l’accappatoio e l’asciugacapelli anticipa il tema che verrà trattato poi nel brano da leggere. Vengono presentati gli obiettivi generali (ossia l’arricchimento del lessico, il ripasso dei pronomi riflessivi e un

² Cfr. i siti *Unsplash* (<https://unsplash.com/license>, visitato il 28/04/2019), *Burst* (<https://burst.shopify.com/legal/terms>, visitato il 12/03/2019), *Gratisography* (<https://gratisography.com/license>, visitato il 12/03/2019).

approfondimento dell'uso dei riflessivi in senso intensivo-affettivo nell'italiano neostandard) e viene proposta un'attività iniziale per "rompere il ghiaccio", incentrata sull'uso "non affettivo" dei verbi riflessivi, in cui si richiede di collegare alcune immagini che raffigurano le azioni della vita quotidiana legate alla cura del corpo ai verbi che le descrivono. Trattandosi semplicemente di un'attività introduttiva, svolta oralmente *in plenum* e a cui è stata attribuita la correzione automatica, i risultati non sono stati inseriti in questa trattazione.

Figura 6.1: Collegamento immagine-verbo

Che cosa fai prima di uscire di casa?
Collega le immagini ai verbi che le descrivono.

	<input type="checkbox"/> lavarsi
	<input type="checkbox"/> radersi/farsi la barba
	<input type="checkbox"/> vestirsi
	<input type="checkbox"/> truccarsi
	<input type="checkbox"/> pettinarsi
	<input type="checkbox"/> alzarsi

Globalità e analisi

La seconda pagina è intitolata *Che cosa mi metto?* e corrisponde alle fasi di globalità e analisi dell'unità didattica. Nella fase della globalità si propongono attività di comprensione testuale. Anche in questo caso l'immagine di un armadio con dei vestiti richiama direttamente il tema del brano che gli apprendenti andranno a leggere, che si intitola *L'appuntamento*³. È tratto da un blog in cui una donna descrive tutte le azioni compiute prima di presentarsi al primo appuntamento con un uomo. Si richiede agli apprendenti di leggere il testo per avere un'idea generale del tema, poi di svolgere un esercizio lessicale, a cui sono stati attribuiti 6 punti sul totale di 12 dell'intera pagina, in cui vengono forniti due elenchi di parole in colonna e viene chiesto di collegare alcune parole tratte dal brano, poste nella colonna di sinistra, con i relativi sinonimi nella colonna di destra. All'esercizio è stata attribuita la correzione automatica, impostando le soluzioni nella fase di progettazione dell'unità. Successivamente viene chiesto di rispondere a 3 domande aperte di comprensione del tema del brano, a ciascuna delle quali è stato attribuito un punto su 12.

L'ultima parte della pagina corrisponde alla fase di analisi, in cui si richiede di inserire in una tabella tutte le parole del brano che riguardano abbigliamento e accessori, prodotti per il corpo e i verbi che descrivono la cura del corpo, per un totale di 3 punti su 12: quest'ultima attività crea le basi per l'approfondimento sui verbi riflessivi nella pagina successiva, seguendo un percorso di presentazione induttiva dei contenuti grammaticali come illustrato nel capitolo 4.

³ Adattamento del brano tratto da: <https://notimetolose.myblog.it/?s=appuntamento&submit=Cerca> (consultato il 12/03/2019)

Figura 6.2: Estratto del brano da leggere



L'APPUNTAMENTO

Leggi il testo, per avere un'idea generale del tema. Non preoccuparti di capire tutto, cerca di cogliere il senso generale.



Sono già le otto e sono ancora sotto la doccia, mi devo asciugare i capelli: questi maledetti capelli, sono sempre spettinata! Accappatoio, mousse e tagliaunghie. Devo stare calma altrimenti sudo e devo di nuovo fare una doccia.

Otto e un quarto. Crema per il corpo. Mi lavo i denti? Sì, me li lavo, ma mi fumo una sigaretta. Speriamo che fumi anche lui. Ecco, lo smalto per le unghie, potevo mettermelo prima, adesso non c'è più tempo: lo farò la prossima volta, sempre che ci sia una prossima volta. (Sì, certo. Sì, certo. Ma non mi sciolino, non sciolino!) "Tempo, tempo!" "metto, è pronto!" momento dell' "Addio"...

Figura 6.3: Collegamento parola-sinonimo

Collega le parole nella colonna di sinistra con il loro significato

a. mousse	○	○	sfruttare al meglio
b. tagliaunghie	○	○	finire in una buca nel terreno
c. ottimizzare	○	○	strumento per tagliare le unghie
d. zircone	○	○	schiuma per capelli
e. trattore	○	○	cristallo che somiglia al diamante
f. infossarsi	○	○	mezzo di trasporto agricolo

Figura 6.4: Esempio delle domande di comprensione

Rispondi alle seguenti domande:

Quali emozioni prova la protagonista del testo prima dell'appuntamento?

Quali parole ti fanno capire questa cosa?

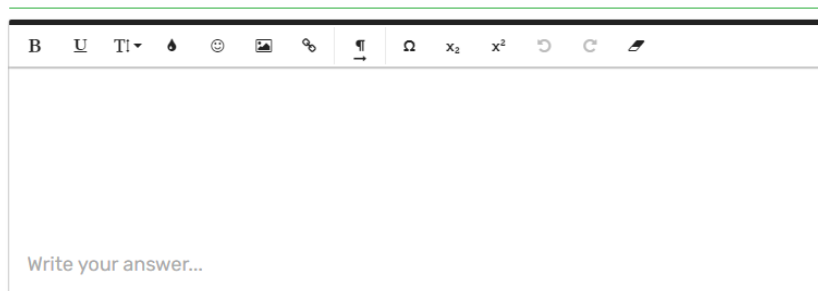


Figura 6.5: Completamento della tabella

Analisi

Trova nel testo tutte le parole che riguardano:

ABBIGLIAMENTO e
ACCESSORI

PRODOTTI PER IL CORPO

VERBI che descrivono la
cura del corpo

Riflessione

In questa fase viene chiesto agli apprendenti di riflettere sulle forme verbali riflessive sino ad ora incontrate. Viene innanzitutto proposto un esercizio di completamento della coniugazione di 3 verbi riflessivi, uno per ciascuna coniugazione del sistema verbale italiano (pari a 12 punti). Trattandosi di un'unità destinata al livello B1, non è stata proposta la distinzione tra le due forme della terza coniugazione (-ire; -isc-), essendo questo un tema generalmente affrontato nei precedenti livelli di competenza linguistica del QCER.

Segue un'attività di ripasso delle regole sulla posizione dei pronomi diretti e indiretti all'interno della frase (pari a 7 punti), in cui, tramite una lista di domande a scelta multipla, viene chiesto agli apprendenti di esplicitare perché il pronome si trovi in una certa posizione all'interno della frase.

Infine, viene chiesto agli apprendenti di leggere alcune frasi contenenti dei verbi riflessivi (cfr. Figura 6.8), e di collocarle in una tabella che racchiude quattro usi possibili della forma riflessiva, tra cui il verbo riflessivo "affettivo" (in questo caso *Mi fumo una sigaretta*), usato colloquialmente per sottolineare la partecipazione psicologica di chi compie l'azione (attività pari a 4 punti).

Figura 6.6: Completamento tabella delle coniugazioni

Di quali pronomi hanno bisogno questi verbi?

Completa la tabella con i pronomi e i verbi mancanti:

LAVARSI	METTERSI	VESTIRSI
io mi lavo		
	tu ti metti	
		lui/lei/Lei si veste
noi ci laviamo		
	voi vi mettete	
		loro si vestono

Figura 6.7: Ripasso delle regole sulla posizione dei pronomi diretti e indiretti

La posizione dei pronomi

1. Nelle seguenti frasi ci sono degli esempi delle regole grammaticali sulla posizione dei pronomi. Scegli a quale regola corrispondono le frasi.

Esempio: cosa mi metto? Il pronome si trova prima del verbo

È questa la pietra che vuoi? L'ho trovata nella foresta. Prendila, è tua !

a Prima del verbo
 b Unito al verbo all'imperativo
 c Unito al verbo all'infinito
 d Unito al verbo al gerundio

e Unito al verbo al participio.

Figura 6.8: Completamento della tabella sugli usi della forma riflessiva

Inserisci nella tabella gli esempi 1-6 che trovi qui di seguito.
NB: Ogni esempio appartiene ad una sola delle 4 categorie (a, b, c, d).

- Mi metto la gonna;
- Mi vesto;
- Ci incontriamo;
- Mi asciugo i capelli;
- Mi fumo una sigaretta;
- Mi trucco.

a) Azione reciproca tra più soggetti	b) azione su parti del corpo o vestiti/accessori che appartengono al soggetto	c) azione del soggetto su se stesso	d) il verbo non è un vero verbo riflessivo. Il pronome riflessivo viene usato colloquialmente solo per sottolineare la partecipazione psicologica di chi compie l'azione

Sintesi

Questa fase dell'unità include due pagine intitolate *Applicazione pratica 1* e *2*, dedicate al reimpiego attivo da parte degli apprendenti delle strutture grammaticali su cui si focalizza

l'unità. Nell'*Applicazione pratica 1* vengono proposti due esercizi, per un totale di 19 punti, nei quali gli apprendenti devono completare delle frasi scegliendo tra i verbi riflessivi "affettivi" dati nella consegna. Nell'*Applicazione pratica 2* sono presenti due esercizi: nel primo si richiede di completare delle frasi con le voci del verbo *fare* alla forma riflessiva (10 punti), mentre nel secondo si richiede di coniugare al passato prossimo il verbo riflessivo all'infinito dato tra parentesi (14 punti).

Figura 6.9: Esempio di attività dell'Applicazione pratica 1

Applicazione pratica



CIBO E BEVANDE
 Completa le frasi con i seguenti verbi: *faccio - fai - mangiamo - facciamo - beve - mangiava - berremo - fanno - mangio - berremo.*

La domenica metto la sveglia alle 13 , mi alzo con calma , accendo la tv , mi metto a tavola praticamente in mutande, e me la vedo mentre mi _____ qualcosa di gustoso

Stasera mi _____ il cous-cous con un sacco di verdure prelibate.

Ma lo trovi il tempo per mangiare? Ti _____ un panino davanti alla tv.

Figura 6.10: Applicazione pratica 2, primo esercizio

Applicazione pratica - 2



Quante cose possiamo fare? Completa le frasi con i **pronomi riflessivi** e le voci del verbo *fare* alla forma riflessiva:

Esempio:
 E se alla fine (farsi, tu) _____ i conti, è meglio che giocare in Borsa » **ti fai** i conti.

Ci sono persone che vivono in 15 dentro un appartamento! È necessario che tutti noi _____ un discorsetto su chi ha permesso e sta permettendo tutto questo.

Noi _____ una diretta in streaming

La mattina presto, quando vado a correre, _____ dei giri completi, quindi vedo tutto il parco.

Figura 6.11: Applicazione pratica 2, secondo esercizio

Coniuga al **passato prossimo** il verbo riflessivo all'infinito che trovi tra parentesi.
NB: l'ausiliare dei verbi riflessivi è sempre *essere*.

TEMPO LIBERO
Esempio: Da poco (farsi, lui) _____ una vacanza in una beauty farm
>> Da poco **si è fatto** una vacanza in una beauty farm.

Lucia: "Ieri sera (guardarsi, io) _____ la registrazione del webcast ".NET ed il test del software "".

Tra il calore delle quattro mura di Giuly, le tre amiche (farsi) _____ i pop-corn e Giuly si è rilassata.

Marco: "Ieri sera (vedersi, io) _____ la puntata di Ballarò registrata martedì."

Verifica

Sono state progettate due attività finalizzate a verificare il raggiungimento degli obiettivi generali dell'unità: conoscere la coniugazione riflessiva dei verbi, conoscere i diversi tipi di azioni che il verbo riflessivo può descrivere e, infine, saper scegliere in base al registro formale o informale della comunicazione tra l'uso della forma non riflessiva o della forma riflessiva "affettiva".

La pagina di *Applicazione pratica 3* contiene un esercizio in cui, data una frase contenente un verbo riflessivo già coniugato, gli apprendenti devono selezionare da una lista di risposte a scelta multipla la funzione che il verbo riflessivo esercita in quella frase: azione reciproca (esempio: *ci incontriamo*), azione compiuta su parti del corpo o vestiti/accessori (esempi: *mi lavo i denti*, *mi metto il cappello*), azione che il soggetto compie genericamente su se stesso (senza la specificazione di parti del corpo o vestiti/accessori, ad esempio: *mi lavo*), oppure un'azione descritta da un verbo riflessivo "affettivo" (esempio: *mi mangio un panino*), per un totale di 15 punti.

Infine, la pagina di *Applicazione pratica 4* contiene un esercizio in cui viene chiesto agli apprendenti di completare delle frasi coniugando i verbi dati all'infinito nella consegna e, una volta identificato il verbo giusto, di scegliere in base al contesto formale o informale in cui è inserita ciascuna frase, tra la forma non riflessiva o la forma riflessiva "affettiva". A tale attività è attribuito un totale di 10 punti.

Figura 6.12: Scelta multipla



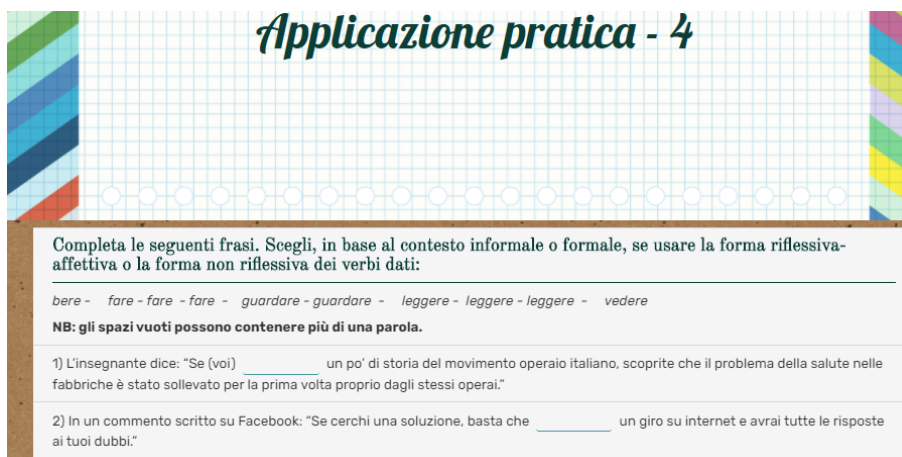
Applicazione pratica - 3

Qual è la funzione del verbo riflessivo? Scegli la risposta giusta:

1) Mi lavo i capelli due o tre volte a settimana.

- a azione reciproca tra più soggetti
- b azione compiuta su parti del corpo/ vestiti/accessori del soggetto
- c azione del soggetto su se stesso
- d azione descritta da un verbo riflessivo "affettivo"

Figura 6.13: Attività di completamento e scelta della forma riflessiva o non riflessiva



Applicazione pratica - 4

Completa le seguenti frasi. Scegli, in base al contesto informale o formale, se usare la forma riflessiva-affettiva o la forma non riflessiva dei verbi dati:

bere - fare - fare - fare - guardare - guardare - leggere - leggere - vedere

NB: gli spazi vuoti possono contenere più di una parola.

1) L'insegnante dice: "Se (voi) _____ un po' di storia del movimento operaio italiano, scoprite che il problema della salute nelle fabbriche è stato sollevato per la prima volta proprio dagli stessi operai."

2) In un commento scritto su Facebook: "Se cerchi una soluzione, basta che _____ un giro su internet e avrai tutte le risposte ai tuoi dubbi."

Rinforzo/recupero

Seguono delle attività facoltative dedicate all'uso dei verbi riflessivi nei modi di dire e alcune attività di produzione scritta e orale, i cui risultati esulano dal tema di questa trattazione.

Figura 6.14: Esempio di attività sui modi di dire

La seguente tabella contiene 4 significati metaforici del verbo *mangiarsi*. Scrivi i numeri delle frasi (1, 2, 3 o 4) in cui il verbo *mangiarsi* è usato con quel significato:

Significati metaforici	Numero della frase (1, 2, 3 o 4)
a. rovinare, distruggere.	
b. esaurire, sprecare;	
c. mangiare molto;	
d. non avere alternative: o così o niente;	

Figura 6.15: Esempio produzione scritta

Produzione scritta

Scrivi un breve racconto di una tua giornata tipica. Prova a riutilizzare il lessico che hai trovato nel testo e anche i verbi riflessivi di tipo "intensivo-affettivo" (es. mi mangio un panino).

B U Tl • 🔊 🌐 📄 🔄 🗑️ Ω x₂ x² ↻ 🖨️

6.2.2. Unità didattica sul congiuntivo

Motivazione

Per la fase di motivazione è stata mostrata una vignetta ironica sul tema dei dubbi sull'uso dell'indicativo vs congiuntivo e sull'istinto di correggere gli errori degli altri, che è stata utilizzata come attività orale di rottura del ghiaccio e condivisione di impressioni ed esperienze personali. In assenza di informazioni relative ai diritti di utilizzo e diffusione dell'opera, la vignetta non verrà riprodotta in questa trattazione e si rinvia al relativo link al profilo *Twitter* dell'autore⁴.

Globalità - analisi

Nelle fasi di globalità e analisi vengono proposte una serie di attività che permettono di familiarizzare con le regole d'uso del congiuntivo nella varietà standard. Le attività sono state progettate per poter essere adatte sia al livello B1, per un primo approccio con il tema del congiuntivo e della coniugazione del tempo presente e passato, sia come attività di

⁴ Astutillo Smeriglia: https://twitter.com/a_smeriglia/status/1081098811567673344 (consultato il 19/01/2019).

ripasso per i livelli superiori, scegliendo di svolgere tutte le attività proposte o creando percorsi personalizzati adattabili al livello della classe.

Nella prima pagina, come avvenuto per la fase di motivazione, come attività orale di rottura del ghiaccio e condivisione di impressioni ed esperienze personali è stata mostrata una vignetta ironica sulla tendenza comune a giudicare gli altri in base all'uso corretto o scorretto del congiuntivo. Anche in questo caso, in assenza di informazioni relative ai diritti di utilizzo e diffusione dell'opera, la vignetta non verrà riprodotta in questa trattazione e si rinvia al relativo link alla pagina che ospita l'immagine.⁵

Sempre nella prima pagina vengono proposti quattro brevi testi da leggere⁶ e viene chiesto di cercare al loro interno i verbi al congiuntivo e di classificare gli esempi trovati nei testi all'interno di una tabella che racchiude 4 ambiti d'uso: congiuntivo retto da verbi di opinione, congiuntivo retto da una frase principale con una forma impersonale, congiuntivo introdotto da una congiunzione subordinante e congiuntivo nelle interrogative indirette. All'attività corrisponde un totale di 4 punti.

Nella seconda pagina vengono presi in esame tutti i casi in cui nello standard il congiuntivo è obbligatorio, chiedendo gli apprendenti di completare delle frasi, coniugando al congiuntivo i verbi dati all'infinito per un totale di 10 punti.

Figura 6.16: Esempio di lettura e completamento della tabella

Testo n.4:

La scrittura deve servire innanzitutto a se stessi, per riflettere e trovare la coscienza di sé, come la meditazione nelle filosofie orientali. A proposito del fatto che la scrittura dev'essere per prima cosa un piacere in sé, vorrei ricordare cosa mi hanno detto Natalia Aspesi: "Penso che scrivere sia un grande piacere personale che ripaga di ogni delusione".

http://guide.supereva.it/case_editrici/interventi/2005/07/218030.shtml

Quando è usato il congiuntivo? Scrivi gli esempi che hai trovato nei testi:

Con un verbo di opinione	Con una forma impersonale	Con una congiunzione subordinante	Nella frase interrogativa indiretta

⁵ *I love Italian language:*
<https://iloveitalianlanguage.blog/2018/04/07/chi-ha-paura-del-congiuntivo> (consultato il 25/05/2019).

⁶ Adattamento brani tratti da:
https://it.wikipedia.org/wiki/Monaco_in_riva_al_mare;
https://it.wikipedia.org/wiki/Vita_su_Marte;
http://guide.supereva.it/viaggi_safari/interventi/2003/03/131025.shtml; http://guide.supereva.it/case_editrici/interventi/2005/07/218030.shtml (consultati il 20/03/2019).

Figura 6.17: Esempio attività di completamento

1. La domanda va spedita alla segreteria una settimana prima che (scadere) _____ i termini della consegna.
2. Tagliate le foglie e fatele asciugare al sole, finché non (diventare) _____ da verde scuro a gialle.

Riflessione

Nella pagina di riflessione sulle strutture grammaticali viene presentato un brano tratto da un blog con la recensione di un romanzo⁷. Nel testo vengono messe in evidenza due frasi, nelle quali il verbo principale è un verbo di opinione, seguito in un caso dal verbo coniugato al congiuntivo e nell'altro caso del verbo coniugato all'indicativo. In un'attività di completamento della regola grammaticale, a cui corrisponde un totale di 6 punti, viene chiesto agli apprendenti di fare ipotesi sulle ragioni alla base della scelta dell'indicativo e del congiuntivo. Al termine dell'attività viene fatto notare agli apprendenti che, trattandosi di un blog caratterizzato da un registro colloquiale, l'autore/autrice ha utilizzato un registro linguistico informale in cui è possibile utilizzare l'indicativo in presenza dei verbi di opinione.


Figura 6.18: Esempio di completamento della regola grammaticale

1. penso che sia geniale per risolvere problemi “seri” di una persona

a) Qual è il verbo della frase principale? _____

b) Quale modo verbale è presente nella frase secondaria? _____

c) La frase ti sembra corretta?



Sintesi

Sono state create 3 pagine di esercizi, ciascuna dedicata ai 3 casi in cui la grammatica italiana prevede la scelta tra congiuntivo e indicativo a seconda del grado di formalità o informalità del contesto (cfr. capitolo 2): l'uso dell'indicativo al posto del congiuntivo

⁷ Adattamento brano tratto da: www.cineblog.it/post/20702/tutta-da-rifare-il-romanzo-di-giorgia-wurth-il-book-trailer/1 (consultato il 20/03/2019).

nelle frasi interrogative indirette, l'uso dell'indicativo al posto del congiuntivo nelle frasi subordinate rette da verbi alla forma impersonale e l'uso dell'indicativo al posto del congiuntivo in frasi subordinate rette da verbi di opinione. In accordo con i docenti referenti delle diverse sedi, in ciascuna classe è stata adottata come attività da svolgere in classe la pagina dedicata alle interrogative indirette, lasciando le altre due pagine come attività facoltative da svolgere eventualmente a casa.

Nello specifico, la pagina sul congiuntivo *vs* indicativo nell'interrogativa indiretta è composta da due esercizi, per un totale di 14 punti. Dopo una breve introduzione, in cui si ricorda che nelle frasi interrogative indirette il congiuntivo corrisponde a uno stile più formale e ad un maggior grado di incertezza, viene proposta una prima attività: agli apprendenti viene chiesto di leggere alcune frasi in cui nella subordinata interrogativa indiretta sono presenti verbi all'indicativo e successivamente viene chiesto di trasformarle per adattare ad una situazione formale, ossia di coniugare i verbi della subordinata al congiuntivo e di apportare qualunque altra modifica ritenuta necessaria (uso del *Lei* e di espressioni di cortesia, eccetera). Nell'attività successiva viene chiesto, invece, di adattare altre frasi ad un contesto informale.

Figura 6.19: Esempio di attività sull'uso del congiuntivo *vs* indicativo nell'interrogativa indiretta

Leggi le seguenti frasi:

- 1) Cominciamo ad andare in giro per la città. Già nei primi momenti mi chiede se io sono la giornalista straniera che è venuta a visitare la città.
- 2) Se si intende lasciare l'auto in una città o in un posto diverso da quello dove si noleggia, bisogna dirlo subito, e chiedere se sono previsti sovrapprezzi in questo caso.
- 3) La ragazza chiede se c'è qualcuno che proviene da fuori e Clem solleva la sua mano.
- 4) Mi chiede se per caso io ho letto "I Libri di Natale" di Dickens.

Ora trasforma le frasi precedenti per adattare ad una situazione formale: stai scrivendo un articolo sulla tua esperienza come studente/ssa in scambio all'estero che sarà pubblicato sul sito della tua università.

In entrambe le attività, in contrasto con l'approccio induttivo su cui si basa il percorso delle unità prodotte, si chiede esplicitamente di coniugare, a seconda dei casi, i verbi al congiuntivo o all'indicativo. Questa scelta di un approccio chiaramente deduttivo ("coniuga al...") a prima vista potrebbe destare delle perplessità. Come evidenziato da

Balboni (2002: 139) le tecniche di manipolazione tipiche dell'approccio formalistico⁸, in cui ci si serve di consegne del tipo "Trasforma al...", mirano ad applicare meccanicamente le regole, piuttosto che ad introdurre una riflessione. Tuttavia, lo stesso Balboni suggerisce che sia possibile inserire tali attività in un approccio di tipo comunicativo e induttivo attribuendo una giustificazione comunicativa alle tecniche di manipolazione: ad esempio, chiedere nella consegna di un esercizio raccontare un film giustifica la trasformazione dei verbi dal presente al passato. Analogamente, nel caso specifico dell'unità didattica sul congiuntivo, lo stesso risultato può essere ottenuto modificando le frasi per adattarle ad una diversa situazione comunicativa che preveda un diverso livello di formalità (esempio: racconta le tue esperienze come studente *Erasmus* ai tuoi amici *vs* ora modifica quanto hai scritto per riferire le stesse esperienze e opinioni personali in un articolo del portale web dell'università).

Questa svolta verso l'approccio deduttivo in queste pagine non interrompe la progressione crescente di difficoltà dell'unità, poiché nell'attività finale, come verrà illustrato nella sezione successiva, spetta agli apprendenti stabilire se utilizzare il congiuntivo o l'indicativo rispetto al grado di formalità del contesto.

Verifica


Infine, sono state progettate due attività finalizzate a verificare il raggiungimento degli obiettivi generali dell'unità: conoscere la coniugazione del congiuntivo presente e passato, conoscere i casi in cui nella varietà standard è prescritto l'uso del congiuntivo e i casi in cui nella varietà neostandard l'uso dell'indicativo al posto del congiuntivo corrisponde ad un grado inferiore di formalità, e, infine, saper scegliere in base al registro formale o informale della comunicazione tra l'uso del congiuntivo o dell'indicativo. Alle attività finali è stato attribuito un totale di 20 punti.

Vengono proposte due attività: nella prima vengono fornite delle frasi complete con verbo coniugato al congiuntivo e per ciascuna frase viene chiesto agli apprendenti di selezionare da una lista a scelta multipla la risposta che spiega perché in quel caso specifico è obbligatorio il congiuntivo (ossia, se è presente un verbo di

⁸ Approccio glottodidattico che ha caratterizzato l'insegnamento delle lingue straniere dal XVIII secolo agli anni Settanta del XX secolo, basato su un percorso di tipo deduttivo, in cui il docente, fonte di informazioni e modello da seguire, fornisce delle regole da cui gli studenti deducono i comportamenti linguistici. Le tecniche didattiche utilizzate sono, oltre ai già citati esercizi di manipolazione ("Trasforma al..."), il dettato e la traduzione tra la lingua materna e la lingua straniera (metodo grammaticale-traduttivo).

opinione/sentimento o una forma impersonale nella frase principale, oppure se è presente una congiunzione subordinante o una frase interrogativa indiretta). Segue una seconda attività di completamento: agli apprendenti è richiesto di completare delle frasi coniugando il verbo della frase secondaria all'indicativo o al congiuntivo a seconda del contesto formale informale in cui è inserita la frase.

Figura 6.20: Esempio della prima attività



Scegli la risposta che spiega perché, secondo lo standard, è obbligatorio il congiuntivo nelle seguenti frasi:

1) Quando facevo in moto quel percorso, mi sembrava una cosa lunga. Adesso mi sembra che non siano passati due minuti.

- a** C'è un verbo di opinione/sentimento nella frase principale
- b** C'è una forma impersonale nella frase principale
- c** C'è una congiunzione subordinante
- d** È una frase interrogativa indiretta

Figura 6.21: Esempio della seconda attività

Indicativo o congiuntivo? Scegli in base al contesto informale o formale della frase:

1) L'insegnante dice: "Mi chiedo se per caso tu (avere) _____ letto "I Libri di Natale" di Dickens."

2) Un giornalista scrive: "Liz Taylor, rimasta evidentemente affascinata da Depardieu, gli chiede se (essere) _____ sposato. L'attore risponde di sì e la Taylor commenta "Peccato!"".

3) Hai fatto un incidente stradale e hai danneggiato l'auto di un'altra persona: "Basta, mi sono stancato! Se vuoi che io ti (pagare) _____ il danno, prendi il modulo per la denuncia!"

6.3. Esiti della sperimentazione

In questa sezione verrà svolta un'analisi quantitativa e qualitativa dei dati raccolti tramite le sperimentazioni delle unità didattiche. Per ciascuna fase dell'unità verranno riportati i dati degli esiti delle attività svolte, suddivisi in base alle diverse sedi di svolgimento della

sperimentazione. Successivamente l'attenzione verrà rivolta ai dati delle attività finali di verifica di ciascuna unità.

Al fine di coinvolgere studenti con un livello di competenza pari al B1, sono stati selezionati gli studenti del I e II anno del corso di laurea triennale dell'Università di Breslavia, del II anno del corso di laurea quadriennale dell'Università Complutense di Madrid e del III e IV anno del corso di laurea quadriennale dell'Università di Murcia. Il profilo della popolazione coinvolta sarà tracciato nella sezione 6.4. Segue un elenco delle classi coinvolte in ciascuna sede:

Unità sui verbi riflessivi

- 07/06/2019: Università di Breslavia, II anno, 5 persone, successivamente ridotte a 4;
- 14/06/2019: Università di Breslavia, I anno, 15 persone;
- 24/09/2019: Università Complutense di Madrid, 22 persone;
- 08/10/2019: Università di Murcia, 6 persone;
- 16/01/2020: Università di Malaga: 5 persone.

Unità sul congiuntivo

- 14/06/2019 Università di Breslavia II anno, 4 persone;
- 09/10/2019 Università di Murcia: 3 persone del IV anno e 2 persone del III anno;
- 13/01/2020 Università di Malaga: II anno: 7 persone.

Prima di procedere all'analisi dei dati raccolti nell'arco di queste sperimentazioni, è necessario premettere quanto segue. Avendo a disposizione un collettivo estremamente ridotto, è bene precisare che lo scopo della sperimentazione e del successivo monitoraggio non è di ricavare conclusioni di carattere generale, valide su vasta scala, per le quali sarebbe stato necessario un collettivo più ampio, bensì di svolgere un primo test dei materiali didattici e valutarne l'efficacia in termini di utilità e contributo motivazionale nell'acquisizione linguistica.

Inoltre, quando si progetta una sperimentazione, è necessario che l'ambiente sperimentale venga tenuto in condizioni costanti, ad eccezione della variabile sperimentale, che invece viene volutamente cambiata. Nel caso della presente ricerca, sono state svolte delle lezioni di lingua italiana di livello B1, aventi tra gli obiettivi linguistici un ripasso dell'uso standard di due tratti grammaticali generalmente affrontati nel livello intermedio, all'interno del quale si inserisce, come elemento di novità, un approfondimento dell'uso

neostandard di tali tratti grammaticali. L'obiettivo della sperimentazione è valutare se e in che misura i partecipanti al termine delle attività siano in grado di formulare delle scelte tra uso standard e neostandard rispetto al grado di formalità/informalità del contesto e di riconoscere l'uso neostandard da parte di parlanti nativi italofoni come il frutto di una scelta e non di una scarsa padronanza delle regole grammaticali.

Una sperimentazione controllata prevede che i partecipanti vengano divisi in due gruppi: un gruppo sperimentale sottoposto ai cambiamenti della variabile e un gruppo di controllo che non sia esposto a quei cambiamenti, la cui presenza garantisce che i dati prodotti dal gruppo sperimentale siano realmente dovuti alla variabile che si sta verificando e non a influenze esterne sconosciute. Nel caso in questione, la prassi per ottenere dei dati sull'efficacia delle attività svolte consisterebbe nel dividere la popolazione in due gruppi: un gruppo che svolge tutte le attività della sperimentazione, incluso l'approfondimento sulla scelta tra uso standard e neostandard rispetto al grado di formalità/informalità del contesto, e un gruppo di controllo a cui non venga fornito tale approfondimento.

Tuttavia, negli accordi siglati con gli atenei esteri coinvolti nella sperimentazione è stato concesso un numero di ore strettamente necessario allo svolgimento delle due unità didattiche, ossia 1 ora e 30 per ciascuna unità, fino ad un massimo di 3 ore totali, che non consentiva di dividere le classi ed effettuare incontri differenziati con il gruppo di sperimentazione e il gruppo di controllo. Di conseguenza, è stato necessario fornire a tutti i partecipanti gli stessi contenuti. Se ne deduce che i dati raccolti non forniscono indicazioni sull'efficacia delle attività di verifica finale in assenza dello svolgimento dell'intera unità didattica.

Inoltre, il totale dei partecipanti alla sperimentazione in taluni casi varia in base alle diverse attività svolte. I fattori che hanno determinato la mancata consegna di alcune attività sono principalmente di natura logistica (ad esempio, l'arrivo in ritardo o, viceversa, l'uscita anticipata dall'aula di alcune persone rispetto all'orario della lezione), mentre solo in pochi casi isolati sono state riscontrate delle difficoltà nel salvataggio delle risposte sulla piattaforma, causati dall'instabilità della connessione internet o da incomprensioni circa la modalità di salvataggio.

Infine, è necessario precisare che prima di ciascuna sperimentazione si è svolto un confronto con i docenti referenti presso i dipartimenti ospiti per valutare se i materiali progettati fossero adeguati al gruppo-classe in cui sarebbero stati sperimentati ed, eventualmente, selezionare alcune attività a cui dare maggiore rilevanza o, invece, da escludere, in base alle competenze linguistiche degli apprendenti. Pertanto, al fine di

adattare le unità al livello dei diversi gruppi-classe coinvolti nella sperimentazione, lo svolgimento nelle diverse sedi delle schede di *Comprensione testuale*, *Riflessione* e *Applicazione pratica* dell'unità sui verbi riflessivi "affettivi" ha subito delle variazioni in base al loro livello di maggiore o minore complessità.

In generale, per tutte le schede che compongono le unità è stata svolta un'analisi comparativa dei dati tramite il calcolo del punteggio medio totalizzato (incluso il calcolo della deviazione standard) e della percentuale delle risposte corrette. Nei casi in cui la sperimentazione è avvenuta in maniera univoca (ossia utilizzando sempre le stesse attività contenute nelle schede) in almeno due delle sedi coinvolte, la trattazione è stata supportata da grafici. Nel caso, invece, delle schede svolte solo in una delle sedi coinvolte (esempio, le schede di *Applicazione pratica 2* e *4* dell'unità dedicata ai verbi riflessivi "affettivi"), i dati sono stati solo brevemente commentati nel corso della trattazione.

Un caso particolare è rappresentato dalla fase di *Riflessione* sopra menzionata, il cui svolgimento è avvenuto in maniera differenziata nelle diverse sedi di sperimentazione e, dunque, i dati delle diverse sedi non sono perfettamente paragonabili. Si è tenuto conto di tale aspetto nella scelta della modalità di analisi dei dati e di cui è stata fornita una descrizione separata per ciascuna sede. Ciò verrà spiegato nel dettaglio nelle sezioni successive.

6.3.1. Unità didattica sui verbi riflessivi

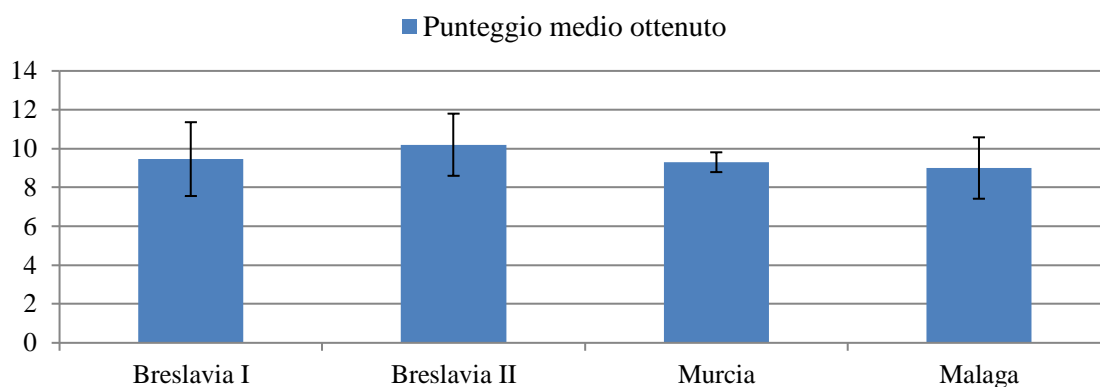
Comprensione testuale

Per facilitare la "rottura del ghiaccio", le attività della fase di comprensione testuale sono state svolte principalmente oralmente. I partecipanti hanno riportato le proprie risposte sul foglio di lavoro elettronico, ma la correzione è avvenuta *in plenum*, ossia coinvolgendo l'intero gruppo-classe nella ricerca delle soluzioni. Pertanto, per questa fase dell'unità si è deciso di tenere traccia solo del punteggio medio totalizzato all'interno di ciascun gruppo classe su un punteggio massimo pari a 12.

Tabella 6.2: Esiti delle attività di comprensione testuale

	Partecipanti	Punteggio medio totalizzato (max: 12)	Deviazione standard	Punteggio min. e max. totalizzato in 12esimi	Percentuale risposte corrette
Breslavia I anno	15	9,46	1,9	5,5 - 12	78,8%
Breslavia II anno	5	10,2	1,6	8 - 12	85%
Madrid	22	Attività svolta oralmente <i>in plenum</i>			
Murcia	6	9,3	0,51	9 - 10	77,7%
Malaga	5	9	1,58	7 - 11	75%

Grafico 1: Esiti attività di comprensione testuale



Come emerge dalla tabella, gli obiettivi di questa fase dell'unità sono stati pienamente raggiunti dai vari gruppi di sperimentazione. I punteggi totalizzati dai partecipanti sono quasi tutti al di sopra della sufficienza, ad esclusione di un unico punteggio pari a 5,5 nel caso di **Breslavia** (I anno).

Osservando i risultati ottenuti nelle sedi di **Breslavia** (I e II anno), **Murcia** e **Malaga**, il punteggio medio ottenuto varia da un minimo di 9 (deviazione standard:1,58) di **Malaga** ad un massimo di 10,2 (deviazione standard:1,6) di Breslavia (II anno). In particolare, il caso di **Murcia** presenta la maggiore omogeneità dei risultati, come si evidenzia dalla deviazione standard inferiore a quella degli altri gruppi, pari a 0,5.

Per quanto riguarda, invece, la sperimentazione svolta a **Madrid**, per questa fase dell'unità non sono stati raccolti dei dati. A causa dell'impossibilità di utilizzare un'aula con postazioni computer, sono stati utilizzati i dispositivi elettronici dei partecipanti (computer portatili, *tablet* e *smartphone*), il che ha comportato un allungamento dei tempi di accesso alla rete e alla piattaforma *Wizer.me*. Pertanto, si è scelto di far svolgere la

comprensione testuale solo oralmente e *in plenum*, proiettando le attività alla parete tramite il computer della cattedra. Ciò ha consentito lo svolgimento sincronico online di tutte le fasi successive dell'unità e il contenimento dei tempi della sperimentazione entro il blocco della lezione.

In generale gli errori rilevati nella produzione scritta delle diverse classi di sperimentazione riguardano contenuti grammaticali non inerenti al tema della lezione. Si citano a titolo di esempio: accenti, articoli (*dal* appuntamento > dall'; l'articolo indeterminativo maschile apostrofato), morfologia (*bucca* >> buca; le frase; *duccia* >> doccia; *andro* >> andrà), accuratezza lessicale (andare *vs* venire; è *rida* >> è divertita), preposizioni rette da verbi (*prova di* >> prova a), uso dei tempi verbali (condizionale semplice al posto del condizionale passato per esprimere il futuro nel passato), collocazioni (essere *in fretta* >> essere di fretta; arrivare *nel* tempo >> *in* tempo).

Riflessione

Questa fase dell'unità rappresenta uno snodo importante del percorso, in cui gli apprendenti dimostrano di saper riflettere induttivamente sulle regole grammaticali che sono alla base dei contenuti linguistici presentati.

Allo scopo di adattare l'unità al livello di competenza dei gruppi-classe coinvolti, come già anticipato, per questa fase si è deciso di differenziare le attività da svolgere nelle sedi di sperimentazione. Di conseguenza, rispetto al punteggio massimo originario di 23 punti, corrispondente al totale delle attività proposte nella fase di riflessione, per ciascuna sede di sperimentazione il punteggio massimo è stato calcolato sulla base delle sole attività svolte. Non essendo, pertanto, possibile confrontare i dati tratti da attività differenti, l'analisi sarà limitata ad un breve commento su quanto è stato svolto in ciascuna sede.

Presso l'Università di **Breslavia** non è stato assegnato lo svolgimento dell'attività di completamento della tabella di coniugazione, pari a 12 punti, pertanto il punteggio è stato calcolato sugli 11 punti rimanenti, corrispondenti all'attività sulla posizione dei pronomi nella sintassi della frase e all'attività di completamento della tabella con le funzioni svolte dai verbi riflessivi. In entrambi i casi gli esiti delle attività mostrano il generale pieno raggiungimento degli obiettivi. Nella classe del I anno e il punteggio medio totalizzato dai 15 partecipanti è pari a 8,3 (deviazione standard:1,3) e varia da un minimo di 7/11 ad un massimo di 10/ 11. La percentuale delle risposte corrette è pari al 75,7%. Nella classe del II anno, che a partire da questa fase dell'unità ha avuto una riduzione del numero dei

partecipanti da 5 a 4, gli esiti sono stati superiori, sia per quanto riguarda il punteggio medio totalizzato, pari a 9,75 (deviazione standard:1,5), compreso tra un minimo di 9/11 e un massimo di 11/ 11, sia per quanto riguarda la percentuale delle risposte corrette, pari all'88,6%.

Presso l'Università Complutense di **Madrid**, poiché il livello linguistico dei partecipanti era pari ad un A2 pienamente raggiunto/ B1.1 in entrata (la sperimentazione si è svolta ad inizio semestre), è stata assegnata la compilazione sia della tabella di coniugazione, sia della tabella conclusiva sulle diverse funzioni dei verbi riflessivi, mentre è stata esclusa l'attività sulla posizione dei pronomi nella frase, essendo un tema non ancora del tutto affrontato nel programma del corso. La fase di riflessione si è basata, pertanto, su un totale di 16 punti. Le attività hanno coinvolto 21 persone e il punteggio medio totalizzato è pari a 12,4 (deviazione standard: 3,5), compreso tra un minimo di 3/16 e un massimo di 16/16. L'elevata variabilità presente tra i punteggi totalizzati si giustifica sulla base dei dati sulla popolazione, in particolare i dati riguardanti la durata dello studio pregresso della lingua italiana, a cui si farà riferimento in seguito nelle sezioni 6.3.3 e 6.4. La percentuale delle risposte corrette è pari al 77,6%.

Per quanto riguarda le università di Murcia e Malaga, lo svolgimento della fase di riflessione ha seguito un percorso diverso dalle sperimentazioni precedentemente citate. La classe dell'Università di **Murcia** è l'unico caso in cui la fase di riflessione è stata svolta interamente, su richiesta della docente referente, come attività di ripasso in vista degli esami. Pertanto, i calcoli sono stati effettuati sul totale originario di 23 punti. Le attività hanno coinvolto 5 persone e gli esiti delle attività mostrano il generale pieno raggiungimento degli obiettivi: il punteggio medio totalizzato è pari a 20 (deviazione standard: 0,7), compreso tra un minimo di 19/23 e un massimo di 21/23. La percentuale delle risposte corrette è pari all'86,9%.

Invece presso l'Università di **Malaga**, è stata svolta su richiesta la sola compilazione della tabella conclusiva sulle diverse funzioni dei riflessivi, per un totale di 4 punti. L'attività ha coinvolto 5 persone, che hanno tutte totalizzato il punteggio massimo pari a 4/4.

Applicazione pratica 1

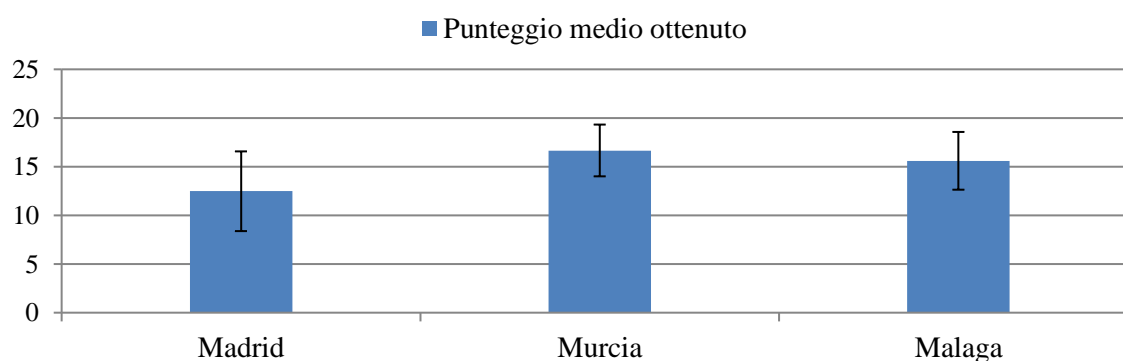
In accordo con il docente referente dello scambio presso l'Università di **Breslavia**, non è stata svolta la sperimentazione di questa parte dell'UD, poiché le attività che la

compongono (cfr. sez. 6.2.1) sono state valutate di livello nettamente inferiore alle competenze linguistiche di entrambi i gruppi classe. Al suo posto è stata svolta con il gruppo del II anno la sperimentazione dell'*Applicazione pratica 2*, a cui si farà riferimento alla sezione successiva. Di seguito verranno riportati nella tabella 6.3 i dati raccolti nelle sedi di Madrid, Murcia e Malaga.

Tabella 6.3: Esiti dell'attività Applicazione pratica 1- verbi riflessivi "affettivi"

	Partecipanti	Punteggio medio totalizzato	Deviazione standard	Punteggio min. e max. totalizzato in 19esimi	Percentuale risposte corrette
Madrid	21	12,4	4	7 - 19	65,6%
Murcia	6	16,6	2,6	12 - 19	87,7 %
Malaga	5	15,6	2,9	11- 19	82,1 %

Grafico 2: Applicazione pratica 1- verbi riflessivi "affettivi"



Osservando il grafico 2 e i dati riportati nella tabella 6.3 emerge che nella classe di **Murcia** sono stati totalizzati i valori più alti relativi sia alla percentuale delle risposte corrette pari all'87,7% (cfr. Madrid: 65,6%; Malaga: 82,1%), sia al punteggio medio, pari a 16,6 (con deviazione standard pari a 2,6), che, confrontato con i valori medi totalizzati nelle altre sedi, è il valore medio più omogeneo. Nella classe di **Madrid**, infatti, il punteggio medio totalizzato è pari a 12,4 con deviazione standard pari a 4, dato riconducibile alla variabilità dei punteggi totalizzati, compresi tra un minimo di 7/19 e il massimo di 19/19. Nella classe di **Malaga**, invece, la variabilità è inferiore: il punteggio medio è pari a 15,6 con deviazione standard pari a 2,9 e i punteggi totalizzati sono compresi tra un minimo di 11/19 e un massimo di 19/19.

Applicazione pratica 2

Come già accennato in precedenza, la sperimentazione della scheda *Applicazione pratica 2* è stata svolta solo presso l'Università di **Breslavia**, in sostituzione della scheda *Applicazione pratica 1*. La scheda è incentrata sulla coniugazione della forma riflessiva del verbo irregolare *fare* e sull'uso del passato prossimo nella forma riflessiva. La sperimentazione è stata svolta dalle 4 persone del II anno, che hanno totalizzato punteggi compresi tra un minimo di 17/24 e un massimo di 22/24 punti e un punteggio medio pari a 20 (deviazione standard: 2,1). La percentuale delle risposte corrette è pari all'83,3%.

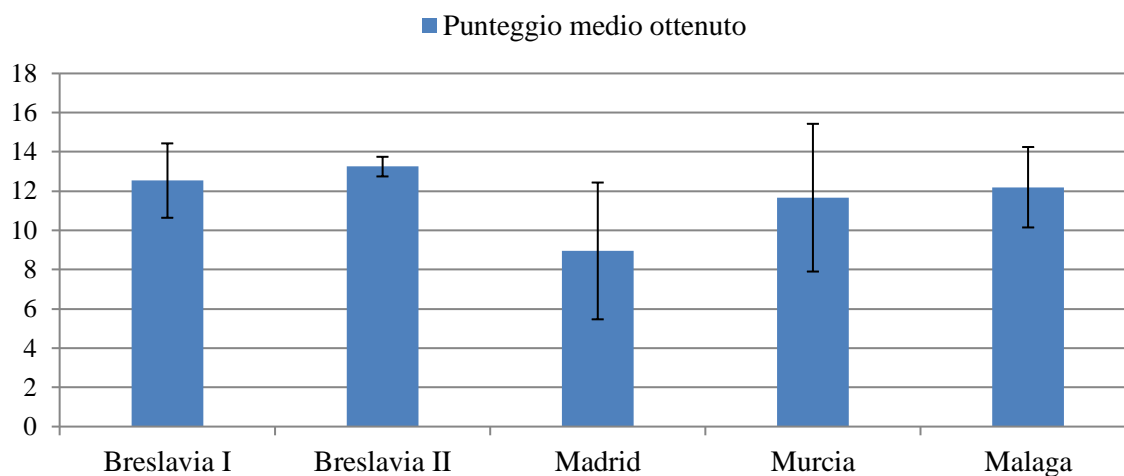
Applicazione pratica 3

Questa fase corrisponde alla verifica finale degli obiettivi dell'unità didattica. La scheda prevede un punteggio massimo di 15 punti. Come si evince osservando i dati riportati nella tabella 6.4 e nella loro rappresentazione grafica (grafico 3), rispetto al totale di 15 punti dell'attività, i punteggi medi totalizzati in ciascuna classe di sperimentazione sono tutti al di sopra della sufficienza e quasi tutti molto vicini al punteggio massimo. Di particolare interesse il caso della classe del II anno di **Breslavia**, dove il punteggio medio totalizzato è pari a 13,2 e la deviazione standard è pari a 0,5, dunque la media ottenuta rispecchia una notevole omogeneità dei risultati all'interno del campione. Ciò è facilmente confermabile osservando i dati del campione, ove i punteggi totalizzati dai partecipanti variano tra i 13 e i 14 punti su 15. In tutti gli altri casi, in cui i valori della deviazione standard superiori allo zero sono riconducibili alla variabilità dei singoli punteggi totalizzati all'interno del campione, di cui si riportano nella tabella solo i valori minimi e massimi registrati.

Tabella 6.4: Esiti dell'attività di verifica finale sull'uso dei riflessivi "affettivi"

	Partecipanti	Punteggio medio totalizzato	Deviazione standard	Punteggio min. e max. totalizzato in 15esimi	Percentuale risposte corrette
Breslavia I anno	13	12,5	2,1	9 - 15	83,5%
Breslavia II anno	4	13,2	0,5	13 - 14	88,3%
Madrid	21	8,9	3,4	4 - 14	59,6%
Murcia	6	11,1	3,7	6 - 15	74,4%
Malaga	5	12,2	2	10 - 14	81,3%

Grafico 3: Esiti dell'attività di verifica finale sull'uso dei riflessivi "affettivi"



Applicazione pratica 4

Questa fase contiene delle attività di rinforzo proposte al termine dell'unità come approfondimento. In accordo con i vari docenti referenti, è stata svolta solo presso l'Università di **Breslavia** (I e II anno), dove gli esiti prodotti in entrambe le classi sono tutti al di sopra della sufficienza rispetto al totale di 10 punti e sono pressoché analoghi: il punteggio medio registrato nella classe del I anno dai 13 partecipanti è pari a 6,7 (deviazione standard pari a 1,9) e varia da un minimo di 3/10 a un massimo di 9/10. Analogamente nella classe del II anno il punteggio medio registrato dai 4 partecipanti è pari a 6,3 (deviazione standard pari a 2,8) e varia da un minimo di 3/10 e un massimo di 8/10.

6.3.2. Congiuntivo

In accordo con i docenti referenti la sperimentazione dell'unità sulla scelta tra indicativo e congiuntivo non è stata svolta a Madrid, né nella classe del I anno a Breslavia, dato che il tema del congiuntivo non era ancora stato trattato in entrambi i corsi di lingua italiana e dunque si è ritenuto che l'unità non fosse adeguata al livello di competenza degli apprendenti. Di seguito verranno presentati gli esiti ottenuti nelle diverse fasi delle unità.

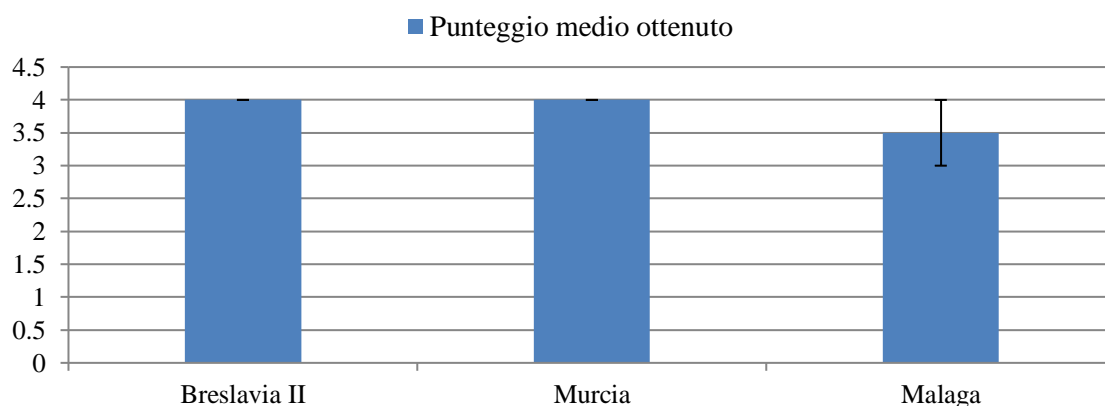
Congiuntivo 1

Gli obiettivi di identificazione dei quattro ambiti d'uso del congiuntivo della fase introduttiva dell'unità sono stati pienamente raggiunti da tutti i partecipanti, con un punteggio medio pari al massimo di 4/4 nelle classi di **Breslavia** e **Murcia**, o di mezzo punto inferiore nella classe di **Malaga**, dove su un totale di 7 apprendenti, 4 persone hanno raggiunto il punteggio massimo di 4/4 e 3 persone hanno raggiunto il punteggio di 3/4. La tabella 6.5 mostra l'omogeneità dei campioni delle tre classi coinvolte.

Tabella 6.5: Esiti delle attività della fase Congiuntivo I

	Partecipanti	Punteggio medio totalizzato	Deviazione standard	Punteggio min. e max. totalizzato in quarti	Percentuale risposte corrette
Breslavia II anno	4	4	0	4	100%
Murcia	6	4	0	4	100%
Malaga	7	3,5	0,5	3-4	89,2%

Grafico 4: Esiti della fase Congiuntivo I



Congiuntivo 2

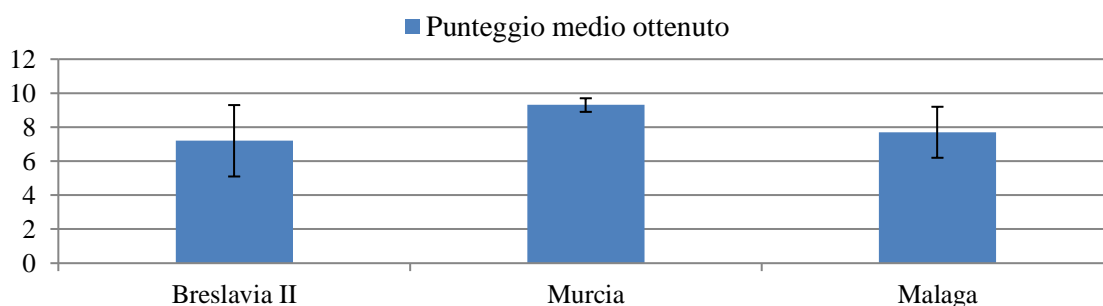
Per quanto riguarda la seconda fase dell'unità, dedicata all'approfondimento dell'uso standard del congiuntivo, osservando i dati riportati nella tabella 6.6, anche in questo caso i punteggi medi totalizzati in ciascuna classe di sperimentazione sono tutti al di sopra della sufficienza rispetto al totale di 10 punti dell'attività. Il punteggio medio più alto è stato totalizzato nella classe di **Murcia**, pari a 9,3 che, confrontato con i punteggi medi totalizzati nelle altre sedi, risulta anche essere la media caratterizzata dalla maggiore omogeneità all'interno del campione, con deviazione standard pari a 0,4. Nella classe di

Breslavia, infatti, il punteggio medio totalizzato è pari a 7,2 con deviazione standard pari a 2,1, mentre nella classe di **Malaga** il punteggio medio è pari a 7,7 e il valore della deviazione standard si attesta a 1,5.

Tabella 6.6: Esiti delle attività della fase Congiuntivo II

	Partecipanti	Punteggio medio totalizzato	Deviazione standard	Punteggio min. e max. totalizzato in 10mi	Percentuale risposte corrette
Breslavia II anno	4	7,2	2,1	4,6-9,2	72,4%
Murcia	5	9,3	0,4	8,6-9,8	93,3%
Malaga	7	7,7	1,5	5,5-9,4	77,3%

Grafico 5: Esiti della fase Congiuntivo II



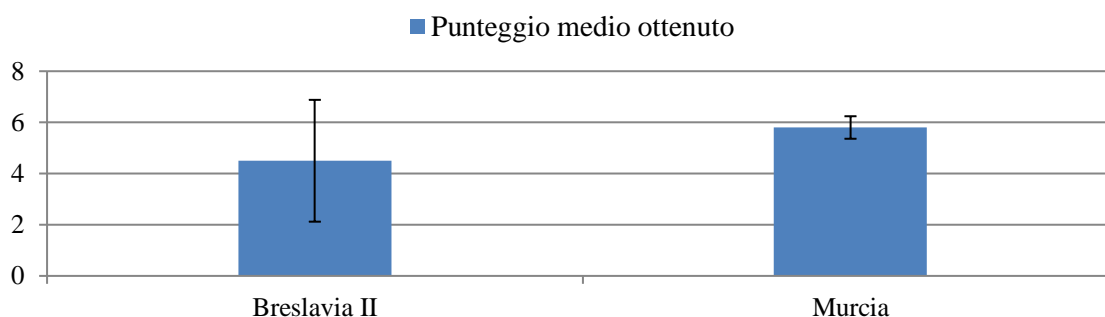
Riflessione

Circa gli esiti della fase di riflessione, mirata alla comprensione induttiva da parte dei partecipanti, dell'uso dell'indicativo al posto del congiuntivo in un contesto informale, ci si limita a rilevare che gli obiettivi sono stati pienamente realizzati in tutte le sedi. Nella classe di **Breslavia** su 4 partecipanti, 2 persone hanno totalizzato il punteggio massimo di 6/6, una persona ha totalizzato 5 punti e una sola persona 1 punto. Molto simile l'andamento nella classe di **Murcia**, dove su 6 partecipanti, 5 hanno totalizzato il massimo di 6/6 e una persona ha totalizzato 5 punti su 6. Nel caso di **Malaga**, per ragioni di tempo l'attività di riflessione è stata svolta in forma collettiva e corretta oralmente; emerge dagli appunti della sperimentazione la nota sull'esito globalmente positivo dell'attività.

Tabella 6.7: Esiti della fase di riflessione

	Partecipanti	Punteggio medio totalizzato	Deviazione standard	Punteggio min. e max. totalizzato su 6	Percentuale risposte corrette
Breslavia II anno	4	4,5	2,3	1-6	75%
Murcia	5	5,8	0,4	5-6	96,6%
Malaga	Attività svolta oralmente <i>in plenum</i>				

Grafico 6: Esiti della fase di riflessione



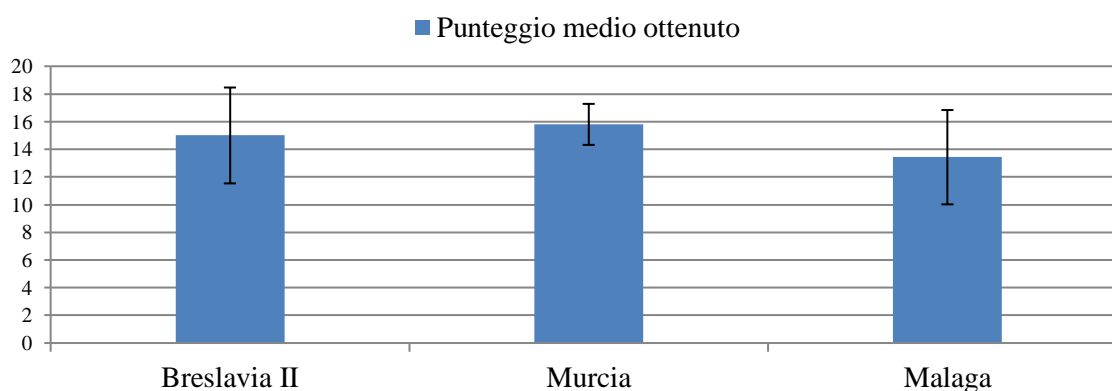
Uso del congiuntivo: formalità vs informalità

Osservando i dati riportati nella tabella 6.8 e nel grafico 7, anche in questo caso i punteggi medi totalizzati in ciascuna classe di sperimentazione sono tutti al di sopra della sufficienza rispetto al totale di 20 punti dell'attività. I valori della deviazione standard, tutti superiori allo zero, sono riconducibili alla variabilità dei singoli punteggi totalizzati all'interno del campione, di cui si riportano nella tabella i valori minimi e massimi registrati. Nello specifico emerge che il punteggio medio più alto sia stato totalizzato nella classe di **Murcia**, pari a 15,8 che, confrontato con i punteggi medi totalizzati nelle altre sedi, risulta anche essere la media più affidabile, dato che il suo valore della deviazione standard (1,4) è inferiore a quello delle altre classi. Nella classe di **Breslavia**, infatti, il punteggio medio totalizzato è pari a 15 con deviazione standard pari a 3,4, mentre–nella classe di **Malaga** il punteggio medio è pari a 13,4 e il valore della deviazione standard si attesta ugualmente a 3,4.

Tabella 6.8: Esiti dell'attività finale sull'uso del congiuntivo vs indicativo

	Partecipanti	Punteggio medio totalizzato	Deviazione standard	Punteggio min. e max. totalizzato in 20esimi	Percentuale risposte corrette
Breslavia anno II	4	15	3,4	12 - 18	75%
Murcia	5	15,8	1,4	14 - 18	79%
Malaga	7	13,4	3,4	7 - 17	67,1%

Grafico 7: Esiti dell'attività finale sull'uso del congiuntivo vs indicativo



Attività extra: indicativo vs congiuntivo nell'interrogativa indiretta

Questa fase contiene delle attività di rinforzo proposte al termine dell'unità come approfondimento. È stata svolta solo presso l'Università di **Murcia**, dove la tempistica dello svolgimento della lezione ha consentito l'inserimento anche di queste attività. Gli esiti prodotti mostrano il pieno raggiungimento degli obiettivi: su un punteggio totale di 14 punti, il punteggio medio totalizzato è pari a 13 (deviazione standard pari a 1), compreso tra un minimo di 12 e un massimo di 14.

6.3.3. Esiti delle attività finali sui verbi riflessivi e sull'uso del congiuntivo vs indicativo

In questa sezione verrà svolto un confronto degli esiti delle attività di verifica svolte al termine delle due unità sperimentate. Prima di mostrare i dati raccolti e al fine di contestualizzarli, si ricorda che nelle attività finali di verifica (rispettivamente *Applicazione*

pratica 3 e Uso del congiuntivo: formalità vs informalità) viene chiesto agli studenti di riutilizzare le strutture grammaticali su cui è incentrata l'unità e fare delle riflessioni di tipo socio-pragmatico (ossia, scegliere di utilizzare o meno una determinata struttura linguistica in base allo specifico contesto comunicativo e al registro formale o informale ad esso correlato).

Allo scopo di confrontare i dati ottenuti, per ciascuna classe di sperimentazione è stata calcolata la percentuale delle risposte corrette totalizzate nelle attività finali di verifica, che permette di fare un confronto degli esiti sia a livello globale tra le diverse sedi coinvolte, sia all'interno di ciascuna sede tra le due attività svolte. L'analisi comparativa escluderà sia la classe del I anno di Breslavia sia la classe di Madrid, dove – come già anticipato – l'unità didattica sul congiuntivo non è stata svolta, poiché l'argomento grammaticale non era ancora stato affrontato nel loro percorso curricolare.

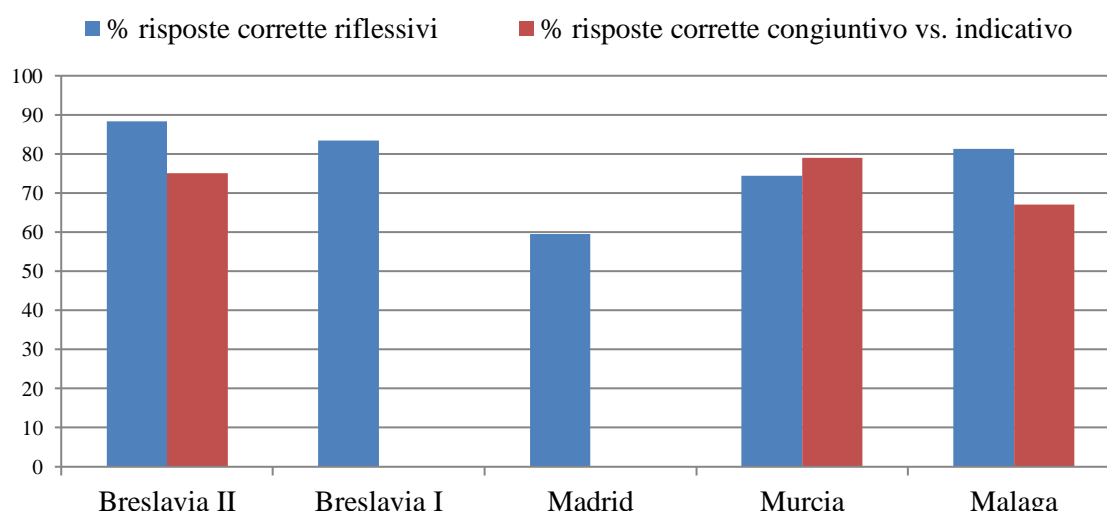
Innanzitutto, per ciascuna classe di sperimentazione è stato calcolato il rapporto tra le risposte corrette date da ciascun partecipante e il totale dei quesiti dell'attività. Ad esempio, nel caso dell'attività finale sui verbi riflessivi, un partecipante ha totalizzato 13 risposte corrette su 15; il rapporto tra risposte corrette e quesiti totali è pari a $13/15 = 0,8666667$.

Dopo aver svolto questa operazione per i punteggi di tutti i partecipanti, è stata calcolata la media dei rapporti ottenuti e, infine, da questa è stata ricavata la percentuale. Le percentuali ottenute sono state utilizzate per realizzare il grafico 8 e sono state inserite nella seguente tabella riassuntiva:

Tabella 6.9: Confronto delle percentuali delle risposte corrette nelle due attività finali di verifica

	Percentuale risposte corrette	
	Attività finale sui verbi riflessivi "affettivi"	Attività finale su congiuntivo vs indicativo
Breslavia I	83,5%	La sperimentazione non è stata svolta
Breslavia II	88,3%	75%
Madrid	59,6%	La sperimentazione non è stata svolta
Murcia	74,4%	79%
Malaga	81,3%	67,1%

Grafico 8: Confronto delle percentuali delle risposte corrette nelle due attività finali di verifica



L'analisi di tali dati è stata integrata dalle risposte date dai partecipanti al quesito *Da quanto tempo studi l'italiano?*, contenuto nel questionario sulle opinioni degli studenti, che sarà illustrato nel dettaglio nella sezione 6.4. Da queste risposte innanzitutto verranno ricavate informazioni generali sulla durata media dello studio pregresso per ciascuna classe. Inoltre, per ciascuna classe di sperimentazione verranno presi in considerazione il punteggio più alto e il punteggio più basso ottenuto nelle due attività di verifica, confrontandoli ove possibile⁹ con la durata dello studio pregresso dell'italiano dichiarata dai singoli partecipanti che li hanno totalizzati. Ciò ha l'obiettivo di determinare se esista un rapporto tra il periodo di studio pregresso e i risultati ottenuti nello svolgimento delle attività.

Prima di commentare i dati, si precisa che il gruppo-classe del II anno dell'università di Breslavia è l'unico in cui le stesse persone hanno partecipato ad entrambe le sperimentazioni, mentre i gruppi-classe di Malaga hanno conservato solo parzialmente la stessa composizione. Di conseguenza, solo nel caso di Breslavia può essere svolta un'effettiva analisi di tipo contrastivo su come lo stesso gruppo di persone abbia svolto l'una e l'altra unità didattica.

⁹ In alcuni casi non è stato possibile attribuire le informazioni sulla durata pregressa dello studio dell'italiano alla persona che ha totalizzato il punteggio più alto o più basso nelle verifiche, poiché alcuni studenti non hanno utilizzato lo stesso nome utente per accedere alle unità didattiche e, successivamente, al questionario. Ciò è dovuto al fatto che, per preservare la privacy degli studenti nella compilazione del questionario sulle opinioni, è stata data loro libera scelta di utilizzare o meno dei nomi utente riconducibili alla loro identità, come ulteriore forma di tutela rispetto alla possibilità offerta al docente dalla piattaforma *Socrative* di visualizzare o oscurare in qualunque momento i nomi utente dei partecipanti al questionario, a seconda delle esigenze della ricerca.

Osservando i dati dell'unità sui verbi riflessivi, emerge che la percentuale più alta, pari all'88,3%, è stata totalizzata dalla classe del II anno di Breslavia, segue la classe di Breslavia del I anno con l'83,5%, Malaga con l'81,3, Murcia con il 74,4% e, infine, Madrid con il 59,6 %.

Per quanto riguarda invece il congiuntivo il punteggio più alto è stato ottenuto dall'università di Murcia con il 79% di risposte corrette, segue il II anno di Breslavia con il 75% e, infine, Malaga con il 67,1%.

Al fine di verificare il rapporto tra la durata dello studio pregresso e gli esiti delle attività di verifica, sono stati osservati i risultati ottenuti dai singoli partecipanti, in particolare per ciascuna classe sono stati osservati i casi delle persone che hanno riportato i punteggi più bassi e più alti. Incrociando questi dati con le informazioni ottenute dal questionario circa la durata dello studio pregresso, nella sperimentazione dell'unità sul congiuntivo non emergono – tranne nel caso di Murcia – differenze sostanziali tra i risultati ottenuti dai componenti della stessa classe aventi una maggiore o minore durata pregressa dello studio dell'italiano, bensì nel confronto tra le diverse sedi.

In generale, è emerso che all'interno della stessa classe persone aventi alle spalle una durata simile, o identica in alcuni casi, hanno totalizzato punteggi molto diversi tra loro (ad esempio, i componenti delle classi di Malaga), mentre nel confronto tra le diverse sedi è emerso che le percentuali più alte delle risposte corrette sono state totalizzate nelle sedi che presentano una durata più lunga dello studio pregresso (ad esempio, i casi di Breslavia e Malaga per i verbi riflessivi e di Murcia per il congiuntivo).

Entrando nel dettaglio, nel caso del II anno dell'università di **Breslavia** si nota che la percentuale di risposta più alta è stata ottenuta nell'unità didattica dei verbi riflessivi (88,3 %) rispetto all'unità sull'uso dell'indicativo vs congiuntivo (75%). Questa differenza può essere spiegata sia in base al fatto che l'uso standard dei verbi riflessivi era già stato svolto al primo anno di corso, sia in base alle risposte fornite nel questionario di valutazione, dove tutti i partecipanti affermano di aver già trovato in precedenza esempi d'uso dei verbi riflessivi in senso "affettivo". Dunque per quanto riguarda l'uso standard dei verbi riflessivi, si tratta di un contenuto già consolidato nella competenza linguistica degli apprendenti del secondo anno, a cui la sperimentazione ha aggiunto ulteriori informazioni e ha fornito un inquadramento teorico circa l'uso riflessivo "affettivo" del neostandard. Al contrario, il tema del congiuntivo è stato trattato in tempi più recenti rispetto alla sperimentazione e, data anche la sua complessità, è comprensibile che permangano ancora dei dubbi (come emergerà dalle risposte date nel questionario nella sezione 6.4).

Inoltre, la classe risulta omogenea dal punto di vista della durata dello studio pregresso della lingua italiana, poiché su un totale di 4 persone la durata media dello studio della lingua italiana è di 2,5 anni, con un valore minimo di 2 anni e massimo di 3 anni e tale omogeneità si rispecchia anche nei singoli risultati ottenuti dai partecipanti nella sperimentazione sui verbi riflessivi “affettivi”.

In questo caso il punteggio più alto è pari 14/15 e i restanti partecipanti hanno totalizzato 13/15. I nomi utente utilizzati sulla piattaforma *Wizer.me* non hanno un corrispondente tra quelli usati per accedere al questionario e dunque non consentono di identificare i partecipanti e attribuire loro i dati sulla durata dello studio dell’italiano. Tuttavia, data la differenza minima sia in termini di durata dello studio che di punteggi ottenuti, questa lacuna risulta irrilevante nella riflessione sul rapporto tra la durata dello studio e i risultati ottenuti.

Nella sperimentazione sul congiuntivo *vs* indicativo, invece, il punteggio più alto raggiunto è pari a 18/20 è stato totalizzato da K.K. e K.J., che dichiarano entrambe di studiare l’italiano da 3 anni, ossia la durata massima dichiarata nella classe. Il punteggio più basso è pari a 12/20 ed è stato totalizzato da due partecipanti a cui non è possibile attribuire i dati sulla durata dello studio dell’italiano, poiché i loro nomi utente sulla piattaforma *Wizer.me* non hanno un corrispondente tra quelli usati per accedere al questionario.

Una riflessione analoga può essere fatta nel caso dell’università di **Malaga**, dove la percentuale delle risposte corrette è pari all’81,3% nel caso dell’unità sui verbi riflessivi. Confrontando questi dati con la durata dello studio dell’italiano, emerge che questa classe è molto disomogenea al suo interno dal punto di vista dello studio pregresso: su un totale di 5 persone la durata media dello studio della lingua italiana è di 2,4 anni, con un valore minimo di 1 anno e massimo di 6 anni.

Tuttavia, come per la classe di Breslavia, poiché l’argomento dei verbi riflessivi viene normalmente trattato tra il livello A2 e il B1, anche le persone aventi un solo anno alle spalle di studio dell’italiano sono riuscite a svolgere con successo l’attività, proprio perché l’argomento era probabilmente già noto. Ciò è confermato dalle risposte fornite nel questionario di valutazione, dove tutti i partecipanti affermano di aver già trovato in precedenza esempi d’uso dei verbi riflessivi in senso “affettivo”. In questo caso non si nota una differenza sostanziale in termini di rapporto tra la durata dello studio pregresso e i risultati ottenuti, poiché il punteggio più basso, ma comunque superiore alla sufficienza e pari a 10/15, è stato raggiunto da S.N. (1 anno di studio) e da S.L. (2 anni di studio) e che

altre due persone - O.R.G. e O.V. - aventi anch'esse 2 anni di studio pregresso dell'italiano hanno totalizzato il punteggio più alto, pari a 14/15.

Nella classe di Malaga che ha partecipato all'unità didattica sull'uso del congiuntivo *vs* indicativo, la percentuale delle risposte corrette è pari al 67,1%. Dal punto di vista della competenza linguistica, la classe, composta da 7 persone, è piuttosto omogenea al suo interno: la durata media dello studio della lingua italiana è di 1,6 anni, con un valore minimo di 1 anno e massimo di 2 anni. Il punteggio più alto, pari a 17/20 è stato totalizzato dalle uniche due persone che hanno partecipato ad entrambe le sperimentazioni, aventi entrambi 2 anni di studio: S.L. che nell'altra sperimentazione aveva totalizzato il punteggio più basso, e di O.R.G., che aveva totalizzato il punteggio più alto anche nell'altra sperimentazione. La persona che ha totalizzato il punteggio più basso (R.D., con punteggio pari a 7/20) non ha compilato il questionario. Poiché non ci sono a disposizione dati sulla sua durata dello studio pregresso, in sostituzione sono stati osservati i punteggi delle persone che hanno dichiarato di aver studiato l'italiano per un solo anno: da ciò emerge che C.B. e M. J. hanno totalizzato entrambe un punteggio superiore alla sufficienza, pari rispettivamente a 13/20 e 14/20.

Nel caso di **Murcia**, alla sperimentazione dell'unità sui verbi riflessivi hanno partecipato 6 persone del III anno, 2 delle quali hanno partecipato anche alla sperimentazione dell'unità didattica sull'uso del congiuntivo *vs* indicativo, insieme a 3 persone del IV anno. La percentuale più elevata delle risposte corrette è quella dell'attività finale sull'uso del congiuntivo *vs* indicativo (79%), caso unico tra le sperimentazioni condotte. Una possibile spiegazione può essere ricercata nei dati sullo studio pregresso della lingua italiana. Infatti, la classe che ha partecipato alla sperimentazione dell'unità sull'uso del congiuntivo *vs* indicativo era composta da 5 persone con una lunga durata dello studio pregresso della lingua italiana, nello specifico la durata media dello studio della lingua italiana è di 3,2 anni, con un valore minimo di 2 anni¹⁰ e un valore massimo di 4 anni, ove anche la durata più breve tra quelle indicate presuppone che il tema dell'uso standard del congiuntivo sia già stato affrontato e che le attività possano essere svolte agevolmente da tutti. Ciò si conferma osservando i punteggi totalizzati, ove in questo caso può essere notato il rapporto tra la durata dello studio pregresso e i risultati ottenuti: il punteggio più alto, pari a 18/20 è stato ottenuto da R.O.G., con 4 anni di studio pregresso,

¹⁰ Nelle risposte date al questionario non è stato chiarito come mai una persona iscritta al III anno abbia dichiarato 2 anni e non 3 di conoscenza pregressa dell'italiano. Si ipotizza un cambiamento nel piano di studi nell'arco del quadriennio o un errore nel conteggio del periodo di studio.

mentre A.O.B., con 2 anni di studio pregresso, ha totalizzato il punteggio più basso, ma comunque superiore alla sufficienza, pari a 14/20.

Nella classe che ha partecipato alla sperimentazione dell'unità sui verbi riflessivi, dei 6 partecipanti del III anno 4 hanno studiato l'italiano per 2 anni e 2 l'hanno studiato per 3 anni. In questo caso, invece, non si nota una differenza sostanziale in termini di rapporto tra la durata dello studio pregresso e i risultati ottenuti: sia A.S., che ha totalizzato il punteggio massimo pari a 15/15, sia A.O.B., che ha totalizzato anche in questa sperimentazione il punteggio più basso, pari a 6/15, hanno entrambe 2 anni di studio pregresso.

Per quanto riguarda, invece, le classi in cui è stata svolta la sola sperimentazione dell'unità sui riflessivi, merita attenzione la percentuale delle risposte corrette pari all'83,5% totalizzata dalla classe del **I anno di Breslavia**, separata di soli 5 punti dagli esiti prodotti dalla classe del II anno dello stesso ateneo (88,3%). Un risultato così elevato potrebbe essere stato determinato da due fattori: innanzitutto il fatto che l'argomento era stato trattato recentemente nell'ambito dell'insegnamento di lingua italiana curriculare, trattandosi del primo anno di studio della laurea triennale; in secondo luogo, all'interno della classe sono presenti alcune persone con una lunga durata di studi pregressi e un caso di bilinguismo familiare, nello specifico la durata media dello studio della lingua italiana è di 27,6 mesi¹¹, con un valore minimo di 7 mesi e un massimo di 120 mesi (ossia, 10 anni).

A questo proposito è interessante notare che anche le persone con una durata dello studio pregresso inferiore ad un anno, compresa tra i 7 e 9 mesi, hanno condotto una sperimentazione di successo, con degli esiti assolutamente più che sufficienti: si citano i casi di A.K. e H.M. che, pur avendo soli 8 mesi pregressi, hanno totalizzato un punteggio pari a 13/15. Non è stato possibile verificare la durata dello studio pregresso delle 2 persone che hanno totalizzato il punteggio massimo pari a 15/15, poiché i nomi utente utilizzati sulla piattaforma *Wizer.me* non hanno un corrispondente tra quelli usati per accedere al questionario e ciò ha impedito la loro identificazione. Il punteggio più basso totalizzato, pari a 9/15, è stato totalizzato da due persone, per una delle quali (J. B.) è nota la durata di un anno di studio pregresso.

Nel caso di **Madrid**, la classe ha totalizzato il punteggio più basso tra le sperimentazioni svolte, pari al 59,6%. Per quanto riguarda la durata degli studi pregressi dell'italiano, su 15 persone la durata media dello studio della lingua italiana è di 14,8 mesi,

¹¹ È stata effettuata la conversione in mesi della durata dello studio pregresso, data l'eterogeneità delle risposte raccolte.

con un valore minimo di 0, nel caso di un principiante assoluto, e un valore massimo di 2 anni.

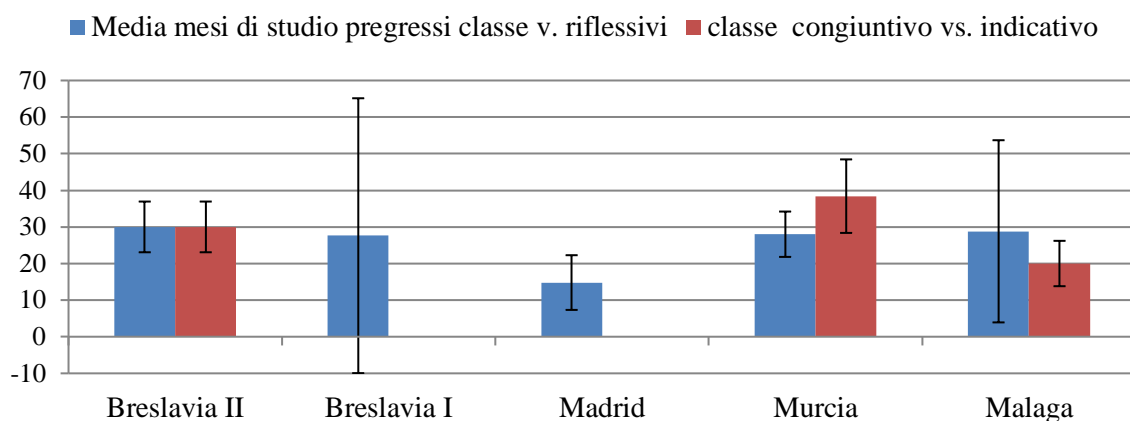
Nella fase preliminare alla sperimentazione, il docente referente a Madrid aveva messo in luce un livello di competenza pari ad A2 raggiunto al termine dell'anno accademico precedente, precisando inoltre che i tre mesi di pausa estiva intercorsa tra la fine del primo anno e lo svolgimento della sperimentazione (settembre 2019) avrebbero potuto incidere negativamente sulle prestazioni dei componenti della classe. Per quanto riguarda il punteggio più basso, pari a 4/15 e ottenuto da 3 persone, solo nel caso di E. è noto che si tratti di un caso di principiante assoluto. Non è stato possibile, invece, ricavare informazioni sulla durata dello studio pregresso né delle altre 2 persone che hanno totalizzato lo stesso punteggio, né di J., che ha totalizzato il punteggio più alto pari a 14/15. In sostituzione di quest'ultimo si è osservato che E. D. E., una delle 3 persone aventi la durata di studio pregresso più lunga, pari a 2 anni, ha ottenuto un punteggio pari a 12/15, mentre anche per le altre 2 persone non è stato possibile risalire alle informazioni sullo studio pregresso per la medesima questione dei nomi utente non coincidenti tra le due piattaforme.

I dati più volte citati circa la durata dello studio pregresso, utilizzati per integrare l'analisi dei risultati ottenuti nelle attività di verifica finale, sono tratti dalle risposte dei partecipanti al quesito *Da quanto tempo studi l'italiano?*, contenuto nel questionario sulle opinioni degli studenti, come già anticipato all'inizio della sezione. Data l'eterogeneità delle risposte raccolte, che variano tra pochi mesi a molti anni di studio, al fine di permettere il confronto dei dati, è stata effettuata una conversione degli anni in mesi e, successivamente, è stata calcolata la media dei mesi di studio pregresso dell'italiano, come illustrato nella tabella 6.10 e nei grafici 9, 10 e 11, che forniscono una visione d'insieme funzionale al confronto tra le classi di sperimentazione di entrambe le unità didattiche.

Tabella 6.10: Media dei mesi di studio pregresso della lingua italiana

	Classe di sperimentazione verbi riflessivi "affettivi"		Classe di sperimentazione congiuntivo vs indicativo	
	Media mesi	Deviazione standard	Media mesi	Deviazione standard
Breslavia I	27,6	37,5	La sperimentazione non è stata svolta	
Breslavia II	30	6,9	30	6,9
Madrid	14,8	7,4	La sperimentazione non è stata svolta	
Murcia	28	6,1	38,4	10
Malaga	28,8	24,8	20	6,1

Grafico 9: Confronto della media dei mesi di studio



Inoltre, i grafici 10 e 11 permettono di focalizzare l'attenzione separatamente sui valori della media dei mesi e la relativa deviazione standard di ciascuna classe di sperimentazione.

Grafico 10: Media mesi di studio pregressi nelle classi di sperimentazione dell'unità sui verbi riflessivi "affettivi"

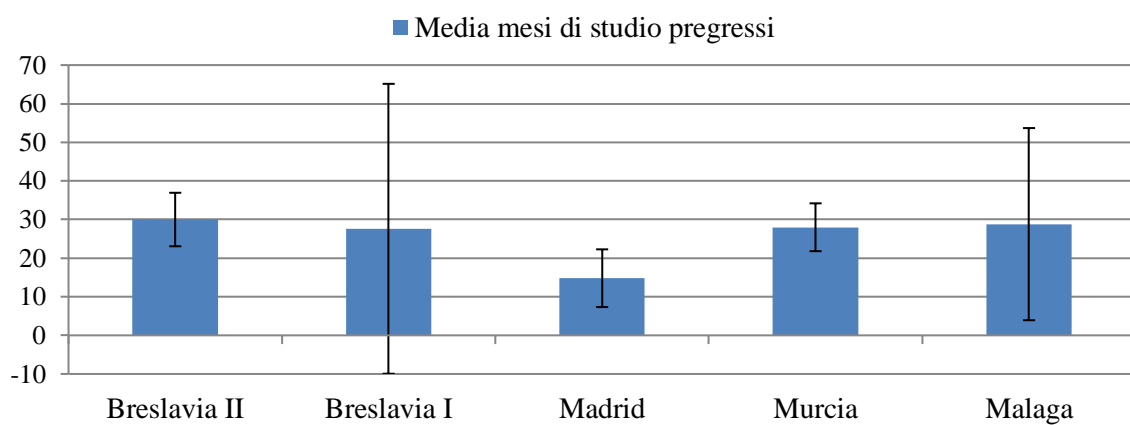
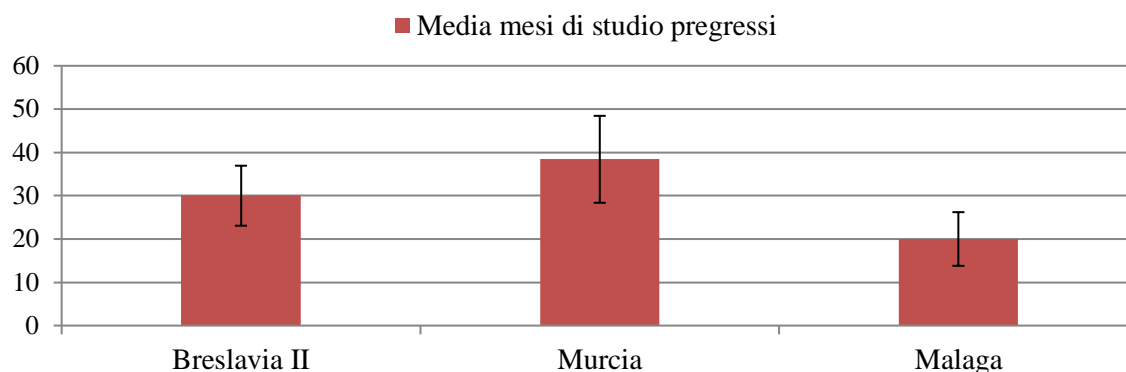


Grafico 11: Media mesi di studio pregressi nelle classi di sperimentazione dell'unità sul congiuntivo vs indicativo



Osservando le tabelle e i grafici precedentemente menzionati, emerge quanto segue. Come già detto in precedenza, il gruppo-classe del **II anno** dell'università di **Breslavia** è l'unico ad aver partecipato ad entrambe le sperimentazioni conservando la stessa composizione interna, dunque i dati raccolti nelle due somministrazioni del questionario sono gli stessi: la media dei mesi di studio pregressi dei 4 partecipanti è pari a 30 (deviazione standard:6,9), dove 2 persone hanno indicato 24 mesi e 2 hanno indicato 36.

Per quanto riguarda l'unità didattica sui verbi riflessivi "affettivi", dati molto simili a quelli menzionati sono stati ottenuti nella classe di **Murcia**, dove la media dei mesi è pari a 28 (deviazione standard: 6,), e i partecipanti hanno indicato una durata compresa tra un minimo di 24 mesi e un massimo di 36. Osservando il grafico 10, anche il valore medio ottenuto nella classe di **Malaga** è molto simile a quelli delle due classi già citate, pari a 28,8 mesi; tuttavia, il valore della deviazione standard pari a 24,8 indica che c'è una maggiore variabilità nel campione, compreso tra un minimo di 12 mesi e un massimo di 72 (ossia, 6 anni).

Particolarmente degni di nota i casi di **Madrid**, dove la media dei mesi è pari a 14,8 (deviazione standard:7,47), e i valori dichiarati variano tra un minimo di zero di un principiante assoluto – unico caso riscontrato – e un massimo di 24 mesi, e della classe del **I anno** di **Breslavia**, dove la media dei mesi è pari a 27,6 (deviazione standard:37,5): il valore estremamente elevato della deviazione standard si spiega probabilmente alla luce del fatto che quella classe presenta il maggior livello di eterogeneità relativamente alla durata dei singoli percorsi di studio della lingua italiana. La durata dichiarata dai singoli partecipanti oscilla, infatti, tra un minimo di 7 mesi e un massimo di 10 anni (120 mesi).

Per quanto riguarda, invece, l'unità didattica sulla scelta tra congiuntivo e indicativo (grafico 11), verranno omessi i dati del II anno di Breslavia, già citati in precedenza. Nella classe di **Murcia** la media dei mesi è pari a 38,4, la più elevata in assoluto tra quelle registrate, con una deviazione standard pari a 10 (la variabilità nel campione è inclusa tra un minimo di 24 e un massimo di 48 mesi). Nel caso di **Malaga**, invece, la media dei mesi è pari a 20, il valore della deviazione standard è pari a 6,1 la variabilità nel campione è inclusa tra un minimo di 12 e un massimo di 24 mesi.

A questo punto è possibile tracciare una conclusione sull'analisi delle attività finali delle unità didattiche. Per quanto riguarda le classi che hanno partecipato alla sperimentazione delle unità didattiche sui verbi riflessivi "affettivi", si può notare che tra le classi di Breslavia (I e II anno), Malaga e Murcia non vi sono molte differenze, sia in termini di durata media del periodo di studio pregresso sia in termini di profitto ottenuto. L'unico caso che si discosta da questi è la classe di Madrid, dove in effetti ad una durata media dello studio pregresso della lingua italiana inferiore rispetto a quella delle altre classi, pari a 14,8 mesi, corrisponde anche un profitto inferiore, con la percentuale delle risposte corrette pari al 59,6%.

Nel caso della sperimentazione dell'unità didattica dedicata al congiuntivo, la graduatoria delle classi che hanno totalizzato il profitto migliore (Murcia, Breslavia II anno e Malaga) rispecchia perfettamente l'ordine decrescente della durata media del periodo di studio pregresso, con Murcia che si attesta a 38,4 mesi, Breslavia II anno a 30 mesi e Malaga a 20.

Dalla correzione emerge che gli obiettivi delle unità sono stati globalmente raggiunti: in entrambe le attività di verifica i punteggi medi ottenuti in ciascuna classe sono superiori alla sufficienza; inoltre, i partecipanti sono in grado di reimpiegare le strutture grammaticali sulle quali è incentrata l'unità e fare delle riflessioni di tipo socio-pragmatico (ossia, scegliere di utilizzare o meno una determinata struttura linguistica in base allo specifico contesto comunicativo e al registro formale/informale ad esso correlato). Salvo poche eccezioni, gli errori rilevati sono principalmente di tipo morfologico (ad esempio, dubbi nella coniugazione dei verbi, nel sistema pronominale ecc.) o relativi a strutture linguistiche che non rappresentavano il *focus* dell'unità.

Confrontando in particolare il grafico 8, che rappresenta la percentuale di risposte corrette totalizzate nelle due attività conclusive dalle diverse classi, e il grafico 9 sulla durata dello studio pregresso della lingua italiana, emerge che tali grafici sono quasi speculari: le classi che hanno svolto con migliore profitto le attività conclusive presentano

anche una durata pregressa maggiore dello studio, il che lascerebbe intendere che l'uso standard di entrambi i temi grammaticali trattati fosse un contenuto già consolidato nella competenza linguistica degli apprendenti, a cui la sperimentazione ha aggiunto ulteriori informazioni e ha fornito un inquadramento teorico circa l'uso nel neostandard. Tuttavia, tali riflessioni saranno arricchite da informazioni aggiuntive nella sezione successiva, dedicata alla descrizione degli esiti del questionario per la raccolta delle opinioni dei partecipanti sulle attività svolte. Tali dati consentiranno di approfondire la riflessione sull'efficacia delle unità didattiche progettate: oltre ai criteri finora già menzionati, ossia i risultati ottenuti nelle singole attività e la durata dello studio pregresso della lingua verranno prese in considerazione anche la conoscenza di altre lingue straniere, la presenza nella L1 e nelle altre lingue straniere conosciute di argomenti grammaticali simili e la percezione di utilità delle attività svolte e della piattaforma *Wizer.me* che le ospita.

6.4. Esiti del questionario sulle opinioni degli studenti

Dopo aver illustrato gli esiti delle unità didattiche, nelle sezioni successive saranno descritti i dati relativi alle opinioni dei partecipanti sulle attività, raccolti tramite la piattaforma *Socrative* (cfr. capitolo 4). Di seguito si dedicheranno le sezioni 6.4.1-6.4.3 all'unità sui verbi riflessivi "affettivi" e le sezioni 6.4.4-6.4.6 all'unità sul congiuntivo *vs* indicativo. Nell'analisi delle risposte date, la trattazione sarà suddivisa in base alle tre parti del questionario: indagine conoscitiva, opinioni sulle attività e opinioni sull'uso della piattaforma *Wizer.me*; per ciascuna sede di sperimentazione verrà dedicato un paragrafo alle risposte date ai primi 7 quesiti a risposta aperta del questionario, mentre i dati raccolti nei restanti quesiti a scelta multipla saranno anche rappresentati graficamente.

6.4.1. Unità didattica sui verbi riflessivi "affettivi": indagine conoscitiva dei partecipanti

I seguenti paragrafi saranno dedicati alle risposte date ai primi 7 quesiti a risposta aperta del questionario, che permetteranno di tracciare un profilo dei partecipanti e della loro familiarità con il tema grammaticale dell'unità.

Tabella 6.11: Profilo dei partecipanti della classe di Breslavia (II anno)

Media mesi di studio pregressi	30
Deviazione standard	6,9
Percentuali risposte corrette nella verifica finale	88,3%
L1 e LS	- L1: polacco; - LS: inglese (4), francese (3), hindi (1), spagnolo (1), tedesco (1).
Utilità della L1/LS nella comprensione del tema?	- Sì (2) la L1, perché ha un argomento grammaticale simile; - un po' (1); - no (1)
Hai già incontrato esempi dell'uso neostandard?	Sì (tutti)

Il questionario è stato somministrato a 4 persone del II anno della laurea triennale, tutte di lingua materna polacca, che indicano le seguenti lingue straniere conosciute, per ciascuna delle quali è riportato tra parentesi il numero di persone: l'inglese (4), il francese (3), l'hindi (1), lo spagnolo (1) e il tedesco (1) (quesiti 1-2). Ciò che emerge osservando i dati raccolti nella tabella 6.11 è l'ipotesi di un rapporto tra la L1, che presenta una struttura simile, la competenza della totalità della classe in due lingue romanze indicate nelle risposte (3/4 indicano la conoscenza del francese e 1/4 dello spagnolo) e il raggiungimento di una percentuale elevata di risposte corrette. A questo proposito, al quesito 6 *Esiste un fenomeno simile nella tua lingua?* 2 persone hanno affermato di sì, una delle due ha precisato: *Sì, ma funziona un po' diverso* e le restanti 2 persone hanno risposto di no¹², mentre al quesito 7 *La tua lingua madre o altre lingue che conosci ti hanno aiutato/a a imparare/capire meglio l'argomento di grammatica di oggi?* 2 persone hanno risposto di sì (*sì; sì, un po'*), una *un po'*, una *no*. A ciò si accompagna una durata dello studio pregresso della lingua italiana tale da permettere la trattazione dell'argomento nel percorso curriculare antecedente alla sperimentazione: 2 persone studiano l'italiano da 2 anni e 2 da 3 anni ed è ipotizzabile che questa differenza di un anno dipenda da percorsi di studio individuale precedenti alla laurea triennale (quesiti 1-3).

Inoltre, osservando i dati emerge la conoscenza passiva dell'uso dei verbi riflessivi in senso affettivo, seppur priva di una consapevolezza teorica: tutti, infatti, hanno affermato di averne già trovato degli esempi in precedenza rispetto alla sperimentazione (quesito 4). A questo proposito in risposta al quesito 5, in cui viene chiesto di indicare uno o più

¹² Dal confronto con il docente referente dello scambio a Breslavia è emerso che in polacco esiste l'uso del verbo riflessivo "affettivo".

contesti d'uso, 2 persone hanno indicato *TV, giornali, libri ecc*, 3 hanno indicato la risposta *Sui social network* e una persona ha indicato *altro*, senza tuttavia specificare di cosa si tratti.

Università di Breslavia – I anno

Tabella 6.12: Profilo dei partecipanti della classe di Breslavia (I anno)

Media mesi di studio pregressi	27,6
Deviazione standard	37,5
Percentuali risposte corrette nella verifica finale	83,5%
L1 e LS	- L1: polacco (tutti), polacco e italiano (1); - LS: inglese (13), francese (3), russo (3), tedesco (3), spagnolo (2), hindi (1), svedese (1)
Utilità della L1/LS nella comprensione del tema?	- Sì (12) la L1, ha un argomento grammaticale simile; - no (2)
Hai già incontrato esempi dell'uso neostandard?	- Sì (11) - No (3)

Il questionario è stato somministrato a 14 persone del I anno della laurea triennale, tutte di lingua materna polacca, tranne un caso in cui è stata indicato il bilinguismo polacco-italiano.

Osservando i dati, a fronte dell'elevata percentuale di risposte corrette, la classe appare molto disomogenea dal punto di vista della durata dello studio pregresso della lingua italiana ed è caratterizzata da una presenza molto ristretta di parlanti di lingue romanze. Rispetto al primo punto, la durata oscilla tra un minimo di 7 mesi e un massimo di 10 anni, a cui si aggiunge un caso di bilinguismo familiare. Nel dettaglio sono state date le seguenti risposte, per ciascuna delle quali è indicato tra parentesi il numero di persone: 1 anno (4); 8 mesi (4), sempre (genitore italiano) (1), 10 anni (1), 8 anni (1), 4 anni (1), 9 mesi (1) e 7 mesi (1).

Per quanto riguarda, invece, le lingue straniere conosciute oltre all'italiano, si nota all'interno della classe una scarsa diffusione delle lingue romanze: inglese (13), francese (3), russo (3), tedesco (3), spagnolo (2), hindi (1) e svedese (1).

Inoltre, osservando i dati emerge la conoscenza passiva dell'uso dei verbi riflessivi in senso affettivo, seppur priva di una consapevolezza teorica: 11 persone hanno affermato di aver già trovato degli esempi di uso dei verbi riflessivi in senso "affettivo", mentre 3 hanno affermato di no. Alcuni hanno aggiunto le seguenti spiegazioni, per le quali è stata

rispettata la grafia originale, senza apportare correzioni e la stessa scelta verrà applicata da qui in avanti: *sì, ma non sapevo che venga usato dai Italiani così spesso; sì, un professore d'italiano ci ha detto questo; un po', ma nessuno me li ha mai spiegati; sì, mi fumo una sigaretta; sì, ma non l'ho saputo per 100%; no, non ho trovato degli esempi di uso dei verbi riflessivi in senso rafforzativo mai; no, a scuola ci fanno fare esercizi con l'uso "appropriato" del riflessivo e su internet non sono mai andata a guardare altri esercizi al riguardo.*

In risposta al quesito 5, in cui viene chiesto di indicare uno o più contesti d'uso, il 24% ha indicato *Parlando con gli italiani*, il 20% *TV, giornali, libri ecc.*, il 16% *Sui social network*, il 16% *Chat/messaggi* e il restante 24% ha indicato *altro*, senza tuttavia spiegare a cosa si riferisse.

A fronte della disomogeneità della durata dello studio pregresso e scarsa diffusione delle lingue romanze tra i partecipanti, anche in questo caso si fa avanti l'ipotesi che la percentuale elevata di risposte corrette possa essere stata influenzata dalla presenza nella L1 di una struttura grammaticale simile. Alla domanda *Esiste un fenomeno simile nella tua lingua?* la totalità dei partecipanti ha risposto *sì*; nello specifico tre persone hanno aggiunto un commento: *Sì, c'è qualcosa abbastanza simile in polacco; Sì, c'è la stessa cosa; Ci sono molti fonemi simili nel polacco, ma non so come spiegarli.* Infine, alla domanda *La tua lingua madre o altre lingue che conosci ti hanno aiutato/a a imparare/capire meglio l'argomento di grammatica di oggi?* 12 persone hanno risposto di *sì*, alcuni hanno motivato la risposta dicendo che in polacco esiste un fenomeno grammaticale simile (*Si, in polacco esiste l'argomento uguale; Si, la mia lingua madre ha la stessa cosa in grammatica; Si', polacco mi ha aiutato; Si, certamente. La conoscenza del francese o anche dell'inglese mi fa diventare più saggio in questo caso*) e 2 hanno risposto di *no*.

Università Complutense di Madrid

Tabella 6.13: Profilo dei partecipanti della classe di Madrid

Media mesi di studio pregressi	14,8
Deviazione standard	7,4
Percentuali risposte corrette nella verifica finale	59,6%
L1 e LS	- L1: spagnolo (18), rumeno (1), finlandese (1) tedesco (1); - LS: inglese (17), tedesco (9), francese (6), spagnolo LS (3), catalano (2), galiziano (2), giapponese (1), olandese (1), portoghese (1), russo (1), svedese (1), valenciano (1).

Utilità della L1/LS nella comprensione del tema?	Sì (tutti): similarità presenti tra l'italiano e lo spagnolo.
Hai già incontrato esempi dell'uso neostandard?	- Sì (13) - No (4)

Il numero di partecipanti al questionario oscilla, a seconda dei quesiti, tra le 15 e le 20 unità¹³. 15 partecipanti su 18 sono di lingua materna spagnola, 1 di lingua materna rumena, 1 finlandese e 1 tedesco. Dalla tabella emerge che la classe ha totalizzato la percentuale più bassa di risposte corrette tra tutte le classi di sperimentazione, nonostante la presenza dei seguenti fattori: la prevalenza di partecipanti di L1 romanze (spagnolo e rumeno), la consapevolezza da parte dei partecipanti della presenza di strutture grammaticali simili nella L1 e la presenza nella classe di una grande varietà di lingue romanze tra le lingue straniere conosciute. Osservando nel dettaglio i fattori sopra menzionati, emerge quanto segue.

Alla domanda *Esiste un fenomeno simile nella tua lingua?* 12 persone hanno risposto sì, 4 persone non ne sono del tutto sicure (esempio: *Non ci avevo mai pensato. Diciamo in teoria sì, ma per me si manifesta in modo diverso*). Successivamente, alla domanda *La tua lingua madre o altre lingue che conosci ti hanno aiutato/a a imparare/capire meglio l'argomento di grammatica di oggi?* tutti hanno risposto di sì, motivando la risposta sulla base delle similarità presenti tra lo spagnolo e l'italiano, al quale, come anticipato, si aggiunge un'ampia varietà di lingue romanze tra le lingue straniere conosciute, come emerge dall'elenco sottostante, in cui tra parentesi è riportato il numero di parlanti: inglese (17), tedesco (9), francese (6), spagnolo LS (3), catalano (2), galiziano (2), giapponese (1), olandese (1), portoghese (1), russo (1), svedese (1) e valenciano (1).

Osservando i dati emerge la conoscenza passiva dell'uso dei verbi riflessivi in senso affettivo, seppur priva di una consapevolezza teorica: 13 persone hanno affermato di aver già trovato degli esempi di uso dei verbi riflessivi in senso "affettivo", alcuni dei quali hanno espresso le seguenti spiegazioni: *Lo avevo sentito ma non comprendevo il senso esatto di quest'uso; Gli esempi che ho visto sono "mi mangio, mi fumo una sigaretta, mi compro"*. 4 hanno risposto di no. In risposta al quesito 5, in cui viene chiesto di indicare uno o più contesti d'uso, il 30% ha risposto *Parlando con gli italiani*, il 26% *TV, giornali*,

¹³ Come già affermato in precedenza, si ricorda che il tempo a disposizione per la compilazione del questionario è stato nettamente inferiore al tempo stimato (10 minuti vs 20-25 min.). Le cause sono da ricercare sia nelle difficoltà tecniche che hanno caratterizzato la lezione, sia nel livello di competenza degli apprendenti, corrispondente ad un A2 – B1.1. in entrata. Le difficoltà di comprensione delle attività hanno, infatti, rallentato lo svolgimento della lezione e ridotto il tempo da dedicare al questionario, che è stato compilato rapidamente da un numero inferiore di persone rispetto al totale dei componenti della classe.

libri ecc., il 14% *Sui social network*, il 4% *Chat/messaggi* e il restante 26% ha indicato *altro*, senza tuttavia spiegare a cosa si riferisse.

Tuttavia, tutti i fattori sopra nominati sembrano aver influito sui risultati ottenuti in maniera più ridotta rispetto a quanto avvenuto nelle altre classi di sperimentazione e probabilmente ciò è dovuto alla durata media del periodo di studio pregresso notevolmente inferiore rispetto ai valori registrati nelle altre classi. A questo proposito sono state date, infatti, le seguenti risposte: 1 anno (8), 2 anni (5), 6 mesi (1), principiante assoluto (1), a cui si aggiungono 3 persone che non hanno specificato o hanno equivocato la domanda: 3 ore a settimana (2), 8 ore a settimana (1), da poco (1).

Università di Murcia

Tabella 6.14: Profilo dei partecipanti della classe di Murcia

Media mesi di studio pregressi	28
Deviazione standard	6,1
Percentuali risposte corrette nella verifica finale	74,4%
L1 e LS	L1: spagnolo; LS: francese (6), inglese (4)
Utilità della L1/LS nella comprensione del tema?	Sì (tutti): similarità presenti tra l'italiano e lo spagnolo.
Hai già incontrato esempi dell'uso neostandard?	Sì (3); No (2); Non so (1)

Il questionario è stato somministrato a 6 persone del III anno del corso di laurea quadriennale, tutte di lingua materna spagnola. Nella classe è stata ottenuta un'alta percentuale di risposte corrette nelle attività finali e osservando i dati raccolti, si può ipotizzare che, vi sia un rapporto tra i risultati ottenuti e una lunga durata media dello studio pregresso della lingua italiana, compresa tra i 2 e i 3 anni, indicati rispettivamente da 4 e 2 persone, a cui si accompagna anche il vantaggio costituito dalla L1 romanza e dalla conoscenza da parte di tutti i partecipanti di un'altra lingua straniera romanza, ossia il francese (6 persone). Segue la conoscenza dell'inglese, indicata da 4 persone. Alla domanda *Esiste un fenomeno simile nella tua lingua?* 5 persone hanno risposto *sì* e una persona *no*. Infine, alla domanda *La tua lingua madre o altre lingue che conosci ti hanno aiutato/a a imparare/capire meglio l'argomento di grammatica di oggi?* tutti hanno risposto di *sì*, motivando la risposta sulla base delle similarità presenti tra l'italiano e lo spagnolo.

Osservando i dati, emerge solo tra alcuni dei partecipanti la conoscenza passiva dell'uso dei verbi riflessivi in senso affettivo, seppur priva di una consapevolezza teorica: 3 persone hanno risposto di aver già trovato degli esempi di uso dei verbi riflessivi in senso "affettivo", 2 *no*, una persona non ne è sicura. In risposta al quesito 5, in cui viene chiesto di indicare uno o più contesti d'uso, una persona ha indicato tre risposte (*Parlando con gli italiani, TV, giornali, libri ecc., Sui social network*), 4 persone hanno indicato *altro*, senza tuttavia spiegare a cosa si riferisse, e una persona non ha dato alcuna risposta.

Università di Malaga

Tabella 6.15: Profilo dei partecipanti della classe di Malaga

Media mesi di studio pregressi	28,8
Deviazione standard	24,8
Percentuali risposte corrette nella verifica finale	81,3%
L1 e LS	- L1: spagnolo (3), francese (2); - LS: inglese (5), francese LS (2), spagnolo LS (2), portoghese (1)
Utilità della L1/LS nella comprensione del tema?	La L1: sì, (tutti); similarità presenti tra l'italiano e il francese (1)
Hai già incontrato esempi dell'uso neostandard?	Sì (tutti)

Il questionario è stato somministrato a 5 persone del II anno del corso di laurea quadriennale, 3 di lingua materna spagnola, 2 di lingua materna francese. Rispetto alle altre classi spagnole, nella classe di Malaga è stata ottenuta la più alta percentuale di risposte corrette nelle attività finali. Analogamente alla classe di Murcia, osservando i dati raccolti si può ipotizzare un rapporto tra i risultati ottenuti e la presenza di due L1 romanze (francese e spagnolo), oltre ad una lunga durata media dello studio progresso della lingua italiana, per la quale sono state date le seguenti risposte: 6 anni (1), 2 anni (2), 1 anno (2). A ciò si accompagna anche il vantaggio costituito dalla conoscenza da parte di tutti i partecipanti di almeno un'altra lingua straniera romanza: francese LS (2), spagnolo LS (2), portoghese (1), a cui si aggiunge per tutti la conoscenza dell'inglese. A questo proposito, tutti hanno affermato che esiste un fenomeno simile nella propria lingua madre e che questa o altre lingue conosciute sono state utili a imparare/capire meglio l'argomento di grammatica della sperimentazione. In particolare, una persona ha motivato la risposta sulla base delle similarità presenti tra italiano e francese.

Infine, rispetto alla conoscenza passiva del fenomeno, tutti hanno affermato di aver già trovato degli esempi d'uso dei verbi riflessivi in senso affettivo in precedenza rispetto alla

sperimentazione. A questo proposito in risposta al quesito 5, in cui viene chiesto di indicare uno o più contesti d'uso, una persona ha indicato *Parlando con gli italiani*, una persona ha indicato *TV, giornali, libri ecc.*, 2 persone hanno indicato *Sui social network* e 3 hanno indicato *altro*, senza dare ulteriori spiegazioni.

6.4.2. Unità didattica sui verbi riflessivi “affettivi”: opinioni sulle attività

Dopo aver tracciato un profilo dei partecipanti alle sperimentazioni tramite i dati raccolti nei quesiti 1-7, nei quesiti 8-25 viene chiesto ai partecipanti di esprimere la propria opinione sulle attività svolte sulla scala di valori (0) *no*, (1) *poco*, (2) *abbastanza*, (3) *sì* o di selezionare la voce *attività non svolta*, nel caso in cui quell'attività non sia stata proposta in classe. Nello specifico, i quesiti riguardano la percezione di utilità delle attività svolte (8-17), la percezione personale di acquisizione dei contenuti (18-21) e l'utilizzo della piattaforma (22-25). Le percentuali delle risposte verranno presentate nei grafici 12-17.

Di seguito la trattazione si focalizzerà sui dati riguardanti la percezione di utilità delle attività svolte (quesiti 8-17): 8) *È stata utile questa attività?* 8) Leggere il testo “L'appuntamento”; 9) Esercizi di collegamento (parole del testo – significato); 10) Rispondere alle domande sul tema del testo; 11) Cercare nel testo i verbi che descrivono la cura del corpo; 12) Completare la coniugazione dei verbi riflessivi; 13) Distinguere le diverse azioni espresse dai verbi riflessivi (*azione del soggetto su se stesso/ su parti del corpo o vestiti/accessori che appartengono al soggetto; azione reciproca tra più soggetti; riflessivo “affettivo”*); 14) Scegliere tra la forma riflessiva-affettiva e la forma non riflessiva; 15) Contenuto extra: i riflessivi nei modi di dire; 16) Produzione scritta e orale; 17) Fare esercizi basati su frasi vere della vita quotidiana (tratte da blog e altre pagine web) che non sono state inventate per “fare grammatica”.

Grafico 12: Percentuali delle risposte date nella classe di Breslavia II ai quesiti sulla percezione di utilità delle attività svolte

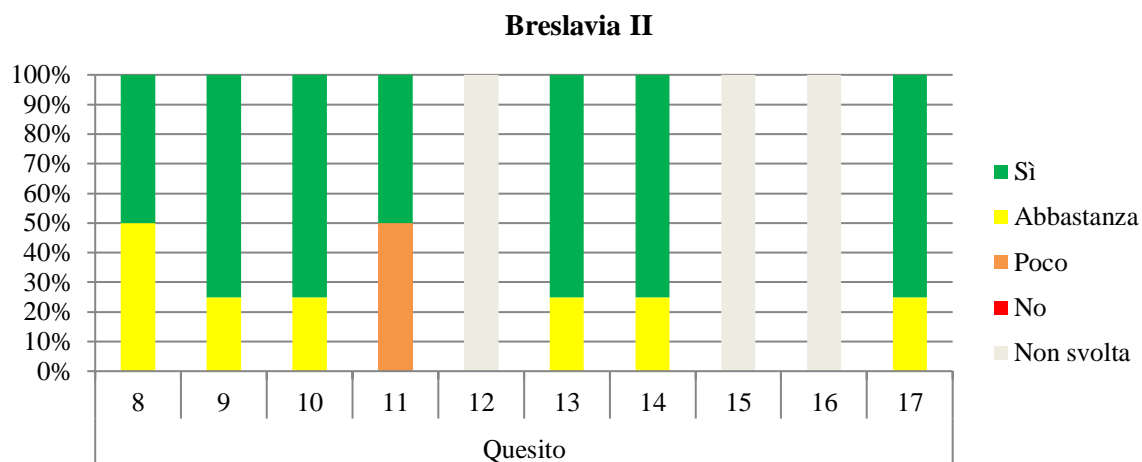


Grafico 13: Percentuali delle risposte date nella classe di Breslavia I ai quesiti sulla percezione di utilità delle attività svolte

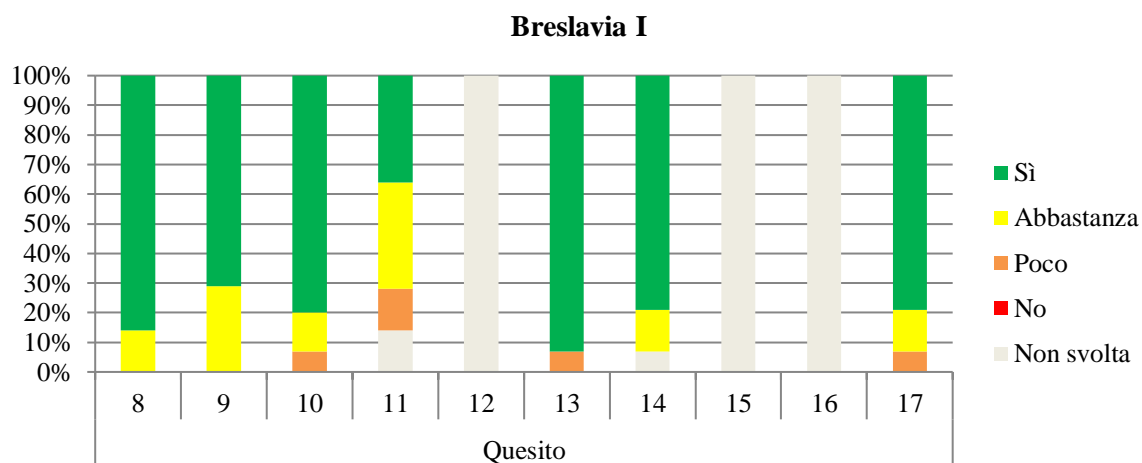


Grafico 14: Percentuali delle risposte date nella classe di Madrid ai quesiti sulla percezione di utilità delle attività svolte

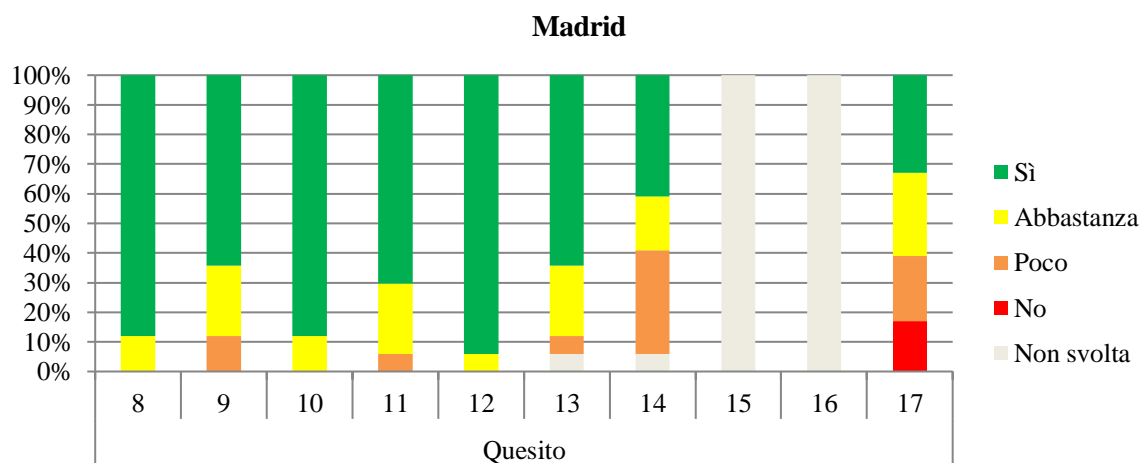


Grafico 15: Percentuali delle risposte date nella classe di Murcia ai quesiti sulla percezione di utilità delle attività svolte

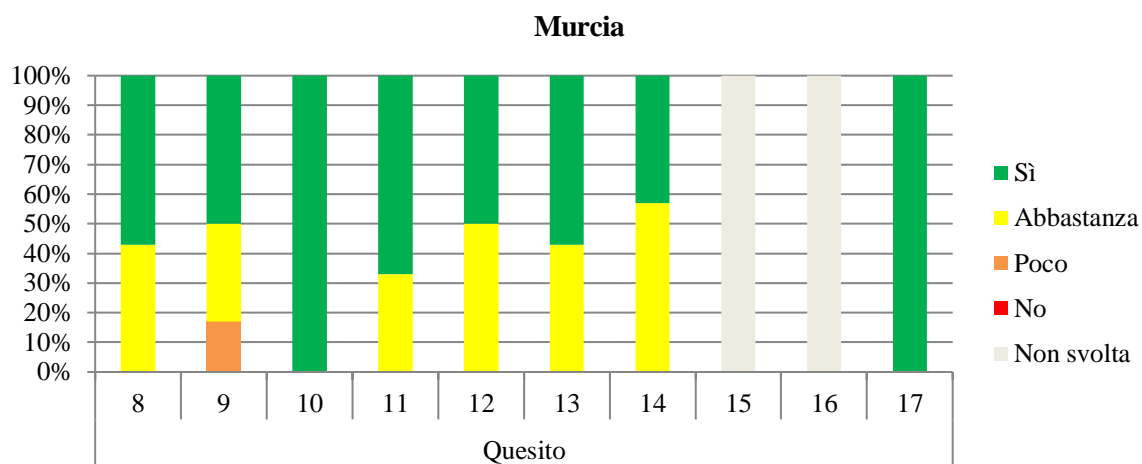


Grafico 16: Percentuali delle risposte date nella classe di Malaga ai quesiti sulla percezione di utilità delle attività svolte

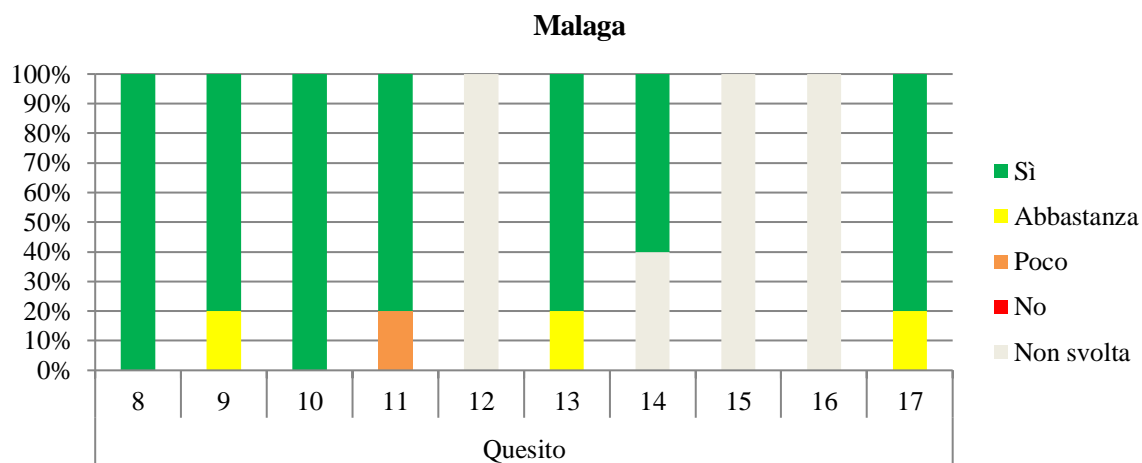
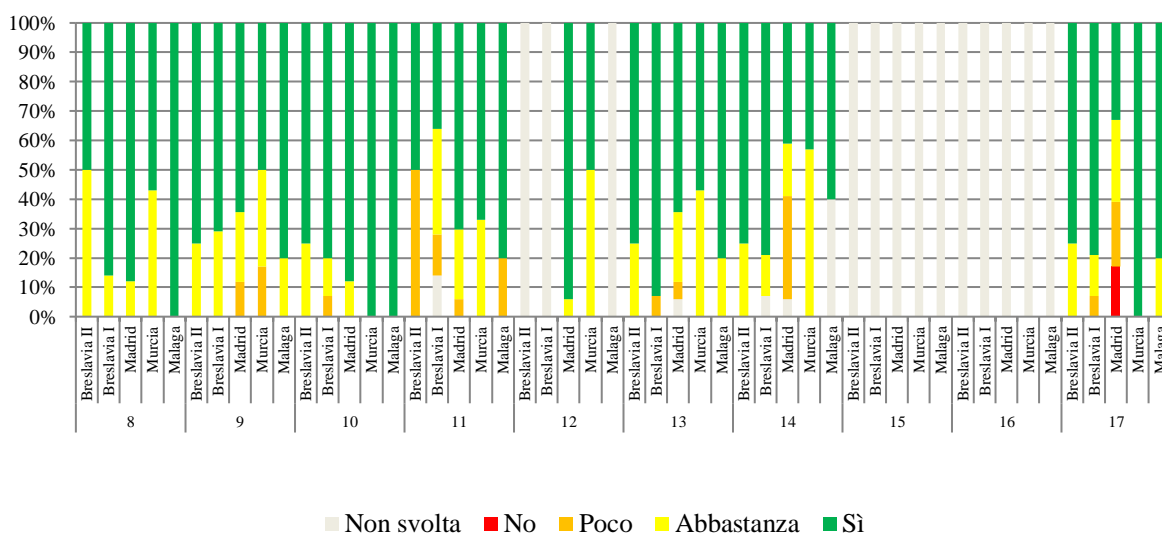


Grafico 17: Panoramica delle percentuali delle risposte date in tutte le classi di sperimentazione ai quesiti sulla percezione di utilità delle attività svolte



Osservando le percentuali delle risposte date¹⁴, emerge a livello globale una generale percezione di elevata utilità delle attività svolte. Di seguito verranno commentati alcuni casi particolari, che hanno trovato spiegazione nei commenti liberi lasciati dai partecipanti al termine delle attività. Al punto 11. *Cercare nel testo i verbi che descrivono la cura del corpo* il fatto che il 50% della classe del II anno di Breslavia e il 20% della classe di Malaga abbiano definito l'attività come *poco utile* è giustificabile sulla base del livello di competenza linguistica molto alto dei partecipanti, per i quali l'attività è stata percepita come banale. Viceversa, nel caso del punto 12 *Completare la coniugazione dei verbi riflessivi*, l'attività è stata percepita come *utile* dal 94% della classe di Madrid, dove il livello di competenza della classe era inferiore rispetto a quello delle altre classi partecipanti, pari ad un A2 pienamente raggiunto, e dunque l'attività di ripasso della coniugazione è stata ritenuta adeguata. Nel caso dei punti 13. *Distinguere le diverse azioni espresse dai verbi riflessivi* e 14. *Scegliere tra la forma riflessiva-affettiva e la forma non riflessiva* si ipotizza che la scelta dell'opzione *attività non svolta* (cfr. punto 13: Madrid 6%; punto 14: Breslavia I: 7%; Madrid: 6%; Malaga: 40%) derivi da una errata interpretazione della domanda da parte dei partecipanti, dato che le attività sono state proposte in tutte le sedi di sperimentazione e, di fatto, sono state svolte da tutti i partecipanti. Diverso, è invece, il caso dei quesiti 15 e 16, che si riferiscono ad attività pensate come supplementari (ossia: 15) *i riflessivi nei modi di dire*; 16) *Produzione scritta*

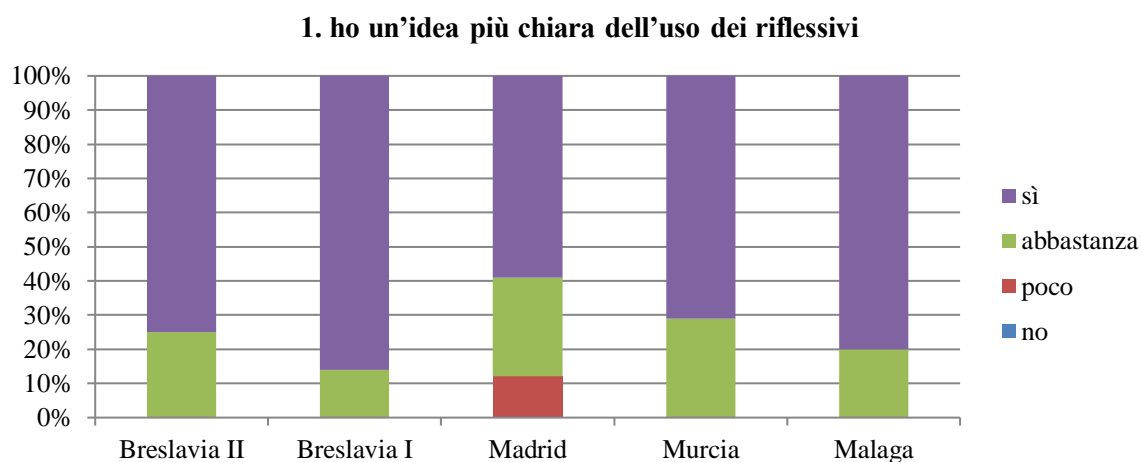
¹⁴ Si precisa che le percentuali del questionario sono state arrotondate all'intero dal sistema di calcolo presente nella piattaforma *Socrative*.

e orale) e che di fatto non sono state svolte in nessuna sede e, dunque, la scelta del descrittore non svolta corrisponde alla reale prassi didattica.

Di seguito la trattazione si focalizzerà sui dati riguardanti la percezione personale di efficacia delle attività svolte nell'acquisizione dei contenuti dell'unità (quesiti 18-21): *Dopo queste attività...: 18) ho un'idea più chiara dell'uso dei riflessivi; 19) riesco a distinguere meglio le situazioni formali da quelle informali; 20) ho capito che in alcune situazioni informali è accettabile usare i riflessivi "affettivi"; 21) ho ancora dei dubbi sull'uso dei riflessivi.*

Osservando il grafico 18 emerge che quasi la totalità dei partecipanti affermano di avere un'idea più chiara sull'uso dei riflessivi, con l'unica eccezione della classe di Madrid, dove vi è una scarsa percentuale (12%) che afferma di avere ancora un'idea poco chiara. A questo proposito va tenuto conto, come già affermato in precedenza, del livello di competenza inferiore rispetto ai componenti delle altre classi.

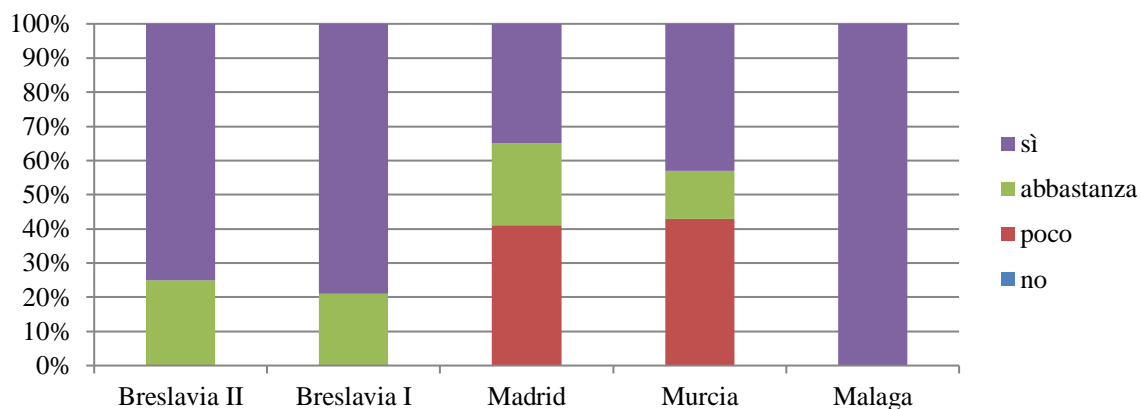
Grafico 18: Percentuali delle risposte date in tutte le classi di sperimentazione al primo quesito sulla percezione personale di acquisizione dei contenuti



Al secondo quesito (grafico 19) *riesco a distinguere meglio le situazioni formali da quelle informali*, la totalità dei partecipanti di Murcia e la maggioranza dei partecipanti di Breslavia (79% per il I anno e 75% per il II anno) afferma di *sì*, nel caso di Madrid il 41% ha risposto *poco*, seguito al 35% *sì* e dal 24% *abbastanza*. Questo dato va considerato alla luce delle considerazioni già fatte sul livello di competenza più basso degli apprendenti del gruppo classe di Madrid. Nel caso di Murcia vi è una parità di risposte (43%) per le voci *sì* e *poco* e la percentuale rimanente ha risposto *abbastanza*.

Grafico 19: Percentuali delle risposte date in tutte le classi di sperimentazione al secondo quesito sulla percezione personale di acquisizione dei contenuti

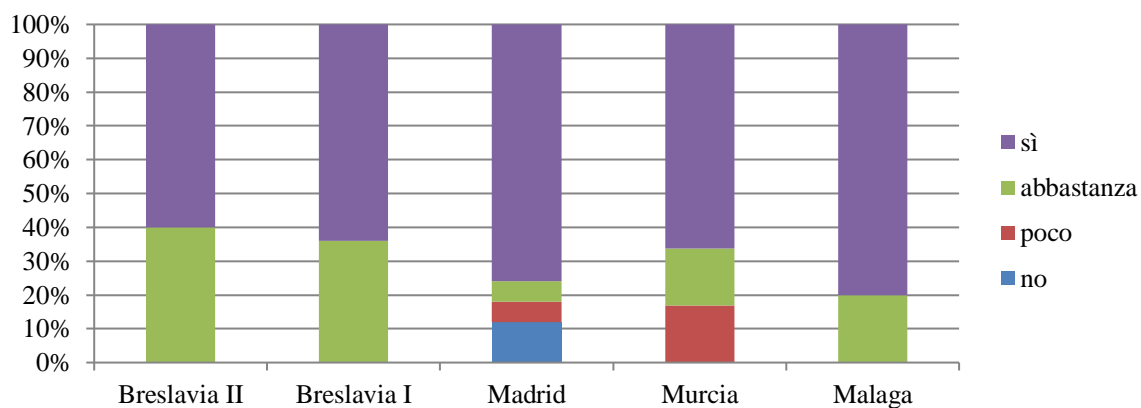
2. riesco a distinguere meglio le situazioni formali da quelle informali



Al terzo quesito (grafico 20) *ho capito che in alcune situazioni informali è accettabile usare i riflessivi "affettivi"*, in ciascuna classe la maggioranza delle risposte è stata attribuita alla voce *sì*, permangono dei dubbi ancora nelle classi di Madrid (12% *no*, 6% *poco* e *abbastanza*) e Murcia (17% *poco* e *abbastanza*).

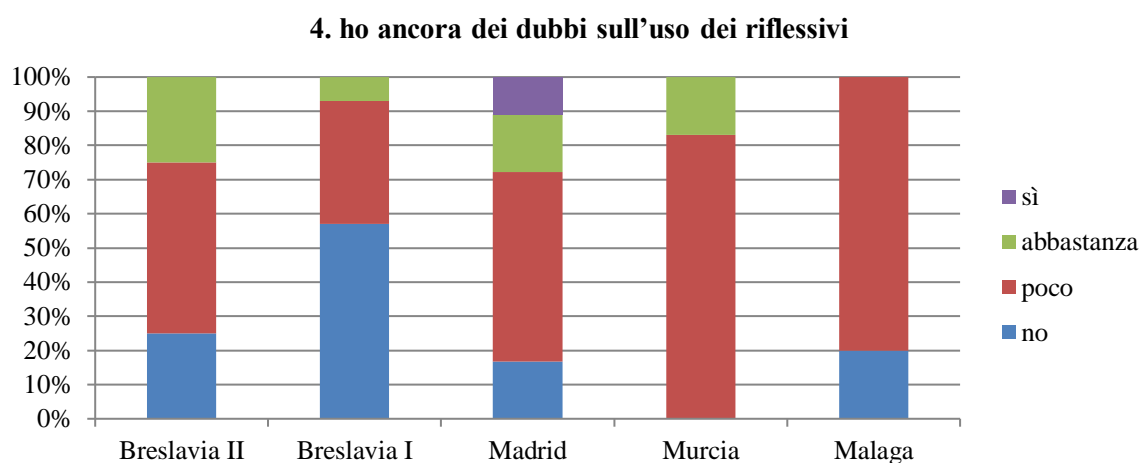
Grafico 20: Percentuali delle risposte date in tutte le classi di sperimentazione al terzo quesito sulla percezione personale di acquisizione dei contenuti

3. ho capito che in alcune situazioni informali è accettabile usare i riflessivi "affettivi"



Al quarto quesito (grafico 21) *Ho ancora dei dubbi sull'uso dei riflessivi "affettivi"*, nelle classi del II anno di Breslavia, Madrid, Murcia e Malaga prevale la risposta *poco* (rispettivamente pari al 50, 56, 83 e 80%), mentre nella classe del I anno di Breslavia prevale la risposta *no* (57%).

Grafico 21: Percentuali delle risposte date in tutte le classi di sperimentazione al quarto quesito sulla percezione personale di acquisizione dei contenuti



I dati ottenuti dai quesiti sulla percezione di utilità delle attività svolte e sulla percezione personale di acquisizione dei contenuti trovano conferma nella sezione del questionario dedicata ai commenti liberi, che vengono riportati qui di seguito. Nella trascrizione è stata rispettata la grafia originale, senza apportare correzioni, indicando tra parentesi la classe in cui sono stati formulati.

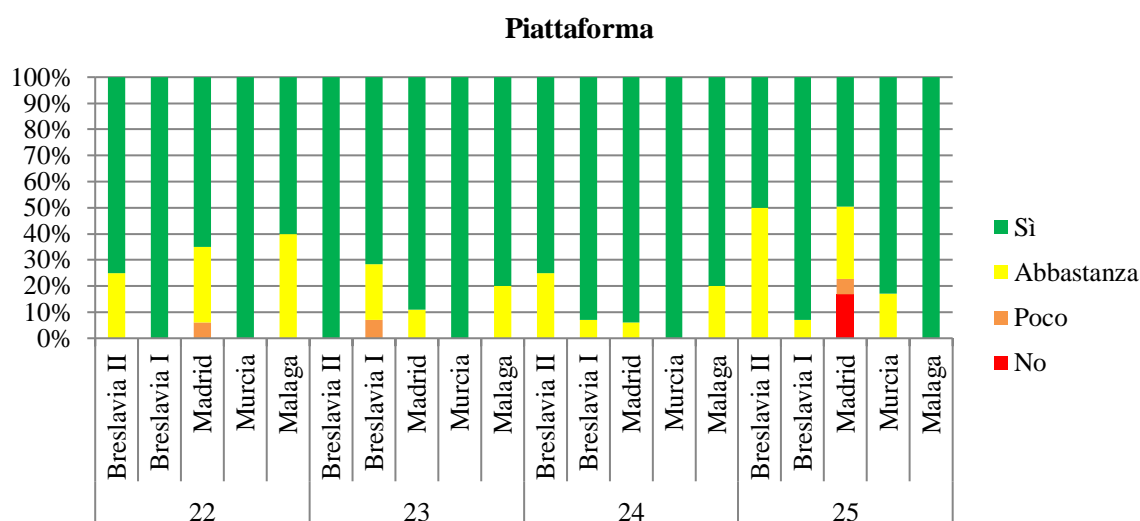
Alcuni di questi commenti esprimono un apprezzamento per l'operato della docente-ricercatrice o per la piacevolezza e utilità dell'attività (*L'attività è stata molto piacevole* :) (Breslavia II anno); *Era molto brava* (Breslavia I anno); *Tutto bene. Mi è piaciuta la lezione; Mi è piaciuta l'attività e ho imparato; Per me, l'esercizio è molto bene per imparare l'italiano* (Madrid); *Penso che siano state delle attività utili e, anzi, ci ha imparato qualcosa di nuovo; Secondo me questi esercizi sono molto utili per aiutarci a comprendere meglio questo tema* (Murcia).

In altri commenti vengono dati suggerimenti sullo svolgimento della lezione (*Penso che Lei possa aggiungere un po' più di conversazione, abbiamo abbastanza grammatica all'università. Vorremmo avere più contatto con il parlato* (Breslavia I anno); *più tempo per capire la grammatica!* (Madrid)), sulle preferenze personali circa le strategie di apprendimento (*È chiara per me la lezione di oggi, ma mi piace fare i esercizi dopo una spiegazione con un libro perché è più chiaro per me la teoria con appunti; la actitud esta muy bien hecha, pero me gustaria mas ejemplos y/o mas definiciones para poder comprenderlo mejor* (Madrid)) e, infine, sono stati inseriti dei commenti sulla grafica e la struttura delle unità, che verranno citati al termine di questa sezione.

6.4.3. Unità didattica sui verbi riflessivi “affettivi”: opinioni sulla piattaforma *Wizer.me*

Per quanto riguarda i quesiti sulla fruizione della piattaforma (22-25), dal grafico 22 emerge la generale soddisfazione degli utenti, le cui percentuali di risposta si dividono tra i descrittori *sì* e *abbastanza*¹⁵. Di seguito verranno commentate alcune particolarità.

Grafico 22: Panoramica delle percentuali delle risposte date in tutte le classi di sperimentazione ai quesiti sull'utilizzo della piattaforma



Circa il punto 22 *Il login e l'accesso alle diverse parti dell'unità è stato facile*, nella classe di Madrid il 6% dei partecipanti ha scelto la risposta *poco*, che presumibilmente si giustifica alla luce dello svolgimento della sperimentazione su dispositivi elettronici appartenenti ai partecipanti (in alcuni casi persino su *smartphone*) e del conseguente rallentamento dei tempi di connessione e apertura delle pagine rispetto all'utilizzo di un'aula computer già predisposta alle attività dall'insegnante.

Al punto 23 *Le istruzioni delle attività erano comprensibili* nella classe del I anno di Breslavia il 7% ha risposto *poco*: tale scelta rispecchia le richieste da parte di alcuni partecipanti di spiegazione delle consegne degli esercizi o di singole parole e trova spiegazione un commento lasciato da un partecipante nella sezione del questionario dedicata ai commenti liberi: *Secondo me, alcune consegne non erano chiare e poco*

¹⁵ Si ricorda che le percentuali del questionario sono state arrotondate all'intero dal sistema di calcolo presente nella piattaforma *Socrative*.

precise, in generale gli esercizi mi sono piaciuti ma le consegne appunto, potrebbero essere più precise.

Nel caso, invece, del punto 25 *È stato utile avere esercizi con correzione automatica*, solo nella classe di Madrid sono stati scelti i descrittori *no* (17%) e *poco* (6%), tuttavia queste risposte non trovano spiegazione nella sezione del questionario dedicata ai commenti liberi, dove invece sono stati formulati vari suggerimenti per il miglioramento delle attività sulla piattaforma, per i quali è stata rispettata la grafia originale, senza apportare correzioni.

Alcuni suggerimenti della classe di Madrid riguardano la progettazione strutturale e grafica delle attività: *Per me, può essere meglio di rafforzare il layout della pagina; Dovresti avere due opzioni in alcuni esercizi per quando ci sono due possibili risposte; Suggesto collocare tre opzioni di risposta nell' esercizio dove sono possibili più risposte.* Questo suggerimento è stato formulato in riferimento ad alcune attività in cui erano forniti nella consegna vari verbi all'infinito, chiedendo ai partecipanti di utilizzarne la forma coniugata per completare alcune frasi e tenendo conto del fatto che, per alcune frasi, fossero possibili più soluzioni: per questi casi specifici i partecipanti suggeriscono di fornire una scelta multipla, invece di scegliere le risposte da un unico elenco di opzioni valido per tutto l'esercizio. Un'alternativa alla scelta multipla viene fornita nella classe di Murcia, dove si suggerisce di progettare l'esercizio facendo in modo che i verbi utilizzati per completare le frasi non siano poi più selezionabili dall'elenco: *In generale penso che sarebbe meglio di potere togliere la risposta quando usi la parola nell'esercizio e potere mettere le diversi risposte.*

6.4.4. Unità didattica sul congiuntivo: indagine conoscitiva dei partecipanti

Analogamente a quanto fatto per l'unità sui verbi riflessivi "affettivi", per ciascuna sede di sperimentazione verrà dedicato un paragrafo alle risposte date ai primi 7 quesiti a risposta aperta del questionario, che permetteranno di tracciare un profilo dei partecipanti e della loro familiarità con il tema grammaticale dell'unità.

Tabella 6.16: Profilo dei partecipanti della classe di Breslavia (II anno)

Media mesi di studio pregressi	30
Deviazione standard	6,9
Percentuali risposte corrette nella verifica finale	75%
L1 e LS	- L1: polacco (tutti); - LS: inglese (4), francese (3), hindi (1), spagnolo (1), tedesco (1)
Utilità della L1/LS nella comprensione del tema?	Tutti (no): né la L1, né le LS, di cui alcuni dichiarano di avere un livello troppo basso
Hai già incontrato esempi dell'uso neostandard?	Sì (3), No (1)

La sperimentazione ha coinvolto le stesse 4 persone del II anno che hanno svolto l'unità didattica sui verbi riflessivi. Pertanto, verranno qui omessi i dati sulla popolazione relativi ai quesiti 1-3, già citati nella sezione dedicata alla sperimentazione dell'altra unità didattica.

Contrariamente a quanto osservato nella stessa classe per lo svolgimento dell'unità sui verbi riflessivi "affettivi", in questo caso non è possibile formulare l'ipotesi di un rapporto tra l'elevata percentuale delle risposte corrette ottenute e l'influenza della L1 e/o delle altre lingue romanze conosciute: alla domanda *Esiste un fenomeno simile nella tua lingua?* tutti hanno risposto di no¹⁶ e concordano nell'affermare che la lingua madre o altre lingue conosciute non siano state utili ad imparare/capire meglio l'argomento grammaticale della sperimentazione. In particolare, una persona ha motivato la risposta, dicendo: *Inglese non aiuta. Spagnolo potrebbe aiutarmi ma ho un livello della lingua troppo bassa.*

Emerge, invece, il dato della conoscenza passiva dell'uso dell'indicativo al posto del congiuntivo in contesti informali, a cui la sperimentazione ha fornito un inquadramento teorico: 3 persone su 4 hanno affermato di aver già trovato degli esempi d'uso dell'indicativo al posto del congiuntivo in precedenza rispetto alla sperimentazione. In risposta al quesito 5, in cui viene chiesto di indicare uno o più contesti d'uso, 3 persone hanno indicato *Parlando con gli italiani*, una persona ha indicato *TV, giornali, libri ecc.*, una ha indicato *Sui social network* e, infine, una persona ha indicato *altro*, senza tuttavia specificare di cosa si tratti.

¹⁶ "Il polacco non possiede la categoria del congiuntivo" (Sosnowski 2010:144).

Tabella 6.17: Profilo dei partecipanti della classe di Murcia

Media mesi di studio pregressi	38,4
Deviazione standard	10
Percentuali risposte corrette nella verifica finale	79%
L1 e LS	L1: spagnolo; LS: francese (5), inglese (3), portoghese (1)
Utilità della L1/LS nella comprensione del tema?	La L1: sì, (tutti); LS: similarità presenti tra italiano, spagnolo e francese (1)
Hai già incontrato esempi dell'uso neostandard?	Sì (2), No (3)

La sperimentazione ha coinvolto nella prima fase dell'unità 6 persone, ridotte poi a 5, a cui è stato somministrato il questionario alla fine della lezione: nel dettaglio, 3 persone del IV anno e 2 persone del III anno del corso di laurea quadriennale; queste ultime hanno partecipato anche alla sperimentazione dell'unità dedicata ai verbi riflessivi "affettivi".

Rispetto alle altre classi che hanno svolto la sperimentazione, nella classe di Malaga è stata ottenuta la più alta percentuale di risposte corrette nelle attività finali. Osservando i dati raccolti si può ipotizzare un rapporto tra i risultati ottenuti e una lunga durata media dello studio progresso della lingua italiana: 2 persone studiano l'italiano da 4 anni, 2 da 3 anni e una persona solo da 2 anni. A ciò si accompagna anche il vantaggio costituito dalla L1 romanza e dalla conoscenza da parte di tutti i partecipanti di almeno un'altra lingua straniera romanza: francese (5) e portoghese (1) a cui si aggiunge la conoscenza dell'inglese (3): tutti e 5 hanno affermato che esiste un fenomeno simile nella propria lingua materna e che questa o altre lingue conosciute sono state utili a imparare/capire meglio l'argomento di grammatica della sperimentazione. In particolare, una persona ha motivato la risposta sulla base delle similarità presenti tra italiano, spagnolo e francese. Dunque, si ipotizza che l'uso standard dei temi grammaticali trattato fosse un contenuto già consolidato nella competenza linguistica degli apprendenti, a cui la sperimentazione ha aggiunto informazioni circa l'uso neostandard in contesti informali già notato da alcuni nella vita quotidiana: alla domanda *Prima delle attività di oggi, avevi già trovato degli esempi di uso dell'indicativo al posto del congiuntivo?* hanno risposto di sì solo 2 persone su 5, che nel quesito successivo hanno indicato entrambe come contesto d'uso *Parlando con gli italiani*; in aggiunta una delle due ha indicato anche *TV, giornali, libri ecc.*

Tabella 6:18: Profilo dei partecipanti della classe di Malaga

Media mesi di studio pregressi	20
Deviazione standard	6,1
Percentuali risposte corrette nella verifica finale	67,1%
L1 e LS	L1: spagnola (5) francese (1); LS: inglese (6), francese (3), spagnolo (1).
Utilità della L1/LS nella comprensione del tema?	- Sì, la L1 (5), per le similarità presenti tra italiano, spagnolo; - no (1);
Hai già incontrato esempi dell'uso neostandard?	- Sì (3), - No (3)

La sperimentazione ha coinvolto 7 persone del II anno della laurea quadriennale: 6 di lingua materna spagnola e una di lingua materna francese. Di queste solo 6 hanno partecipato alla valutazione delle attività svolte tramite il questionario: 5 di lingua materna spagnola e la persona di lingua materna francese.

Rispetto alle altre classi coinvolte nella sperimentazione, la classe di Malaga presenta i valori più bassi sia della percentuale delle risposte corrette, sia della durata media dello studio pregresso: 4 persone studiano l'italiano da 2 anni e 2 persone da 1 anno. Riguardo a queste ultime, una delle due era in scambio *Erasmus* dalla Francia, dunque la breve durata pregressa dello studio dell'italiano dipende dal piano di studi dell'università di provenienza; nell'altro caso non è stata data una motivazione. Sulla base di questi dati, è possibile ipotizzare che vi sia un rapporto tra la percentuale di risposte corrette, inferiore a quella delle altre classi, e la breve durata pregressa dello studio dell'italiano, soprattutto nel caso dei partecipanti con un solo anno pregresso. Una simile durata non presuppone, infatti, che il tema dell'uso standard del congiuntivo sia già stato affrontato e che le attività possano essere svolte agevolmente.

Tale dato viene compensato dalla percezione di utilità delle altre lingue romanze conosciute: 5 persone su 6 affermano che la lingua materna o altre lingue conosciute – in particolare il francese LS (3) e lo spagnolo LS (1) – sono state utili a imparare/capire meglio l'argomento di grammatica e, in particolare, una persona ha motivato la risposta sulla base delle similarità presenti tra italiano e spagnolo (*sì, perché lo spagnolo anche ha il congiuntivo come tempo verbale*). Rispetto alla domanda *Esiste un fenomeno simile nella tua lingua?* 4 persone hanno risposto di *sì*, una *non lo so* e una persona si è astenuta.

Infine, per quanto riguarda la competenza passiva del fenomeno in questione, alla domanda *Prima delle attività di oggi, avevi già trovato degli esempi di uso dell'indicativo al posto del congiuntivo?* la classe si divide a metà tra 3 sì e 3 no. Nello specifico, le persone che hanno risposto affermativamente hanno poi selezionato nel quesito 5 i seguenti contesti d'uso, dove tra parentesi si riporta il numero di preferenze ricevute: *Parlando con gli italiani* (1), *TV, giornali, libri ecc.* (1), *Sui social network* (1), *Chat/messaggi* (1) e, infine, 2 persone hanno indicato *altro*, senza aggiungere altre spiegazioni¹⁷.

6.4.5. Unità didattica sul congiuntivo: opinioni sulle attività

Dopo aver tracciato un profilo dei partecipanti alle sperimentazioni tramite i dati raccolti nei quesiti 1-7, nei quesiti 8-20 viene chiesto ai partecipanti di esprimere la propria opinione sulla seguente scala di valori: (0) *no*; (1) *poco*; (2) *abbastanza*; (3) *sì* o di selezionare la voce *attività non svolta*, nel caso in cui quell'attività non sia stata proposta in classe. Le percentuali delle risposte date ai quesiti sulla percezione di utilità delle attività svolte (8-11), sulla percezione personale di acquisizione dei contenuti (12-15) e sull'utilizzo della piattaforma *Wizer.me* (16-20) verranno presentate nei grafici 23-31. Di seguito la trattazione si focalizzerà sui dati riguardanti la percezione di utilità delle attività svolte (quesiti 8-11) *È stata utile questa attività?:* 8) Cercare nei 4 brevi testi gli esempi d'uso del congiuntivo; 9) Ripassare i casi in cui il congiuntivo è obbligatorio; 10) Fare ipotesi sulla scelta tra congiuntivo e indicativo in base al contesto formale/informale; 11) Fare esercizi basati su frasi vere della vita quotidiana (tratte da blog e altre pagine web) che non sono state inventate per “fare grammatica”. Osservando le percentuali delle risposte date, emerge a livello globale, una generale percezione di elevata utilità delle attività svolte.

¹⁷ Per una eventuale revisione del questionario, si deduce la necessità di modificare la dicitura “Altro” con “Altro: spiega qui la tua risposta”.

Grafico 23: Percentuali delle risposte date nella classe di Breslavia II ai quesiti sulla percezione di utilità delle attività svolte

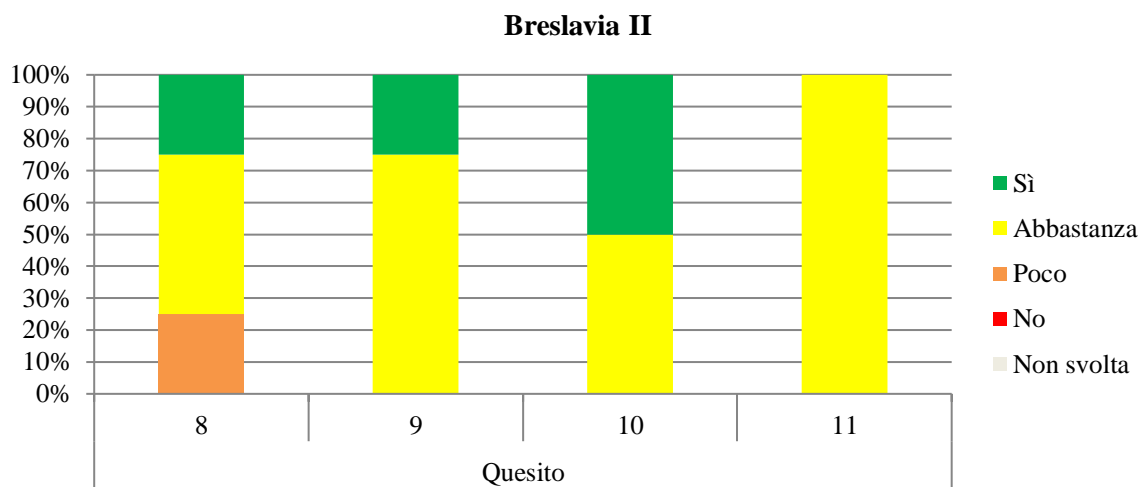


Grafico 24: Percentuali delle risposte date nella classe di Murcia ai quesiti sulla percezione di utilità delle attività svolte

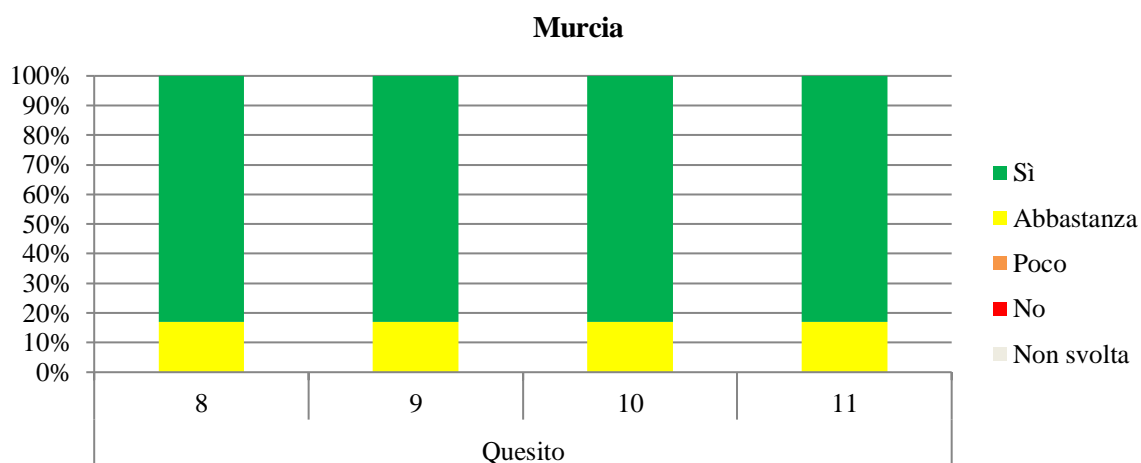


Grafico 25: Percentuali delle risposte date nella classe di Malaga ai quesiti sulla percezione di utilità delle attività svolte

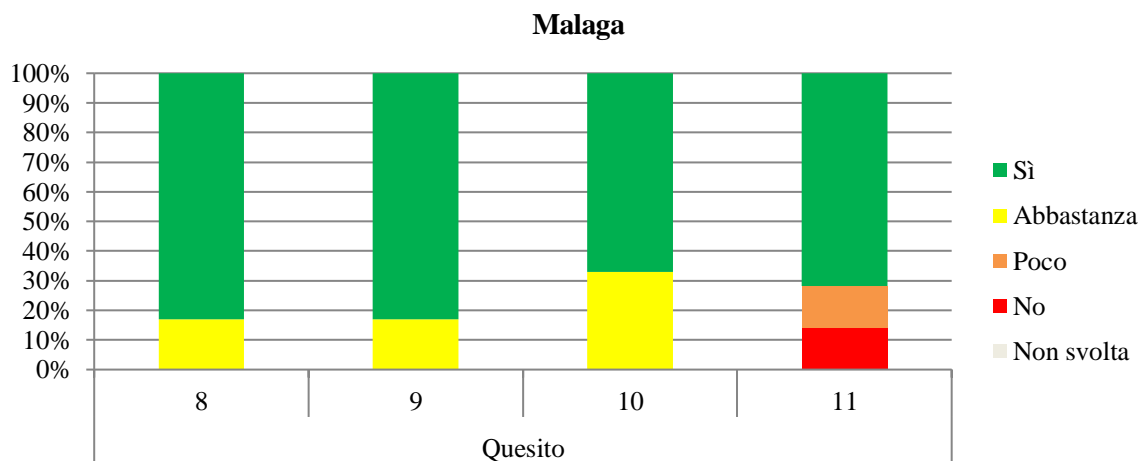
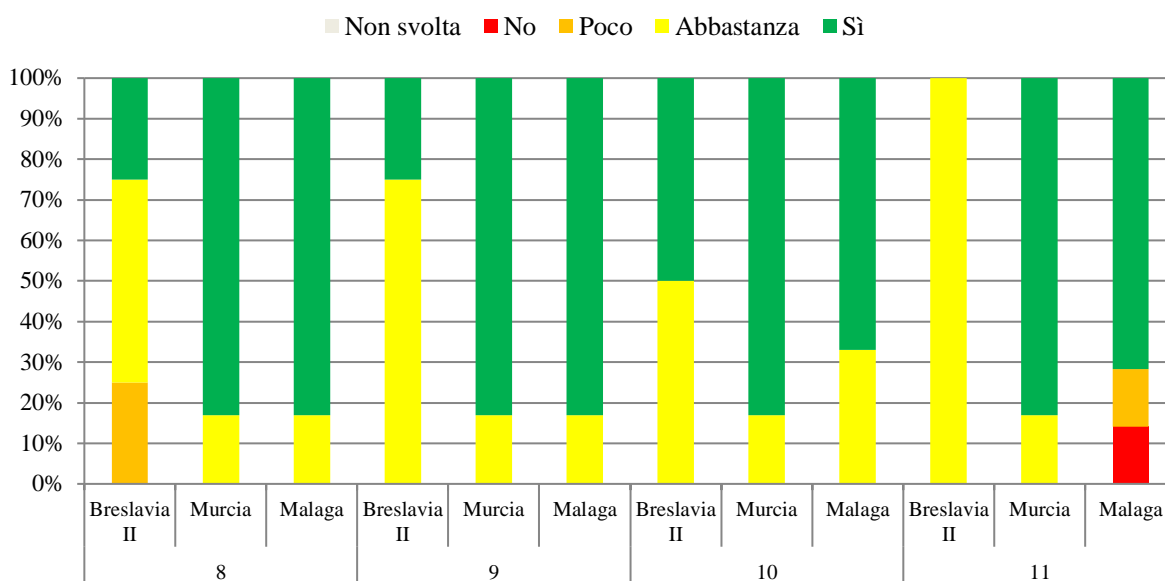


Grafico 26: Panoramica delle percentuali delle risposte date in tutte le classi di sperimentazione ai quesiti sulla percezione di utilità delle attività svolte



Di seguito verranno commentati solo due casi in cui le risposte comprendono anche i descrittori *poco* e *no*. Nel primo caso, rispetto al punto 8. *Cercare nei 4 brevi testi gli esempi d'uso del congiuntivo* nella classe di Breslavia le risposte date si suddividono tra *poco utile* (25%), *abbastanza utile* (50%) e *sì* (25%): data l'assenza di spiegazioni nella sezione dedicata ai commenti personali, si ipotizza che il 25% delle preferenze date alla voce *poco utile* possa essere dovuto alla durata pregressa dello studio della lingua dei componenti di questa classe (cfr. sezione 6.3.3), per cui la semplice individuazione nel testo delle forme verbali sia un'attività banale, non percepita come sfidante e come favorevole allo sviluppo delle competenze.

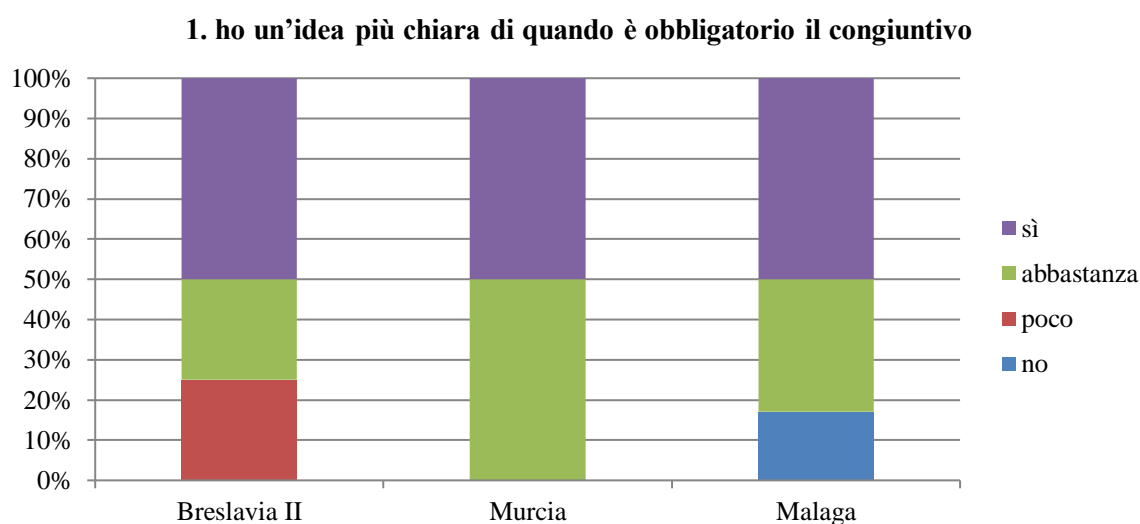
Nel secondo caso, rispetto al punto 11. *fare esercizi basati su frasi vere della vita quotidiana (tratte da blog e altre pagine web) che non sono state inventate per "fare grammatica"*, nella classe di Malaga le risposte date si suddividono tra *no* (14%), *poco utile* (14%) e *sì* (71%)¹⁸. Data anche in questo caso l'assenza di spiegazioni nella sezione dedicata ai commenti personali, si ipotizza che il 14% delle preferenze date alla voce *no* dipenda dal fatto che, semplicemente, dal punto di vista dell'apprendente l'uso di testi autentici non influisca nella percezione di utilità dell'attività, a differenza di altri fattori menzionati nel questionario (esempio, la chiarezza delle tracce degli esercizi).

¹⁸ Si ricorda, come già specificato in precedenza, che le percentuali del questionario sono state arrotondate all'intero dal sistema di calcolo presente nella piattaforma *Socrative*.

Di seguito la trattazione si focalizzerà sui dati riguardanti la percezione personale di efficacia delle attività svolte nell'acquisizione dei contenuti dell'unità (quesiti 12-15), ossia: *Dopo queste attività...: 12) ho un'idea più chiara di quando è obbligatorio il congiuntivo; 13) riesco a distinguere meglio le situazioni formali da quelle informali; 14) ho capito che in alcune situazioni informali è accettabile non usare il congiuntivo; 15) ho ancora dei dubbi sulla scelta tra congiuntivo e indicativo.*

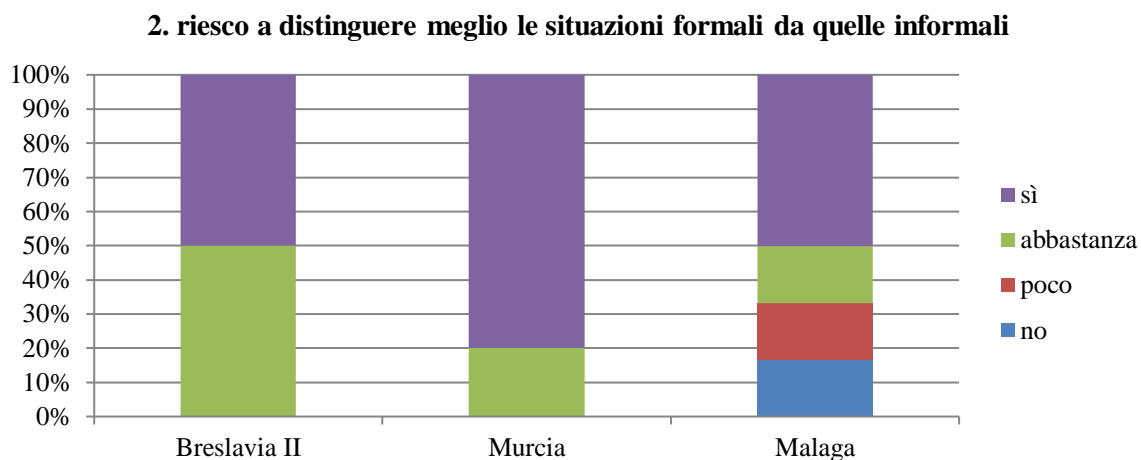
Al primo quesito (grafico 27) *ho un'idea più chiara di quando è obbligatorio il congiuntivo*, in tutte le classi di sperimentazione il 50% dei partecipanti ha risposto *sì*; nel caso di Murcia vi è parità tra le voci *sì* e *abbastanza*, mentre solo nella classe di Malaga il 17% ha risposto *no*.

Grafico 27: Percentuali delle risposte date in tutte le classi di sperimentazione al primo quesito sulla percezione personale di acquisizione dei contenuti



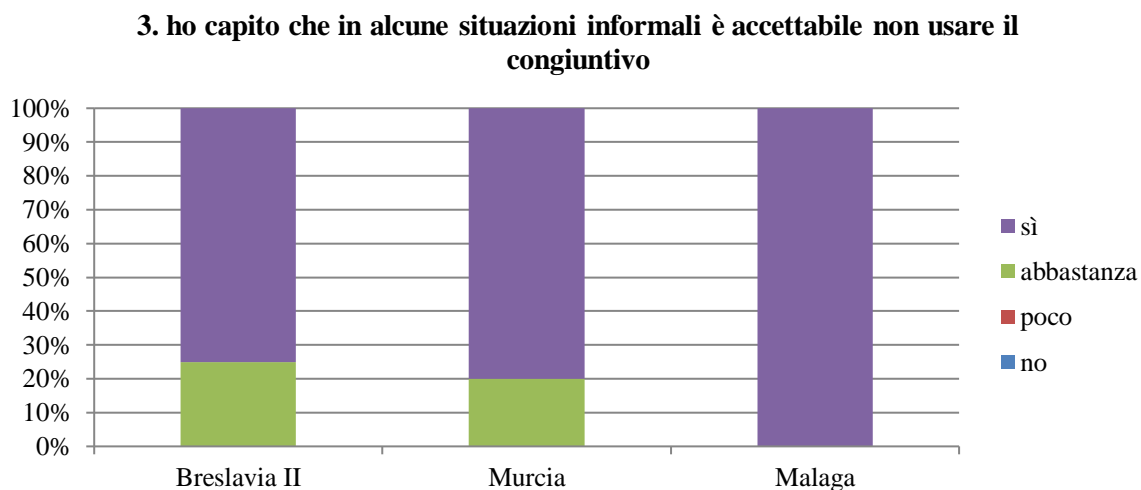
Al secondo quesito (grafico 28) *riesco a distinguere meglio le situazioni formali da quelle informali* nella classe di Murcia prevale il *sì* (80%), nella classe di Breslavia vi è parità tra le voci *sì* e *abbastanza*; restano alcuni dubbi nella classe di Malaga (le voci *no*, *poco* e *abbastanza* hanno ricevuto il 16,6% delle preferenze, mentre la maggioranza si è espressa per il *sì*).

Grafico 28: Percentuali delle risposte date in tutte le classi di sperimentazione al secondo quesito sulla percezione personale di acquisizione dei contenuti



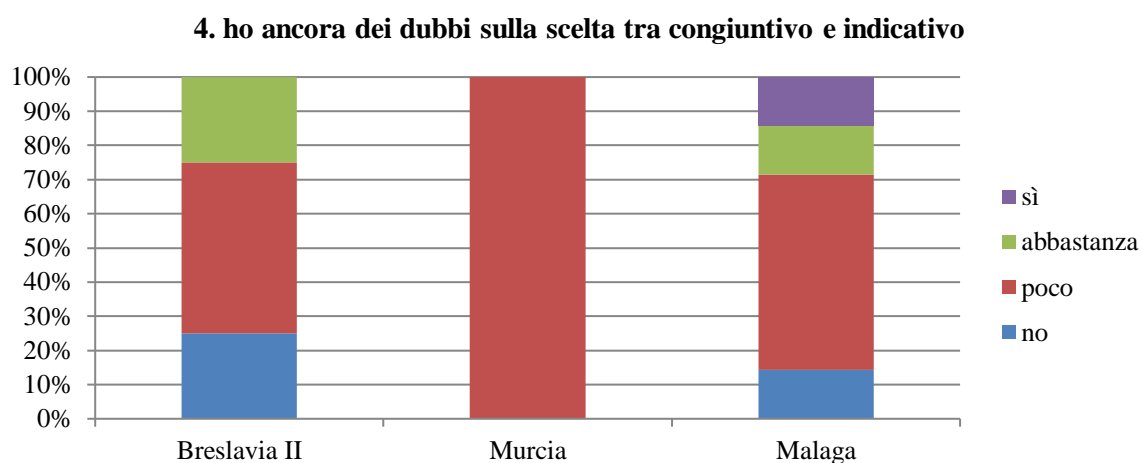
Osservando il grafico 29, emerge che, al termine della verifica, è chiaro a tutti gli apprendenti coinvolti che in alcune situazioni informali è accettabile non usare il congiuntivo. La percentuale del *sì* è pari al 100% a Malaga, all' 80% a Murcia e al 75% a Breslavia e in questi ultimi due casi le percentuali rimanenti corrispondono alla voce *abbastanza*.

Grafico 29: Percentuali delle risposte date in tutte le classi di sperimentazione al terzo quesito sulla percezione personale di acquisizione dei contenuti



Infine dal grafico 30 emerge che la maggioranza dei partecipanti in tutte le classi ha ancora dubbi nella scelta tra congiuntivo e indicativo: Breslavia: *sì* 50 % e *abbastanza* 25%; Murcia 100% *poco* e Malaga 57 % *poco*, mentre le restanti voci hanno totalizzato ciascuna il 14,3%.

Grafico 30: Percentuali delle risposte date in tutte le classi di sperimentazione al quarto quesito sulla percezione personale di acquisizione dei contenuti



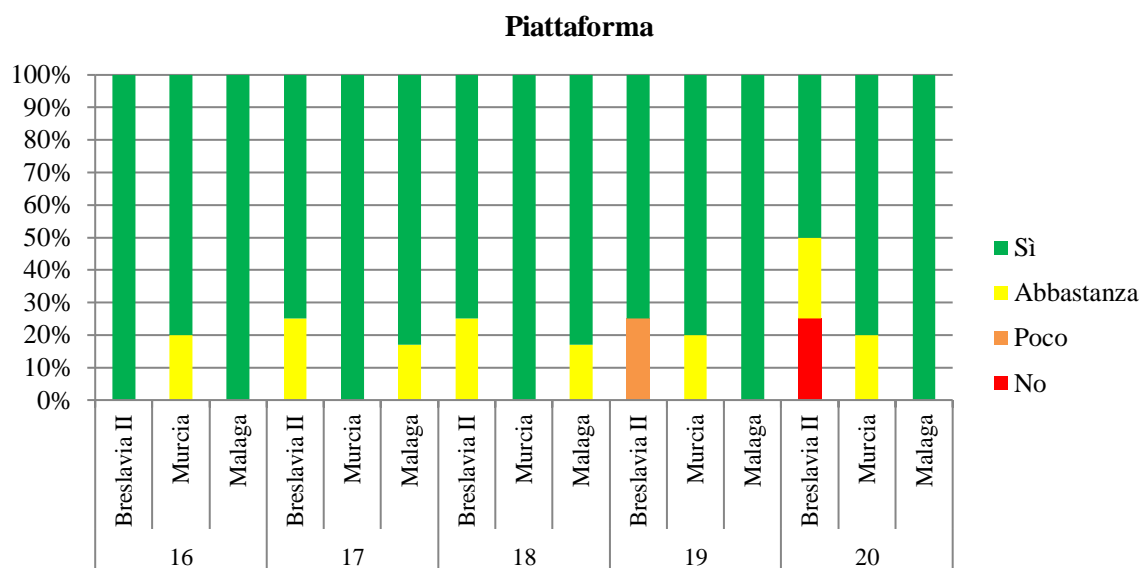
Pertanto, sono ancora presenti dei dubbi sui casi in cui la grammatica prescrive nello standard l'uso del congiuntivo, sulla scelta tra indicativo e congiuntivo e sulla distinzione tra situazioni formali e informali, mentre invece appare chiaro che in alcuni contesti informali è accettabile non usare il congiuntivo. Dunque, le attività svolte sono state generalmente apprezzate e il tema generale dell'unità è stato compreso, sebbene necessiti ancora di ulteriore pratica e riflessione, come confermano le risposte libere date nella sezione dedicata ai commenti personali, per i quali si è conservata la grafia originaria nella trascrizione: *Ho bisogno di praticare di più; Bisogno di imparare dei verbi irregolari; Tutto è bene* (Murcia); *È più facile per lo studio; Tutto ha stato bene molto facile a comprendere* (Malaga).

6.4.6. Unità didattica sul congiuntivo: opinioni sulla piattaforma Wizer.me

Segue la trattazione dei dati relativi ai punti 16-20 del questionario sulle opinioni dei partecipanti sull'uso della piattaforma: 16) Il login e l'accesso alle diverse parti dell'unità è stato facile; 17) Le istruzioni delle attività erano comprensibili; 18) Le attività si visualizzano bene sullo schermo; 19) È stato utile avere esercizi con correzione automatica; 20) Mi è piaciuta l'idea delle attività online. Dal grafico 31 emerge la generale soddisfazione degli utenti, le cui percentuali di risposta si dividono, nella quasi totalità dei casi, tra i descrittori *sì* e *abbastanza*. A conferma di ciò nella classe di Malaga un utente ha lasciato il seguente commento: *Mi piacciono le attività online perché sono più dinamiche.*

Solo nel caso del punto 19. È stato utile avere esercizi con correzione automatica e del punto 20. Mi è piaciuta l'idea delle attività online nella classe di Breslavia una percentuale ridotta, pari in entrambi i casi al 25%, è stata attribuita rispettivamente ai descrittori *poco* e *no*, tuttavia nella sezione dei commenti liberi non sono state date spiegazioni in proposito.

Grafico 31: Panoramica delle percentuali delle risposte date in tutte le classi di sperimentazione ai quesiti sull'utilizzo della piattaforma



6.5. Riflessioni finali sugli esiti delle sperimentazioni

A conclusione di questa analisi dei dati, si possono fare le seguenti riflessioni, che si baseranno sui contenuti delle tabelle 6.19 e 6.20 a seguire, che riassumono i dati già presentati nel capitolo.

Come già affermato in precedenza, nella sezione 6.3.3, dedicata al confronto degli esiti delle attività di verifica delle due unità, i punteggi medi ottenuti e le percentuali di risposte corrette erano stati posti principalmente in relazione alla durata dello studio pregresso della lingua, da cui era generalmente emerso che, ad una durata media dello studio pregresso della lingua italiana inferiore rispetto a quella delle altre classi, corrisponde anche un profitto inferiore (cfr. i dati della classe di Madrid nella sperimentazione dei verbi riflessivi “affettivi” e i dati della classe di Malaga nella sperimentazione dell’unità sul congiuntivo: in entrambi i casi le classi erano presenti alcune persone con una durata dello studio pregresso non superiore ad un anno).

Successivamente, gli esiti delle attività di verifica sono stati confrontati anche con i dati emersi dal questionario sulle opinioni dei partecipanti, prendendo in considerazione anche la conoscenza di altre lingue straniere, la presenza nella L1 e nelle altre lingue straniere conosciute di argomenti grammaticali simili e la familiarità dei partecipanti con gli argomenti trattati, sia dal punto di vista dell'uso standard che dell'uso neostandard.

Per quanto riguarda il possibile rapporto tra la L1 e le LS conosciute dai partecipanti e i risultati da loro ottenuti, in tutte le classi delle università spagnole coinvolte i partecipanti hanno dichiarato che la lingua madre e le altre lingue conosciute sono state utili a capire meglio gli argomenti grammaticali trattati, motivando la risposta sulla base delle similarità presenti tra italiano e spagnolo (sia come L1, che come LS, nel caso di studenti *Erasmus*) e in taluni casi tra anche tra italiano e francese come LS, essendo quest'ultima la lingua straniera più diffusa tra i partecipanti delle classi di Murcia. La classe di Madrid, invece, si caratterizza per la presenza al suo interno di una certa varietà di lingue romanze oltre allo spagnolo: una persona ha indicato il rumeno come L1, mentre tra le lingue straniere conosciute sono state indicate il francese, il catalano, il galiziano, il portoghese e il valenciano.

Tuttavia, se per le classi di Murcia e Malaga, dove sono state ottenute delle alte percentuali di risposte corrette nelle attività finali, si può affermare che ad una lunga durata media dello studio progresso della lingua italiana si accompagna anche la percezione, da parte dei partecipanti, del vantaggio costituito dalla L1 romanza e dalla conoscenza di altre lingue straniere romanze, nelle quali sono state trovate analogie con i temi trattati nella sperimentazione, nel caso di Madrid, invece, in assenza di una lunga durata pregressa dello studio dell'italiano, la sola presenza di varie lingue romanze all'interno della classe non ha costituito un fattore determinante e sufficiente per il raggiungimento di risultati equiparabili a quelli delle altre classi, dato che, come già affermato, nella classe è stata totalizzata la percentuale più bassa di risposte corrette tra tutte le classi di sperimentazione.

Tale riflessione sembra confermarsi osservando i dati delle classi di Breslavia, dove si è verificata una situazione analoga a quella delle classi di Murcia e Malaga, con il raggiungimento di percentuali elevate di risposte corrette. A prima vista il profilo dei partecipanti polacchi alla sperimentazione presenta alcune differenze rispetto a quello dei partecipanti spagnoli, ossia la L1 slava e la presenza molto ristretta di parlanti di lingue romanze come LS diverse dall'italiano: in entrambe le classi di Breslavia la lingua straniera più frequente tra i partecipanti è l'inglese, segue una ristretta minoranza che ha indicato il francese o lo spagnolo e altre lingue non romanze. Tutti hanno affermato nelle

risposte del questionario che le lingue straniere conosciute non sono state utili alla comprensione, incluse le persone che hanno una lingua romanza nel proprio curriculum, poiché la loro conoscenza di quelle lingue era troppo limitata per poter fornire un supporto effettivo alla comprensione.

Tuttavia, nonostante queste differenze nel profilo linguistico, un fattore che ha contribuito al raggiungimento di alte percentuali di risposte corrette nel caso dell'unità dedicata ai verbi riflessivi "affettivi" è probabilmente la presenza in polacco di una struttura grammaticale simile, utilizzata nel registro informale, come dichiarato dai partecipanti nelle risposte al questionario.

Sulla base di quanto affermato, emerge che le classi di sperimentazione che hanno svolto con migliore profitto le attività conclusive sull'uso neostandard dei due tratti grammaticali sono quelle che presentano, in generale, i seguenti fattori: innanzitutto una più lunga durata pregressa dello studio della lingua italiana, a cui si aggiunge la presenza nella L1 di un fenomeno grammaticale simile a quello affrontato nell'unità. Quest'ultimo fattore si è verificato sia nel polacco che nelle varie lingue romanze dei partecipanti coinvolti: tuttavia, come mostrato dai dati precedentemente presentati, specie nel caso della classe di Madrid, la sola analogia tra l'italiano e la L1 e/o le diverse lingue romanze conosciute non ha costituito un fattore determinante per lo svolgimento positivo delle attività.

Oltre a ciò, la maggior parte dei partecipanti delle classi di sperimentazione, ad eccezione di pochi, ha dichiarato di aver già incontrato nella vita quotidiana alcuni esempi dell'uso neostandard degli argomenti trattati nelle unità, senza, tuttavia, averne alcuna cognizione teorica che ne spiegasse il funzionamento. Se ne deduce che l'uso standard di entrambi i temi grammaticali trattati fosse un contenuto già presente nella competenza linguistica degli apprendenti e che l'impatto delle unità si sia concretizzato non solo nell'aver aggiunto ulteriori informazioni circa l'uso standard, ma soprattutto nell'aver fornito un inquadramento teorico riguardo all'uso neostandard, che sino ad allora era percepito dagli studenti come errato o inspiegabile, in quanto fuori dalla norma linguistica che era stata loro trasmessa.

Attraverso l'uso delle unità i partecipanti hanno dimostrato di saper riutilizzare le strutture grammaticali su cui erano incentrate le unità e di saper scegliere di utilizzare o meno una determinata struttura linguistica in base allo specifico contesto comunicativo e al registro formale o informale ad esso correlato.

Infine, attraverso il questionario hanno comunicato il proprio apprezzamento per l'operato della docente-ricercatrice e hanno fornito anche un riscontro sulle attività svolte, dando dei suggerimenti per il miglioramento dello svolgimento della lezione, della fruizione della piattaforma dal punto di vista della grafica e della struttura delle attività.

Tabella 6.19: Dati relativi alle classi di sperimentazione dell'unità sui verbi riflessivi "affettivi"

	Breslavia II	Breslavia I	Madrid	Murcia	Malaga
Media mesi (studio pregresso)	30	27,6	14,8	28	28,8
Deviazione standard	6,9	37,5	7,4	6,1	24,8
% risposte corrette verifica finale	88,3%	83,5%	59,6%	74,4%	81,3%
L1 e LS	L1: polacco; LS: inglese (4), francese (3), hindi (1) spagnolo (1), tedesco (1)	L1: polacco (tutti); polacco e italiano (1) LS: inglese (13), francese (3), russo (3), tedesco (3), spagnolo (2), hindi (1) svedese (1)	L1: spagnolo (18), rumeno (1), finlandese (1) tedesco (1). LS: inglese (17), tedesco (9), francese (6), spagnolo LS (3), catalano (2), galiziano (2), giapponese (1), olandese (1), portoghese (1), russo (1), svedese (1), valenciano (1)	L1: spagnolo; LS: francese (6), inglese (4)	L1: spagnolo (3), francese (2); LS: inglese (5), francese LS (2), spagnolo LS (2), portoghese (1)
Utilità della L1/LS nella comprensione del tema?	Sì (2) la L1, perché ha un argomento grammaticale simile; un po' (1); no (1)	Sì (12) la L1, ha un argomento grammaticale simile; no (2)	Sì (tutti): similarità presenti tra l'italiano e lo spagnolo.	Sì (tutti): similarità presenti tra l'italiano e lo spagnolo.	La L1: sì, (tutti); similarità presenti tra l'italiano e il francese (1)
Hai già incontrato esempi dell'uso neostandard?	Sì (tutti)	Sì (11) No (3)	Sì (13) No (4)	Sì (3); 2 (no); 1 (non so)	Sì (tutti)

Tabella 6.20: Dati relativi alle classi di sperimentazione dell'unità sul congiuntivo vs indicativo

	Breslavia II	Murcia	Malaga
Media mesi (studio pregresso)	30	38,4	20
Deviazione standard	6,9	10	6,1
% risposte corrette verifica finale	75%	79%	67,1%
L1 e LS	L1: polacco; LS: inglese (4), francese (3), hindi (1) spagnolo (1), tedesco (1)	L1: spagnolo; LS: francese (5), inglese (3), portoghese (1)	L1: spagnola (5) francese (1); LS: inglese (6), francese (3), spagnolo (1).
Utilità della L1/LS nella comprensione del tema?	Tutti (no): né la L1, né le LS, di cui alcuni dichiarano di avere un livello troppo basso	La L1: sì, (tutti); LS: similarità presenti tra italiano, spagnolo e francese (1)	La L1: sì, (5), per le similarità presenti tra italiano, spagnolo; no (1)
Hai già incontrato esempi dell'uso neostandard?	Sì (3), No (1)	Sì (2), No (3)	Sì (3), No (3)

CONCLUSIONI

Il tema di questo lavoro si intreccia strettamente con la questione della ricerca di uno standard linguistico, una questione ricorrente nella storia dell'italiano e che perdura almeno dall'Unità d'Italia. Come esposto nei capitoli 1 e 2, la ristrutturazione del sistema linguistico italiano ha fatto emergere a partire dalla seconda metà del Novecento una nuova varietà che “si candida ad occupare, dopo secoli di ostracismo, il baricentro dell'intero sistema linguistico italiano o per lo meno a condividere con lo standard ufficiale il crisma della norma” (Sabatini 1985: 175). Alla descrizione dei tratti distintivi di tale varietà linguistica si affianca anche la necessità di definire quali di questi fenomeni grammaticali siano da considerarsi effettivamente costitutivi di un nuovo standard linguistico e, di conseguenza, quali di essi vadano trasmessi nell'insegnamento dell'italiano a stranieri.

Questa ricerca si inserisce nell'ambito di questa riflessione e, sulla base dei principali studi condotti sul neostandard e citati nel corso della trattazione, propone un possibile approccio per l'insegnamento dell'italiano, che consiste nel presentare i fenomeni grammaticali del neostandard come segnali di un cambiamento tendenziale dell'italiano (in accordo con quanto proposto da Sobrero 1992), avendo cura però di fare una distinzione basata sulla funzionalità.

Per alcuni fenomeni il distacco dalla norma grammaticale tradizionale produce un effettivo apporto funzionale alla lingua, poiché senza di essi non sarebbe possibile esprimere un certo concetto o raggiungere un determinato scopo pragmatico; altri, invece, abbassano il registro linguistico senza apportare di fatto alcun vantaggio sul piano funzionale, poiché sono solo un'alternativa arbitraria a forme standard perfettamente funzionanti (in accordo con quanto proposto da Prandi 2020). I fenomeni che rientrano nella prima categoria andrebbero trasmessi nell'insegnamento come parte integrante della nuova varietà standard; tutti gli altri possono essere menzionati come varianti sub-standard, seppur diffuse (cfr. capitolo 2).

Questa proposta non nega l'importanza dell'elaborazione di concetti come “italiano neostandard” (cfr. Berruto 1987) o “italiano dell'uso medio” (cfr. Sabatini 1985), ossia la revisione dello standard ereditato dai secoli passati, un'esigenza esistente da secoli e più volte affiorata nel corso della storia. Tuttavia, se da un lato le descrizioni di Francesco Sabatini e di Gaetano Berruto hanno avuto il pregio di documentare anche dal punto di

vista storico i cambiamenti avvenuti nell'italiano, dall'altro includere nelle liste tutto ciò che si distaccava dalle norme grammaticali ha portato di fatto all'accostamento di risorse funzionali essenziali a forme che invece non apportano alcun vantaggio funzionale e che producono solo una caduta di registro. Ne deriva che persistono ancora dei dubbi dal punto di vista della messa a punto di uno standard dell'uso comune che dia sicurezza linguistica ai parlanti e un modello definito a chi ne valuta gli usi.

Rispetto alle pubblicazioni già citate nei capitoli precedenti sul tema del neostandard, che principalmente si focalizzano solo sulla frequenza d'uso o sulla percezione di correttezza da parte dei parlanti, questa ricerca offre dunque un criterio la cui applicazione permette di non cadere né nell'errore di rifiutare in blocco tutto ciò che differisce dalla norma prescrittiva, né nell'errore opposto, ossia di ritenere che tutti gli errori di oggi saranno le regole di domani in base ai soli dati sulla frequenza o sulla preferenza dei parlanti.

Una volta chiarita la questione dello standard e dell'insegnabilità, è necessario reperire delle fonti da cui ricavare dei testi che contengano i tratti grammaticali oggetto di studio. Pertanto, nel capitolo 3 sono state illustrate le ragioni per le quali i corpora possono rappresentare delle valide risorse per l'apprendimento della lingua e, nel caso specifico, per il reperimento di testi autentici. Sono stati presi in esame i vari ambiti di applicazione dei corpora alla didattica delle lingue, dedicando particolare attenzione all'uso dei corpora per la realizzazione di materiali didattici (cfr. le pubblicazioni contenute in Sinclair 2004, Granger, Meunier 2008, Meunier, Granger 2008, Aijmer 2009, Hinkel 2011, gli atti dei convegni *Teaching and Language Corpora* (TaLC) e la bibliografia ivi contenuta).

Contrariamente alla maggior parte degli studi prodotti in questo ambito, in cui prevale l'uso diretto dei corpora da parte degli apprendenti per verificare delle ipotesi sul funzionamento della lingua, questo studio si inserisce nel filone del loro uso diretto da parte dei docenti per attingere lessico o fraseologia da riutilizzare in aula (in particolare per l'italiano L2/LS si è fatto riferimento a Corino, Marengo 2009, Merlo 2009, Kennedy, Miceli 2010, Guidetti, Lenzi, Storchi 2012, Lyding *et al.* 2013 e Ivanovska-Naskova 2018). Da questo punto di vista l'innovatività di questa ricerca risiede nel fatto di aver percorso un terreno ancora poco battuto, ossia l'utilizzo di testi tratti dai corpora per la costruzione di intere unità didattiche mirate all'acquisizione di determinate strutture grammaticali.

Dopo aver illustrato i criteri di riferimento utilizzati per l'inquadramento teorico di linguistica italiana e l'applicazione dei corpora, l'esposizione è stata articolata secondo i

seguenti ambiti: criteri di scelta del corpus più idoneo agli scopi della ricerca (capitolo 4) e i risultati della sua interrogazione (capitolo 5); aspetti innovativi nella progettazione delle unità didattiche (capitoli 4 e 6), e infine l'analisi incrociata degli esiti delle sperimentazioni delle UD e del questionario (capitolo 6).

Innanzitutto è necessario precisare che l'identificazione del corpus PAISÀ come la risorsa più idonea per gli scopi della ricerca è il frutto di un compromesso tra i parametri fissati per la valutazione dei corpora dell'italiano scritto attualmente disponibili. Sebbene il corpus PAISÀ non emerga rispetto alle altre risorse prese in esame né per le dimensioni (250 milioni di parole) né in particolare per quanto riguarda l'attualità dei materiali, che si fermano al 2010, si è ritenuto che lo scarto di dieci anni dalla conclusione della costruzione del corpus al momento dell'elaborazione di questo studio non compromettesse la ricerca di fenomeni sintattici che, come più volte ricordato, sono di fatto antichi e che hanno iniziato a emergere sempre più nella lingua scritta a partire dalla seconda metà del Novecento. Inoltre, si è ritenuto che i limiti sopra menzionati fossero controbilanciati dai seguenti aspetti: la presenza di licenze *Creative Commons*, che consentono il reimpiego dei testi del corpus per vari scopi, tra cui la progettazione delle attività didattiche e, in secondo luogo, il fatto che il corpus fosse stato progettato affinché potesse fungere da risorsa di supporto all'apprendimento dell'italiano.

Questa valutazione è stata fatta tenendo conto degli obiettivi della ricerca: documentare empiricamente la diffusione dei tratti linguistici elencati da Sabatini (1990) attraverso il reperimento di testi scritti nel corso degli ultimi 20-30 anni al massimo (coerentemente alla pubblicazione degli studi sull'italiano dell'uso medio/neostandard), che fossero liberamente riutilizzabili per produrre unità didattiche senza incorrere nella violazione del *copyright*. Di fatto PAISÀ è risultata essere l'unica risorsa che soddisfa tutte queste esigenze, seppur parzialmente, come sopra descritto.

L'interrogazione del corpus è stata resa possibile tramite la progettazione di alcune stringhe di ricerca in linguaggio CQL (Evert 2009, Nesi, Thompson 2014). Prendendo in esame 13 dei 14 tratti linguistici elencati da Sabatini nel 1990 (cfr. sezione 4.7.2), sono state elaborate 57 stringhe di ricerca, il cui numero è nettamente superiore ai tratti elencati da Sabatini. Come è stato illustrato nei capitoli 4 e 5, l'elaborazione di più stringhe si è resa necessaria principalmente in due casi: per i tratti linguistici articolati in sottofenomeni, che chiaramente vanno cercati separatamente, e nel caso in cui le specifiche caratteristiche sintattiche o morfologiche di un determinato tratto hanno inevitabilmente richiesto la progettazione di più stringhe di ricerca (si veda ad esempio il caso dei verbi in

forma pronominale per indicare partecipazione affettiva, per i quali un'unica stringa di ricerca non avrebbe potuto coprire tutte le forme pronominali usate per l'intera coniugazione). Quindi, il numero di stringhe finali in alcuni casi è stato correlato linguisticamente alla presenza di sottofenomeni, in altri casi dall'oggettiva impossibilità di coprire con un'unica stringa di ricerca tutte le occorrenze di un fenomeno, a causa della sua complessità intrinseca.

Sulla funzionalità delle stringhe è doveroso fare alcune precisazioni. È necessario infatti ricordare che quando si utilizzano delle stringhe che si basano sull'annotazione sintattica di ricerca, le occorrenze ottenute per ciascun fenomeno e di conseguenza i valori della frequenza e della deviazione standard dipendono notevolmente dalla formulazione delle stringhe di ricerca. Quest'ultima, a sua volta, è influenzata o dai limiti imposti dalle opzioni attualmente a disposizione nell'uso del linguaggio CQL, che di fatto rendono estremamente complesso il reperimento di determinati fenomeni grammaticali, o dalla necessità di ricercare un equilibrio tra la formulazione di una stringa che garantisca una copertura ad ampio spettro di un fenomeno grammaticale e il rischio di reperire anche grandi quantità di casi che, pur presentando delle analogie strutturali con il fenomeno indagato, in realtà non siano inerenti (cfr. capitolo 4).

Un'ulteriore riflessione meritano le procedure adottate per il calcolo della frequenza all'interno del corpus dei fenomeni grammaticali oggetto di studio, al fine di determinarne i più ricorrenti. Questo aspetto è di particolare importanza per la presente ricerca, poiché, come è stato descritto, la frequenza dei fenomeni grammaticali nel corpus è stato uno dei criteri adottati per la selezione dei fenomeni grammaticali sui quali costruire dei materiali didattici per l'italiano per stranieri. Richiamando brevemente quanto esposto nel capitolo 5, per ciascuno dei tre campioni da 200 occorrenze scaricati dal corpus è stato innanzitutto calcolato il rapporto tra il numero di occorrenze salvate dopo l'eliminazione di eventuale rumore (corrispondente a casi non inerenti e dopponi) e le 200 occorrenze iniziali del campione. Successivamente è stata calcolata la media dei tre rapporti ottenuti, che abbiamo chiamato "media campione". Tale media campione può essere concepita come il tasso di occorrenze effettivamente corrispondenti al fenomeno grammaticale indagato rispetto al "rumore".

Poi si è proceduto al calcolo della deviazione standard campione, che fornisce una misura di come siano distribuiti i dati nel campione rispetto alla media, ossia permette di sapere se la media costituisce una rappresentazione affidabile dei dati. Si è scelto di

calcolare la deviazione standard campione e non l'errore standard della media per pura semplicità di calcolo.

Al fine di rendere paragonabili i dati relativi alla frequenza e alla deviazione standard di fenomeni grammaticali aventi dati di partenza molto eterogenei, sono state, inoltre, calcolate la frequenza stimata, che è data dal prodotto tra il totale delle occorrenze inizialmente ottenute per il fenomeno indagato tramite l'interrogazione del corpus PAISÀ e la media campione, e la frequenza stimata (o normalizzata) per 1 milione di parole. Quest'ultima è ottenuta moltiplicando la frequenza stimata per 1.000.000 e dividendo poi il risultato per il numero totale di token del corpus PAISÀ.

Ciò è stato svolto nella consapevolezza che né il calcolo della media campione né la stima delle frequenze normalizzate rappresentino procedure standard. Normalmente per quantificare i casi inerenti al fenomeno ricercato rispetto al cosiddetto "rumore" le stime di frequenza si basano su un campione unico. Tuttavia, si è ritenuto che la scelta di avere tre campioni su cui calcolare dapprima la frequenza stimata (al netto del rumore) e, successivamente, la normalizzazione per milione di parole consentisse di disporre di una base di dati più eterogenea e, pertanto, di poter formulare delle affermazioni sulla frequenza dei fenomeni basate – si auspica – su una maggiore fondatezza. Da questo punto di vista si ritiene che tale scelta, che differisce da quanto viene normalmente svolto nelle ricerche in questo ambito, rappresenti un aspetto metodologicamente innovativo di questo lavoro.

Per completezza si aggiunge che si è scelto di applicare la procedura di normalizzazione anche ai valori della deviazione standard campione, in modo che fosse in proporzione alla frequenza, calcolando la deviazione standard campione stimata, data dal prodotto tra il totale delle occorrenze inizialmente ottenute da PAISÀ e la deviazione standard campione, e la deviazione standard stimata per 1 milione di parole, data dal prodotto tra deviazione standard campione stimata e 1.000.000, diviso per il numero totale di token di PAISÀ, ossia: $\text{deviazione standard campione stimata} * 1.000.000 / 265.062.636$. Anche in questo caso, si ha la consapevolezza che la procedura non è standard; tuttavia si è resa necessaria per riportare una misura di dispersione da accompagnare alle stime di frequenze normalizzate.

La quantità di occorrenze reperite nel corpus per i fenomeni oggetto di studio non ha sempre consentito di realizzare per ciascun fenomeno, come stabilito, un macrocampione da 600 occorrenze da suddividere in 3 campioni da 200 occorrenze ciascuno. Di conseguenza, i casi illustrati nel capitolo 5 presentano delle differenze rispetto ai seguenti

fattori: il numero di campioni ricavati da ciascuna stringa (3 campioni vs un campione unico), il totale iniziale di ciascun campione (200 occorrenze o meno), la differenza tra il totale iniziale di occorrenze scaricate per ciascun campione e il totale delle occorrenze salvate al termine della ripulitura del campione dai casi non inerenti (ad esempio, 176 occorrenze su 200), eccetera. Per tali ragioni, la procedura di normalizzazione per 1 milione di parole è stata necessaria per poter confrontare i dati relativi alla frequenza e alla deviazione standard di fenomeni grammaticali altrimenti difficilmente confrontabili, vista l'eterogeneità appena descritta dei dati di partenza di ciascun campione.

Tramite le procedure sopra citate è stato possibile classificare i fenomeni grammaticali in tre raggruppamenti, rispetto alla loro frequenza bassa, media o alta nel corpus. Ciò ha permesso di comprendere quali tra i fenomeni grammaticali elencati da Sabatini (1990) siano maggiormente utilizzati nella produzione scritta di media formalità o informalità dagli italofoeni nativi e, di conseguenza, abbiano una maggiore probabilità di essere trovati nei testi scritti da parte degli apprendenti di lingua italiana L2/LS.

Una volta individuati i fenomeni grammaticali più ricorrenti nel corpus, questi sono stati sottoposti a un secondo criterio di selezione, ossia sono stati confrontati con i contenuti grammaticali della manualistica per stranieri attualmente in commercio (cfr. Słapek 2017). In questo modo sono stati individuati i fenomeni finora esclusi dalle pubblicazioni esistenti e per i quali fosse quindi rilevante e utile creare dei materiali didattici. Entrando nel dettaglio, la rilevanza e l'utilità sono state valutate non solo rispetto ai contenuti delle pubblicazioni esistenti, ma anche rispetto alle diverse possibilità di impiego dei materiali didattici da progettare.

Al fine di realizzare unità didattiche utilizzabili in varie fasi del percorso di apprendimento, sia come primo approccio con il tema, sia negli approfondimenti dei livelli successivi al progredire della competenza linguistica, si è fatta strada l'idea di creare due unità didattiche incentrate su argomenti grammaticali collocati a cavallo tra più livelli del Quadro Comune Europeo di Riferimento per la conoscenza delle lingue (QCER) (Consiglio d'Europa 2018, 2002), che potessero rappresentare un oggetto di interesse – nonché una fonte di dubbi sull'uso della lingua – su più livelli.

Per tale ragione, l'attenzione è stata rivolta all'uso dell'indicativo al posto del congiuntivo (che include sottofenomeni a frequenza alta, media e bassa nel corpus) e all'uso del verbo in forma pronominale riflessiva in senso affettivo (es. *io mi mangio un panino*; fenomeno a media frequenza nel corpus), ritenendo che la produzione di materiali

didattici su questi temi avesse una maggiore rilevanza glottodidattica rispetto alla produzione relativa ad un singolo argomento specifico collocato in un solo livello.

Dal punto di vista della progettazione delle unità la presente ricerca presenta due aspetti innovativi nell'ambito della glottodidattica e dell'applicazione dei corpora alla didattica delle lingue. Il primo è l'intento, già menzionato, di compensare l'attuale assenza di materiali didattici incentrati sul neostandard per i livelli inferiori al C1-C2 (cfr. capitoli 4 e 6) tramite dei materiali didattici inediti, utilizzabili sia come primo approccio, sia come attività di rinforzo e recupero di temi grammaticali già trattati, scegliendo quali attività svolgere dalla più semplice alla più complessa. Il secondo aspetto innovativo consiste nel fatto che la presente ricerca si distacca dalla frequente tendenza all'uso diretto dei corpora da parte degli apprendenti e si inserisce in un ambito poco praticato dell'uso diretto dei corpora da parte dei docenti, ossia l'utilizzo da parte del docente dei testi estratti dal corpus per la costruzione di materiali didattici, nello specifico intere unità didattiche mirate all'acquisizione di determinate strutture grammaticali (cfr. capitolo 3).

Dopo aver messo in luce gli aspetti innovativi della progettazione dei materiali, prima di passare alla riflessione finale sugli esiti delle attività di verifica delle unità, verranno qui richiamate alcune precisazioni fatte nel corso di questa ricerca circa i limiti relativi allo svolgimento di tali sperimentazioni.

L'obiettivo della sperimentazione è valutare se e in che misura i partecipanti al termine delle attività siano in grado di formulare delle scelte tra uso standard e neostandard rispetto al grado di formalità/informalità del contesto e di riconoscere l'uso neostandard da parte di italofoeni nativi come il frutto di una scelta e non di una scarsa padronanza delle regole grammaticali.

La prima riflessione da fare riguarda il numero di partecipanti coinvolti. Avendo a disposizione un collettivo estremamente ridotto, lo scopo delle sperimentazioni delle unità didattiche e della raccolta delle opinioni dei partecipanti è di svolgere un primo test dei materiali didattici e valutarne la percezione di efficacia da parte dei partecipanti in termini di utilità e contributo motivazionale nell'acquisizione linguistica, nonché i limiti e gli aspetti da migliorare. Pertanto, non è possibile ricavare da questa ricerca conclusioni di carattere generale, valide su vasta scala, per le quali sarebbe stato necessario il coinvolgimento di un collettivo più ampio.

Proseguendo le riflessioni sull'ampiezza del collettivo, si precisa che il totale dei partecipanti alla sperimentazione in taluni casi ha subito delle variazioni in base alle diverse attività svolte. I fattori che hanno determinato la mancata consegna di alcune

attività sono principalmente di natura logistica (ad esempio, l'arrivo in ritardo o, viceversa, l'uscita anticipata dall'aula di alcune persone rispetto all'orario della lezione), mentre solo in pochi casi isolati sono state riscontrate delle difficoltà nel salvataggio delle risposte sulla piattaforma, causati dall'instabilità della connessione internet o da incomprensioni circa la modalità di salvataggio.

Di seguito verranno, invece, esposti i limiti di carattere metodologico. Quando si progetta una sperimentazione, è necessario che l'ambiente sperimentale venga tenuto in condizioni costanti, ad eccezione della variabile sperimentale, che invece viene volutamente cambiata. Una sperimentazione controllata prevede, infatti, che i partecipanti vengano divisi in due gruppi: un gruppo sperimentale sottoposto ai cambiamenti della variabile e un gruppo di controllo che non sia esposto a quei cambiamenti, la cui presenza garantisce che i dati prodotti dal gruppo sperimentale siano realmente dovuti alla variabile che si sta verificando e non a influenze esterne sconosciute.

Nella presente ricerca l'ambiente sperimentale consiste nelle lezioni di lingua italiana di livello B1, aventi tra gli obiettivi linguistici un ripasso dell'uso standard di due tratti grammaticali generalmente affrontati nel livello intermedio, all'interno del quale si inserisce, come elemento di novità, un approfondimento dell'uso neostandard di tali tratti grammaticali.

Nel caso in questione, la prassi per ottenere dei dati sull'efficacia delle attività svolte consisterebbe nel dividere la popolazione in due gruppi: un gruppo che svolge tutte le attività della sperimentazione, incluso l'approfondimento sulla scelta tra uso standard e neostandard rispetto al grado di formalità/informalità del contesto, e un gruppo di controllo a cui non venga fornito tale approfondimento. Al termine del percorso, entrambi i gruppi verrebbero sottoposti alle stesse attività di verifica, al fine di valutare se vi siano delle differenze tra i risultati ottenuti da chi è stato sottoposto alla variabile sperimentale e chi non lo è stato.

Pur nella piena consapevolezza della necessità di tali procedure per ottenere risultati significativi, non è stato di fatto possibile dividere le classi ed effettuare incontri differenziati con il gruppo di sperimentazione e il gruppo di controllo. Ciò è dovuto al fatto che negli accordi siglati con gli atenei esteri coinvolti nella sperimentazione è stato concesso il numero di ore strettamente necessario allo svolgimento delle due unità didattiche, ossia 1 ora e 30 per ciascuna unità, fino ad un massimo di 3 ore totali. Di conseguenza, le classi sono state coinvolte nella loro interezza ed è stato necessario fornire a tutti i partecipanti gli stessi contenuti. Se ne deduce che i dati raccolti non forniscono

indicazioni sull'efficacia delle attività di verifica finale in assenza dello svolgimento dell'intera unità didattica. Infine, non essendo stati svolti test di significatività nelle sperimentazioni delle unità didattiche, soltanto studi futuri, basati sul coinvolgimento di collettivi più grandi ed eterogenei potranno verificare la validità di quanto ottenuto.

Fatte queste precisazioni sui limiti dello svolgimento delle sperimentazioni, si procederà ora all'analisi incrociata degli esiti delle sperimentazioni delle unità didattiche, svolte presso i dipartimenti di italianistica di alcuni atenei europei, e degli esiti del questionario per la raccolta delle opinioni degli studenti sulle attività svolte (capitolo 6), mettendo in luce quanto fattori come la durata dello studio pregresso della lingua italiana, la conoscenza di una o più lingue romanze e la presenza di tratti grammaticali simili nella L1 o nelle altre LS studiate possano aver avuto un impatto sugli esiti delle sperimentazioni condotte in aula.

Gli esiti delle attività di verifica di ciascuna unità (ossia, i punteggi medi ottenuti e le percentuali di risposte corrette) sono stati confrontati con i dati emersi dal questionario sulle opinioni dei partecipanti, prendendo in considerazione i seguenti fattori: la lingua materna degli apprendenti (L1), la conoscenza di altre lingue straniere (LS), la durata dello studio pregresso dell'italiano, la familiarità dei partecipanti con gli argomenti trattati, sia dal punto di vista dell'uso standard che dell'uso neostandard e infine la presenza nella L1 e nelle altre LS conosciute di argomenti grammaticali simili.

Dall'analisi incrociata di questi dati è emerso che le classi di sperimentazione che hanno svolto con migliore profitto le attività conclusive delle due unità sull'uso neostandard sono, in generale, quelle che presentano una più lunga durata pregressa dello studio della lingua italiana, a cui si aggiunge la presenza nella L1 o nelle LS conosciute di un fenomeno grammaticale simile a quello affrontato nell'unità. Quest'ultimo fattore si è verificato nel caso dell'unità sui riflessivi, sia per il polacco che per le varie L1 romanze dei partecipanti coinvolti (principalmente spagnolo e francese), consentendo di ipotizzare un possibile rapporto tra la L1 e/o le LS conosciute dai partecipanti e i risultati da loro ottenuti.

Sebbene dai dati raccolti emerga che i profitti migliori si accompagnino sempre ad una lunga durata dello studio pregresso della lingua italiana, anche nel caso in cui il fenomeno grammaticale oggetto di studio non ha un corrispettivo effettivo nella L1 (ad esempio, il congiuntivo non esiste in polacco, ma la verifica finale dedicata a questo argomento è stata svolta con successo dalla classe del secondo anno di Breslavia), non è possibile affermare il contrario: in assenza di una lunga durata pregressa dello studio dell'italiano la sola

analogia tra l'italiano e la L1 e/o le diverse lingue romanze conosciute sembra non costituire un fattore determinante per lo svolgimento positivo delle attività. Ciò è emerso in particolare nella classe di Madrid, dove a fronte della più breve durata dello studio pregresso e della più ampia varietà di L1/LS romanze in aula (L1: spagnolo e rumeno; LS: francese, catalano, galiziano, portoghese e valenzano) è stata totalizzata la percentuale più bassa di risposte corrette tra tutte le classi di sperimentazione.

Ciò permette di ipotizzare un possibile rapporto tra la durata pregressa dello studio, la L1 e le LS conosciute dai partecipanti e i risultati positivi da loro ottenuti, ove il primo fattore sembra avere un peso maggiore sulla riuscita delle attività, anche in assenza di analogie tra la L1 e/o le diverse lingue romanze conosciute.

Per quanto riguarda le conoscenze pregresse specifiche sul tema, la maggior parte dei partecipanti delle classi di sperimentazione ha dichiarato di aver già incontrato nella vita quotidiana alcuni esempi dell'uso neostandard degli argomenti trattati nelle unità, senza tuttavia avere alcuna cognizione teorica che ne spiegasse il funzionamento.

Pertanto, rispetto a quest'ultimo aspetto citato, il contributo delle unità didattiche progettate si è concretizzato nell'aver fornito un inquadramento teorico non solo circa l'uso standard di alcuni fenomeni grammaticali – in taluni casi come primo approccio, in altri come integrazione e potenziamento di un argomento già parzialmente noto – ma anche riguardo al loro uso neostandard, che sino ad allora era percepito dai partecipanti come errato o inspiegabile, essendo in contrasto con la norma linguistica che era stata loro trasmessa.

Attraverso l'uso delle unità i partecipanti hanno dimostrato di saper riutilizzare le strutture grammaticali su cui erano incentrate le unità e di saper scegliere di utilizzare o meno una determinata struttura linguistica in base allo specifico contesto comunicativo e al registro formale o informale ad esso correlato. I materiali progettati possono, dunque, fungere da supporto per la formazione degli apprendenti di lingua italiana L2/LS, favorendo una percezione realistica delle stratificazioni sociolinguistiche della lingua usata dai parlanti.

Quanto esposto finora permette di affermare che studi futuri, basati sul coinvolgimento di collettivi più eterogenei e più grandi, che consentano la creazione del gruppo sperimentale e del gruppo di controllo, potranno verificare la validità di quanto è stato ottenuto in questo studio. Inoltre gli esiti delle sperimentazioni, il riscontro sulle attività svolte fornito dai partecipanti tramite il questionario, e in particolare i suggerimenti per lo svolgimento della lezione e per una migliore fruizione della piattaforma dal punto di vista

della grafica e della struttura delle attività, costituiscono una base importante per modificare le unità didattiche prodotte e dare vita in futuro ad uno studio più ampio.

Riferimenti bibliografici

- Acquaviva P. (2000), *La grammatica italiana: il lavoro comincia adesso*, «Lingua e stile», 2, pp. 249–272.
- Aijmer K. (2009, a c. di), *Corpora and Language Teaching*, John Benjamins, Amsterdam/Philadelphia.
- AIL: Accademia Italiana di Lingua (2015), *AIL presentazione ed introduzione. 6 esami di italiano, da A1 a C2, & 2 esami di italiano commerciale per tutti i livelli di conoscenza*, Firenze. Disponibile alla pagina:
<http://www.acad.it/downloads/sillabo.html>. (consultato il 10/01/2019).
- Alhujaylan H. (2014), *Teaching collocations and lexical phrases: A data driven learning approach*, contributo presentato alla 11a conferenza internazionale su *Teaching and Language Corpora* (TaLC11), Lancaster University, 20-23 luglio 2014. Disponibile online: <http://ucrel.lancs.ac.uk/talc2014/doc/TALC2014-abstract-book.pdf> (consultato il 20/12/2019).
- Alisova T. (1972), *Strutture semantiche e sintattiche della proposizione semplice in italiano*, Sansoni, Firenze.
- Ammon U. (1986), *Explication der Begriffe 'Standardvarietät' und 'Standardsprache' auf normtheoretischer Grundlage*, in *Sprachlicher Substandard. Vol. I*, a c. di G. Holtus, E. Radtke (1986-1990), pp. 1-62.
- Ammon U. (2003), *On the social forces that determine what is standard in a language and on conditions of successful implementation*, in «Sociolinguistica», 17, pp. 1-10.
- Amor S. (2002), *Authenticity and Authentication in Language Learning: Distinctions, Orientations, Implications*, Peter Lang, Frankfurt.
- Andorno C. (2003), *La grammatica italiana*, Mondadori, Milano.
- Andorno C., Bernini G., Giacalone Ramat A., Valentini A. (2003), *Sintassi*, in *Verso l'italiano*, a c. di A. Giacalone Ramat, Carocci, Roma, pp. 116-127.
- Antonelli G. (2007), *L'italiano nella società della comunicazione*, Il Mulino, Bologna.
- Antonelli G. (2011), *Lingua*, in *Modernità italiana. Cultura, lingua e letteratura dagli anni Settanta a oggi*, a c. di A. Acribo, E. Zinato, Carocci, Roma, pp. 15–52.
- Arcangeli M., Giugni S., Cardillo G., D'Angelo K., Ferrari S., Greco N., Lucatorno A. (2014), *Attestato ADA. Piano dei corsi*, Alma & Società Dante Alighieri, Firenze.
- Asor Rosa A., Serianni L., Trifone P. (1994, a c. di), *Storia della lingua italiana. Vol. I: I luoghi della codificazione; vol. II: Scritto e parlato; vol. III: Le altre lingue*, Einaudi, Torino.
- Aston G., Bernardini S., Stewart D. (2004, a c. di), *Corpora and language learners*, John Benjamins, Amsterdam/Philadelphia.
- Babych S. (2012), *Exploring Patterns of Textual Cohesion in Multilingual Corpora and their Application for Teaching Language and Translation*, contributo presentato alla 10a conferenza internazionale su *Teaching and Language Corpora* (TaLC10), Uniwersytet Warszawski, 11-14 luglio 2012. Disponibile online: <http://talc10.ils.uw.edu.pl/files/2011/09/BABYCHSVitlana1.pdf> (consultato il 20/12/2019).
- Baker M. (1993), *Corpus linguistics and translation studies: implications and applications*, in *Text and Technology: In Honour of John Sinclair*, a c. di M. Baker, G. Francis e E. Tognini-Bonelli, John Benjamins, Amsterdam, pp. 233–352.

- Baker M. (1995), *Corpora in translation studies: an overview and some suggestions for future research*, «Target», 7/2, pp. 223-243.
- Baker M. (1999), *The role of corpora in investigating the linguistic behaviour of professional translators*, «International Journal of Corpus Linguistics», 4/2, pp. 281–298.
- Baker P., Hardie A., McEnery T. (2006), *A Glossary of Corpus Linguistics*, Edinburgh University Press, Edinburgh.
- Balboni P. E. (1994), *Didattica dell'italiano a stranieri*, Bonacci, Roma.
- Balboni P. E. (1999), *Dizionario di glottodidattica*, Guerra Edizioni, Perugia.
- Balboni P. E. (2002), *Le sfide di Babele. Insegnare le lingue nelle società complesse*, UTET, Torino.
- Balboni P. E. (2008), *Imparare le lingue straniere*, Marsilio, Venezia.
- Bambini V., Trevisan M. (2012), *EsploraCoLFIS: Un'interfaccia web per le ricerche sul Corpus e Lessico di Frequenza dell'Italiano Scritto (CoLFIS)*, Quaderni del Laboratorio di Linguistica, 11, Scuola Normale Superiore, Pisa, pp. 1-16.
- Banfi E. (2003 a c. di), *Italiano di cinesi. Percorsi acquisizionali*, FrancoAngeli, Milano.
- Banfi E., Bernini G. (2003), *Il verbo*, in *Verso l'italiano. Percorsi e strategie di acquisizione*, a c. di A. Giacalone Ramat, Carocci, Roma, pp. 70-115.
- Barbera M. (2007), *Per la storia di un gruppo di ricerca. Tra bmanuel.org e corpora.unito.it*, in *Corpora e linguistica in Rete*, a c. di M. Barbera, E. Corino, C. Onesti, Guerra Edizioni. Perugia, pp. 3-20.
- Barbera M. (2013, a c. di), *Una introduzione ai NUNC: storia della creazione di un corpus*, in *Molti occhi sono meglio di uno: saggi di linguistica generale 2008-12*, Qu.A.S.A.R., Milano, pp. 97-114.
- Barbera M. (2015, a c. di), *La linguistica dei corpora in Italia all'alba del terzo millennio*, in *Quanto più la relazione è bella: saggi di storia della lingua italiana 1999-2014*, bmanuel.org. glottologica et philologica-series maior, Youcanprint, Torino, pp. 117-138.
Disponibile online: https://iris.unito.it/retrieve/handle/2318/1670619/420144/Barbera_Relazione%3dSaggiStoriaLinguaItaliana_2015.pdf (consultato il 29/01/2018)
- Barbera M., Marellò C. (2008), *Tra scritto-parlato, Umgangssprache e comunicazione in rete: i corpora NUNC*, in «Studi di Grammatica Italiana», 27 (2008, recte 2011), a c. di A. Antonini, S. Stefanelli. Atti del convegno *Per Giovanni Nencioni. Convegno Internazionale di Studi* (Pisa - Firenze, 4-5 maggio 2009), Le Lettere, Firenze, pp. 157-185.
- Barni M., Bandini, A., Sprugnoli, L., Lucarelli, S., Scaglioso, A.M., Strambi, B., Fusi, C., Arruffoli, A.M. (2009). *Linee guida CILS. Certificazione di Italiano come Lingua Straniera* (presentazione di M. Vedovelli), Guerra Edizioni, Perugia. Disponibile alla pagina: <https://goo.gl/toCDN9> (consultato il 10/01/2019).
- Baroni M. (2010), *Corpora di italiano*, in *Enciclopedia dell'italiano*, a c. di R. Simone, G. Berruto, P. D'Achille (2010-1), Istituto della Enciclopedia Italiana, Roma, pp. 300-303. Disponibile online: http://www.treccani.it/enciclopedia/corpora-di-italiano_%28Enciclopedia-dell%27Italiano%29/ (consultato il 29/01/2018).
- Baroni M., Bernardini S. (2004), *BootCaT: Bootstrapping corpora and terms from the web*, in *Proceedings of Fourth International Conference on Language Resources and Evaluation (LREC 2004)*. Atti del convegno (Lisboa, 26-28 maggio 2004), ELRA - European Language Resources Association, Paris, pp.1313-1316.
- Baroni M., Bernardini S. (2006, a c. di), *Wacky! Working papers on the web as corpus*, Gedit, Bologna. Disponibile online alla pagina <http://wackybook.sslmit.unibo.it/wackycontents.html> (05/02/2019)

- Baroni M., Bernardini S., Comastri F., Piccioni L., Volpi A., Aston G., Mazzoleni M. (2004), *Introducing the La Repubblica corpus: A large, annotated, TEI(XML)-compliant corpus of newspaper Italian*, in *Proceedings of Fourth International Conference on Language Resources and Evaluation (LREC 2004)*. Atti del convegno (Lisboa, 26-28 maggio 2004), ELRA - European Language Resources Association, Paris, pp. 1771 – 1774.
- Baroni M., Bernardini S., Ferraresi A., Zanchetta E. (2009), *The WaCky Wide Web: A Collection of Very Large Linguistically Processed Web-crawled Corpora*, in «*Journal of Language Resources and Evaluation*», 43/3, pp. 209-226.
- Baroni M., Ueyama M. (2006), *Building general- and special-purpose corpora by Web crawling*, in *Proceedings of the NIJL International Workshop on Language Corpora (NIJL, 2006)*. Atti del convegno (Tokyo, marzo 2006), pp. 31 – 40.
- Bazzanella C. (2014), *Linguistica cognitiva. Un'introduzione*, Laterza, Roma-Bari.
- Begotti P. (2011), *Didattizzazione di materiali autentici e analisi di manuali di italiano per stranieri*, FILIM - Formazione degli Insegnanti di Lingua Italiana nel Mondo. Disponibile online: https://www.itals.it/sites/default/files/Filim_didattizzazione_analisi_teoria.pdf (consultato il 06/05/ 2017).
- Begotti P., (2006), *Didattizzazione di materiali autentici e analisi dei manuali di italiano per stranieri*, modulo telematico in ambito del Laboratorio Itals, Venezia. Disponibile online: http://venus.unive.it/filim/materiali/accesso_gratuito/Filim_didattizzazione_analisi_teoria.pdf (consultato il 06/05/ 2017)
- Bembo P. (1989[1966]), *Prose della volgar lingua. Gli Asolani. Rime*, a c. di C. Dionisotti, Tea, Torino.
- Benincà P., Salvi G. / Frison L. (1988), *L'ordine degli elementi della frase e le costruzioni marcate*, in *Grande grammatica italiana di consultazione. Vol. I: La frase. I sintagmi nominale e preposizionale*, a c. di L. Renzi, Salvi G., Cardinaletti A., Il Mulino, Bologna, pp. 129-239.
- Benko V (2016), *Two years of Aranea: Increasing counts and tuning the pipeline*, in *Proceedings of the 10th LREC Conference (LREC 2016)*. Atti del convegno (Portorož, 23-28 maggio 2016), ELRA - European Language Resources Association, Paris, pp. 4245-4248.
- Benko V. (2014a), *Compatible Sketch Grammars for Comparable Corpora*, in *Proceedings of the XVI EURALEX International Congress: The User In Focus*, a c. di A. Abel, C. Vettori, N. Ralli. Atti del convegno (Bolzano/Bozen, 15–19 luglio 2014), Eurac Research, Bolzano/Bozen, pp. 417-430.
- Benko V. (2014b), *Aranea: Yet Another Family of (Comparable) Web Corpora*, in *Text, Speech and Dialogue. Proceedings of the 17th International Conference (TSD 2014)*, a c. di P. Sojka, A. Horák, I. Kopeček, K. Pala. Atti del convegno (Brno, 8-12 settembre 2014), Springer International Publishing, Basel, pp. 257-264.
- Benucci A. (2007, a c. di), *Sillabo di Italiano per Stranieri. Una proposta del Centro Linguistico dell'Università per Stranieri di Siena*, Guerra Edizioni, Perugia.
- Berisso M. (2000), *Livelli linguistici e soluzioni stilistiche. Sondaggi sulla nuova narrativa italiana 1991- 1998*, in «*Lingua e Stile*», 3, pp. 471- 504.
- Berretta M. (1986), *Riprese anaforiche e tipi di testo: il monologo espositivo*, in *Parallela2. Aspetti della sintassi dell'italiano contemporaneo*, a c. di K. Lichem, E. Mara, S. Knaller, Narr, Tübingen, pp. 47-59.
- Berretta M. (1987), *Bricolage di parole in lingua straniera*, «*IeO*», 5, pp. 222-226.

- Berretta M. (1992), *Deissi e anafora nella conversazione*, in *Grammatica del parlare e dell'ascoltare a scuola*, a c. di L. Brasca, M. L. Zambelli, La Nuova Italia, Firenze, pp. 13-31.
- Berretta M. (1993), *Morfologia*, in *Introduzione all'italiano contemporaneo; vol. I: Le strutture; vol. II: La variazione e gli usi* a c. di A.A. Sobrero), La Nuova Italia, Roma/Bari, pp. 193–245.
- Berretta M. (1995), *Ordini marcati dei costituenti maggiori di frase: una rassegna*, in «Linguistica e filologia», 1, pp.125-170.
- Berretta M. (2002), *Quello che voglio dire è che: le scisse da struttura topicalizzanti a connettivi testuali*, in *La parola al testo. Scritti per Bice Mortara Garavelli*, a c. di G.L. Beccaria, C. Marellò, Edizioni dell'Orso, Alessandria, pp. 15-31.
- Berruto G. (1983), *L'italiano popolare e la semplificazione linguistica*, «Vox Romanica» 42, pp. 38-79.
- Berruto G. (1985a), *Per una caratterizzazione del parlato: l'italiano parlato ha un'altra grammatica?*, in *Geschprochenes Italienisch in Geschichte und Gegenwart*, a c. di G. Holtus, E. Radtke, Narr, Tübingen, pp. 120-53.
- Berruto G. (1985b), *'Dislocazioni a sinistra' e 'grammatica' dell'italiano parlato*, in *Sintassi e morfologia della lingua italiana d'uso. Teorie ed applicazioni descrittive*, a c. di A. Franchi De Bellis, L. M. Savoia, Bulzoni, Roma, pp. 59-82.
- Berruto G. (1986), *Le dislocazioni a destra in italiano*, in *Tema-Rema in Italiano*, a c. di H. Stammmerjohann, Narr, Tübingen, pp. 55-69.
- Berruto G. (1987), *Sociolinguistica dell'italiano contemporaneo*, La Nuova Italia Scientifica, Roma.
- Berruto G. (1990), *Semplificazione linguistica e varietà sub-standard*, in *Sprachlicher Substandard. Vol. III*, a c. di G.Holtus, E. Radtke (1986-1990), pp. 17-43.
- Berruto G. (1993a), *Le varietà del repertorio*, in *Introduzione all'italiano contemporaneo. Vol. II*, a c. di A. A. Sobrero, pp. 3-36.
- Berruto G. (1993b), *Varietà diamesiche, diastratiche, diafasiche*, in *Introduzione all'italiano contemporaneo. Vol. II*, a c. di A. A. Sobrero, pp. 37-92.
- Berruto G. (2010), *Italiano standard*, in *Enciclopedia dell'italiano*, a c. di R. Simone, G. Berruto, P. D'Achille (2010-1). Disponibile alla pagina: [https://www.treccani.it/enciclopedia/italiano-standard_\(Enciclopedia-dell'Italiano\)/](https://www.treccani.it/enciclopedia/italiano-standard_(Enciclopedia-dell'Italiano)/) (consultato il 06/12/2017).
- Berruto G. (2011a), *Substandard*, in *Enciclopedia dell'italiano*, a c. di R. Simone, G. Berruto, P. D'Achille (2010-1). Disponibile alla pagina: https://www.treccani.it/enciclopedia/substandard_%28Enciclopedia-dell%27Italiano%29/ (consultato il 06/12/2017).
- Berruto G. (2011b), *Varietà*, in *Enciclopedia dell'italiano*, in Simone R., Berruto G., D'Achille P., (2010-1, a c. di). Disponibile alla pagina: [https://www.treccani.it/enciclopedia/variet%C3%A0_\(Enciclopedia-dell%27Italiano%29\)/](https://www.treccani.it/enciclopedia/variet%C3%A0_(Enciclopedia-dell%27Italiano%29)/) (consultato il 06/12/2017).
- Berruto G. (2012 [1987]), *Sociolinguistica dell'italiano contemporaneo*, La Nuova Italia Scientifica, Roma, 1987, Carocci, Roma, 2012.
- Berruto G. (2012), *Sul parlante nativo (di italiano)*, in *Idem, Saggi di sociolinguistica e linguistica*, a c. di G. Bernini, B. Moretti, S. Schmid, T. Telmon, Edizioni dell'Orso, Alessandria, pp. 87–99.
- Berruto G., Cerruti M. (2014), *Manuale di sociolinguistica*, UTET, Torino.
- Bettoni C., Di Biase B. (2005), *Sviluppo obbligatorio e progresso morfosintattico. Un caso di Processabilità in italiano L2*, in «ITALS: Didattica e linguistica dell'italiano come lingua straniera», a c. di P. E. Balboni, M. Santipolo, 3/7, pp. 27-48.

- Bettoni, C., Di Biase, B. (2015), *Processability theory: theoretical bases and universal schedules*, in *Grammatical Development in Second Languages: Exploring the Boundaries of Processability Theory*, a c. di C. Bettoni, B. Di Biase, The European Second Language Association, Amsterdam, pp. 19-80.
- Biber D. (2012), *Corpus-Based and Corpus-driven Analyses of Language Variation and Use*, in *The Oxford Handbook of Linguistic Analysis (1 ed.)*, a c. di B. Heine e H. Narrog, Disponibile online:
<https://eclass.uoa.gr/modules/document/file.php/MEDIA139/Readings%20//%20%CE%91%CE%BD%CE%B1%CE%B3%CE%BD%CF%8E%CF%83%CE%BC%CE%B1%CF%84%CE%B1/Biber%2C%20Corpus-Based%20and%20Corpus-Driven%20Analyses%20of%20Language%20Variation%20and%20Use.pdf>
 (consultato il 24/01/2020).
- Biber D., Conrad S., Reppen R. (1998), *Corpus Linguistics: Investigating Language Structure and Use*, Cambridge University Press, Cambridge.
- Bolasco S. (1999), *Analisi multidimensionale dei dati. Metodi, strategie e criteri d'interpretazione*, Carocci, Roma.
- Bolwler L., Pearson J. (2002), *Working with Specialized Language. A practical Guide to Using Corpora*, Routledge, London- New York.
- Bonomi I. (1996), *La narrativa e l'italiano dell'uso medio*, in «Studi di Grammatica Italiana», 16, pp.321-338.
- Bonomi I. (1997), *La grammatica del parlato in alcuni scrittori contemporanei*, in *Norma e lingua in Italia: alcune riflessioni fra passato e presente*, Istituto Lombardo di Scienze e Lettere Milano, pp. 167-183.
- Bonomi I. (2002), *L'italiano giornalistico. Dall'inizio del '900 ai quotidiani on line*, Cesati, Firenze.
- Bonomi I., Mauroni E. (2003), *L'innovazione grammaticale in testi scritti di fine millennio*, in *Italia linguistica anno Mille. Italia linguistica anno Duemila*, a c. di N. Maraschio, T. Poggi Salani, Atti del XXXIV Congresso Internazionale di Studi della SLI (Firenze, 19-21 ottobre 2000), Bulzoni, Roma, pp. 491-519.
- Bonvino E. (2004), *I materiali didattici per l'insegnamento dell'italiano L2: criteri di selezione*, ICON. Disponibile online:
<http://grafiche.radioshock.info/uni/corsoscuolaita/2lezione/Criteri%20di%20selezione%20materiali%20didattici.pdf> (consultato il 06/05/ 2017).
- Borghetti C., Castagnoli S., Brunello M. (2010), *I generi del web tra tradizione e innovazione: un'analisi linguistica sulla base del corpus PAISÀ*. Presentazione invitata alla conferenza *Scritto e parlato, formale e informale: La comunicazione mediata dalla rete* (Università di Torino, Torino, 29-30 ottobre 2010).
- Bortolini U., Pizzuto E. (1997, a c. di), *Il progetto CHILDES-Italia, Contributi di ricerca sulla lingua italiana*, Edizioni del Cerro, Tirrenia.
- Bortolini U., Zampolli A. (1971), *Lessico di frequenza della lingua italiana contemporanea: prospettive metodologiche*, in *L'insegnamento dell'Italiano in Italia e all'Estero*, a c. di M. Medici, R. Simone. Atti del IV Convegno internazionale di Studi (Roma, 1-2 giugno 1970), II, Bulzoni, Roma, pp. 639-648.
- Breen M. P. (1985), *Authenticity in the language classroom*, «Applied linguistics», 6, pp. 60-70.
- Brezina V., Gablasova D., Marín Cervantes I. (2018), *Using corpora to teach sociolinguistics: A practical workshop*, workshop tenuto alla 13a conferenza internazionale su Teaching and Language Corpora (TaLC13), Cambridge University, 18-21 luglio 2018. Disponibile online:

- <https://www.educ.cam.ac.uk/events/conferences/talc2018/programme/BookofAbstractsFinalversion.pdf> (consultato il 20/12/2019).
- Bresnan J. (2001), *Lexical-functional syntax*, Blackwell, Malden, (MA).
- Bruni F. (1984), *L'italiano. Elementi di storia della lingua e della cultura*, UTET, Torino.
- Buchholz S., Marsi E. (2006), *CoNLL-X Shared Task on Multilingual Dependency Parsing*, in *Proceedings of the Tenth Conference on Computational Natural Language Learning (CoNLL-X '06)*. Atti del convegno (New York 8-9 giugno 2006), Association for Computational Linguistics, Stroudsburg, pp. 149–164.
- Budassi, M. (2017), *Estrazione del 'che' polivalente da un corpus POS-tagato: Limiti e possibilità*, «Analisi Linguistica e Letteraria», 25, 147-172.
- Budinich L. (2016), *La variazione semasiologica di sottocodice: un esempio di analisi lessicale corpus-based*, in «CHIMERA. Romance Corpora and Linguistic Studies», 3/1, pp. 23-56.
- Calleri D., Chini M., Cordin P., Ferraris P. (2003), *Confronti tra l'acquisizione di italiano L1 e l'acquisizione di italiano L2*, in *Verso l'italiano*, a c. di A. Giacalone Ramat, Carocci, Roma, pp. 220-253.
- Cardillo G., Vecchio P. (2015a, a c. di), *Nuovo sillabo della Certificazione PLIDA. Livelli A1-C2*. Roma: Società Dante Alighieri. Disponibile alla pagina: <https://goo.gl/bgGxxH> (consultato il 10/01/2019).
- Cardillo G., Vecchio P. (2015b a c. di), *Livello PLIDA B1 - Quaderno delle specifiche*. Società Dante Alighieri. Disponibile alla pagina: <https://plida.it/certificazione-plida/documenti.html?start=0> (consultato il 10/01/2019).
- Cardillo G., Vecchio P. (2016 a c. di), *Livello PLIDA B2 - Quaderno delle specifiche*. Società Dante Alighieri. Disponibile alla pagina: <https://plida.it/certificazione-plida/documenti.html?start=0> (consultato il 10/01/2019).
- Cardillo G., Vecchio P. (2017 a c. di), *Livello PLIDA C1 - Quaderno delle specifiche*. Società Dante Alighieri. Disponibile alla pagina: <https://plida.it/certificazione-plida/documenti.html?start=20> (consultato il 10/01/2019).
- Cardillo G., Vecchio P. (2018 a c. di), *Livello PLIDA A2 - Quaderno delle specifiche*. Società Dante Alighieri. Disponibile alla pagina: <https://plida.it/certificazione-plida/documenti.html?start=0> (consultato il 10/01/2019).
- Cardinale U. (2011), *Manuale di scrittura giornalistica*, UTET Università, Torino.
- Castellani A. (1991), *Italiano dell'uso medio o italiano senz'aggettivi?*, in «Studi Linguistici Italiani», 17, pp. 233-256.
- Castellani A. (1991), *Italiano dell'uso medio o italiano senza aggettivi?*, in «Studi linguistici italiani», 17, pp. 233-56.
- Charkova K., Charkova D., *A linguistically informed, systematic, and in-depth approach to developing data driven language learning materials*, contributo presentato alla 11a conferenza internazionale su *Teaching and Language Corpora (TaLC11)*, Lancaster University, 20-23 luglio 2014. Disponibile online: <http://ucrel.lancs.ac.uk/talc2014/doc/TALC2014-abstract-book.pdf> (consultato il 20/12/2019).
- Chini M. (2012b), *Linguistica educativa e linguistica acquisizionale: qualche suggerimento dalla ricerca su L2*, in *Linguistica educativa*, a c. di S. Ferreri, Bulzoni, Roma, pp.123-140.
- Chini M. (2005), *Che cos'è la linguistica acquisizionale*, Carocci, Roma.
- Chini M. (2011), *Qualche riflessione sulla didattica di L2 ispirata alla recente ricerca acquisizionale*, in «Italiano LinguaDue», 3/2, pp. 1-22. Disponibile online: <http://riviste.unimi.it/index.php/promoitals/article/view/1912> (consultato il 20/08/2020).

- Chini M. (2012a), *Alcune ricadute didattiche degli studi acquisizionali sull'italiano L2*, in «Rassegna Italiana di Linguistica Applicata. Numero monografico L'italiano L2 nella scuola secondaria di secondo grado», a cura di C. Bosisio, B. Cambiagli, M. Zanola, 44/ 1, pp. 69-91.
- Chini M. (2012c), *Spunti per la didattica linguistica a partire da alcuni recenti filoni di studio di L2*, in *(Far) apprendere, usare e certificare una lingua straniera. Studi in onore di Serena Ambroso*, a c. di E. Bonvino, E. Luzi, A.R. Tamponi, Bonacci, Roma, pp. 61-74.
- Chini M. (2016), *Elementi utili per una didattica dell'italiano L2 alla luce della ricerca acquisizionale*, «Italiano LinguaDue», 8/2, pp 1-18. Disponibile online: <https://riviste.unimi.it/index.php/promoitals/article/view/8172> (consultato il 20/08/2020).
- Christ O. (1994a), *The IMS Corpus Workbench Technical Manual*. Institut für maschinelle Sprachverarbeitung, Universität Stuttgart.
- Christ O. (1994b), *The XKwic User Manual*. Institut für maschinelle Sprachverarbeitung, Universität Stuttgart.
- Cignetti L. (2006), *L'ordine delle parole nello scritto e nel parlato (con alcune osservazioni sul fenomeno della "doppia dislocazione"*, in *Parole frasi testi tra scritto e parlato*, a c. di A. Ferrari, in «Cenobio», 55/3, pp. 207-214.
- Cignetti L. (2010), *Dativo etico*, in *Enciclopedia dell'italiano*, a c. di R. Simone, G. Berruto, P. D'Achille (2010-1), Disponibile alla pagina: [https://www.treccani.it/enciclopedia/dativo-etico_\(Enciclopedia-dell'Italiano\)/](https://www.treccani.it/enciclopedia/dativo-etico_(Enciclopedia-dell'Italiano)/) (consultato il 21/09/2020)
- Ciliberti A. (1994), *Manuale di glottodidattica*, La Nuova Italia, Firenze.
- Comodi A. (1995), *Materiali autentici: selezione e uso nella didattica dell'italiano come lingua straniera*, Guerra edizioni, Perugia.
- Consiglio d'Europa (2002), *Quadro comune europeo di riferimento per le lingue: apprendimento insegnamento, valutazione*, traduzione di F. Quartapelle e D. Bertocchi, La Nuova Italia, Firenze.
- Consiglio d'Europa (2018), *Common European framework of reference for languages: learning, teaching, assessment. Companion volume with new descriptors*.
- Coonan C. M. (2000), *La ricerca azione*, Università Ca'Foscari, Venezia.
- Cordin P. (2001), *La frase: i pronomi personali*, in *Grande grammatica italiana di consultazione. Vol. I: La frase. I sintagmi nominale e preposizionale*, a c. di L. Renzi, Salvi G., Cardinaletti A., Il Mulino, Bologna, pp. 549-606.
- Corino E. (2014), *Didattica delle lingue corpus based*, «EL.LE», 3/2, pp. 231-256.
- Corino E., Marellò C. (2009), *Didattica con i corpora di italiano per stranieri*, «Italiano LinguaDue», pp. 279-285.
- Cortelazzo M. (1972), *Avviamento critico allo studio della dialettologia italiana. Vol. III: Lineamenti di italiano popolare*, Pacini, Pisa.
- Cortelazzo M. (1976 [1972]), *Avviamento critico allo studio della dialettologia italiana. Vol III: Lineamenti di italiano popolare*, Pacini, Pisa.
- Cortelazzo M. (1984), *Perché "a mí me gusta" sì e "a me mi piace" no?*, in *Umgangssprache im Iberoromanischen. Festschrift für Heinz Kröll*, a c. di G. Holtus, E. Radtke, Narr, Tübingen, pp. 25-28.
- Cortelazzo M. A. (2000a), *La lingua italiana di fine millennio*, in *Idem, Italiano d'oggi*, Esedra, Padova, pp. 9-24.
- Cortelazzo M. A. (2000b), *Per la storia dell'italiano scolastico*, in *Idem, Italiano d'oggi*, Esedra, Padova, pp. 91-109.

- Cortelazzo M. A. (2012), *Evoluzione della lingua, percezione del cambiamento, staticità della norma*, in *I sentieri della lingua. Saggi sugli usi dell'italiano tra passato e presente*, a c. di C. Di Benedetto, S. Ondelli, A. Pezzin, S. Tonello, V. Ujčić, M. Viale, Esedra, Padova, pp. 15–20.
- Coseriu E. (1973), *Lezioni di linguistica generale*, Boringhieri, Torino.
- Cosme C., Gilquin, G. (2008), *Free and bound prepositions in a contrastive perspective. The case of with and avec*, in *Phraseology: An Interdisciplinary Perspective*, a c. di S. Granger, F. Meunier, John Benjamins, Amsterdam, pp. 259-274.
- Coxhead A. (2008), *Phraseology and English for academic purposes: Challenges and opportunities*, in *Phraseology Foreign Language Learning and Teaching*, a c. di F. Meunier, S. Granger, John Benjamins, Amsterdam / Philadelphia, pp. 149-161.
- Cresti E. (2000), *Corpus di italiano parlato*, Accademia della Crusca, Firenze.
- Cresti E., Panunzi A. (2013), *Introduzione ai corpora dell'italiano*, il Mulino, Bologna.
- Curado-Fuentes A. (2012), *Keyword Decoding Reliability for News B1-level Reading Comprehension*, contributo presentato alla 10a conferenza internazionale su *Teaching and Language Corpora (TaLC10)*, Uniwersytet Warszawski, 11-14 luglio 2012. Disponibile online: <http://talc10.ils.uw.edu.pl/files/2011/09/CURADO-FUENTES1.pdf> (consultato il 20/12/2019).
- Curry N., Goodman O. (2018), *Using the Cambridge learner corpus to develop learning materials*, workshop tenuto alla 13a conferenza internazionale su *Teaching and Language Corpora (TaLC13)*, Cambridge University, 18-21 luglio 2018. Disponibile online: <https://www.educ.cam.ac.uk/events/conferences/talc2018/programme/BookofAbstractsFinalversion.pdf> (consultato il 20/12/2019).
- D'Achille P. (1990), *Sintassi del parlato e tradizione scritta della lingua italiana. Analisi di testi dalle origini al secolo XVIII*, Bonacci, Roma.
- D'Achille P. (1994), *L'italiano dei semicolti*, in: *Storia della lingua italiana. Vol. II Scritto e parlato*, a c. di A. Asor Rosa, L. Serianni, P. Trifone, Einaudi, Torino, pp. 41-79.
- D'Achille P. (2003), *L'italiano contemporaneo*, Il Mulino, Bologna.
- D'Achille P. (2010), *Lingua d'oggi*, in *Enciclopedia dell'italiano, Vol. I*, Istituto dell'Enciclopedia Italiana, Roma, pp. 793–799.
- D'Achille P. (2016), *Architettura dell'italiano di oggi e linee di tendenza*, in *Manuale di linguistica italiana*, a c. di S. Lubello, De Gruyter, Berlino/Boston, pp. 165-189.
- D'Achille, P., Proietti, D., Viviani, A. (2005), *La frase scissa in italiano: aspetti e problemi*, in *Tipologia linguistica e società. Considerazioni inter- e intralinguistiche. Due Giornate italo-danesi di studi linguistici (Roma, 27–28 novembre 2003)*, a c. di P. D'Achille, I. Korzen, Cesati, Firenze, pp. 249–279.
- Daloiso M. (2009), *I fondamenti neuropsicologici dell'educazione linguistica*, Cafoscarina, Venezia.
- Daloiso M. (2015), *Scienze del linguaggio e educazione linguistica*, Bonacci/Loescher, Torino.
- Danesi M. (1988), *Neurolinguistica e glottodidattica*, Liviana, Padova.
- Danesi M. (1998), *Il cervello in aula!*, Guerra edizioni, Perugia.
- Dardano M. (1994), *Profilo dell'italiano contemporaneo*, in *Storia della lingua italiana. Vol. II: Scritto e parlato* a c. di L. Serianni, P. Trifone, Einaudi, Torino, pp. 343–432.
- Dardano M. (2020, a c. di), *Sintassi dell'italiano antico II. La prosa del Duecento e del Trecento. La frase semplice*, Carocci, Roma.
- De Cesare A. M. (2006), *'C'è la tua bambina con i fiammiferi'. Funzioni del costrutto presentativo c'è... che...*, in *Parole frasi testi tra scritto e parlato*, a c. di A. Ferrari, in «Cenobio», 55/3, pp. 215-221.

- de Haan P. (1992), *The optimum corpus sample size?*, in *New Directions in English Language Corpora: Methodology, Results, Software Developments*, a c. di G. Leitner, Mouton de Gruyter, Berlin, pp. 3-19.
- De Mauro T. (1970), *Per lo studio dell'italiano popolare unitario*, in *Lettere da una tarantata*, a c. di A. Rossi, De Donato, Bari, pp. 43-75.
- De Mauro T. (1997 [1980]), *Guida all'uso delle parole*, Editori Riuniti, Roma.
- De Mauro T. (2011 [1983]), *Storia della lingua Italiana*, Laterza, Bari.
- De Mauro T. (2014), *Storia linguistica dell'Italia repubblicana dal 1946 ai nostri giorni*, Laterza Roma/Bari.
- De Mauro T., Mancini F., Vedovelli M., Voghera M. (1993), *Lessico di frequenza dell'italiano parlato*, Etaslibri, Milano.
- De Roberto E. (2010), *Strutture assolute*, in *Enciclopedia dell'italiano*, a c. di R. Simone, G. Berruto, P. D'Achille (2010-1). Disponibile alla pagina: [https://www.treccani.it/enciclopedia/strutture-assolute_\(Enciclopedia-dell'Italiano\)/](https://www.treccani.it/enciclopedia/strutture-assolute_(Enciclopedia-dell'Italiano)/) (consultato il 23/09/2020).
- De Santis C. (2003), *La parola alle narratrici (e ai narratori)*, in *Quaderni dell'Osservatorio Linguistico, vol. I*, a c. di F. Frasnè et al., Franco Angeli, Milano, pp. 54-96.
- Della Putta P., Visigalli M. (2010), *Come aiutare l'acquisizione della morfosintassi del sintagma nominale italiano: confronto fra tre percorsi glottodidattici rivolti a studenti anglofoni*, «Italiano LinguaDue», 2/2, pp. 23-52. Disponibile online: <http://riviste.unimi.it/index.php/promoitals/article/view/821/1062> (consultato il 20/08/2020).
- Diadori P., Palermo M., Troncarelli D. (2015), *Insegnare italiano come seconda lingua*, Carocci, Roma.
- Diadori P., Semplici S., Troncarelli D. (2020, a c. di), *Didattica di base dell'italiano L2*, Carocci, Roma.
- Doughty C. (2003), *Instructed SLA: constraints, compensation, and enhancement*, in *The handbook of second language acquisition*, a c. di C.J. Doughty, M.H. Long, Blackwell, Oxford, pp. 256-310.
- Egerland V. (1996), *The syntax of past participles. A generative study of nonfinite constructions in ancient and modern Italian*, Lund University Press, Lund.
- Ellis R. (1997), *SLA research and language teaching*, Oxford University Press, Oxford.
- Evans V. (2007), *A Glossary of Cognitive Linguistics*, Edinburgh University Press, Edinburgh.
- Evert S. (2009), *The CQP Query Language Tutorial*. Institute for Natural Language Processing, University of Stuttgart. Disponibile online: <https://pdfs.semanticscholar.org/6a0b/ed8affc49daf0f0b7ccfb35da5abc8f732fa.pdf> (consultato il 28/07/2020).
- Ferrari A. (1999), *L'extra-posizione a destra in italiano, con osservazioni sul francese*, in *Linguistica testuale comparativa*, a c. di G. Skytte, F. Sabatini, Museum Tusulanum Press, København, pp. 111-140.
- Ferrari A. (2003), *Le ragioni del testo. Aspetti sintattici e interpuntivi dell'italiano contemporaneo*, Accademia della Crusca, Firenze.
- Ferrari A. (2008, a c. di), *L'interfaccia lingua testo. Natura e funzioni dell'articolazione informativa dell'enunciato*, Edizioni dell'Orso, Alessandria.
- Ferraris S. (1999), *Imparare la sintassi*, Mercurio, Bergamo.
- Fiorentino G. (2010), *Che polivalente*, in *Enciclopedia dell'italiano*, in Simone R., Berruto G., D'Achille P., (2010-1, a c. di). Disponibile alla pagina:

- [http://www.treccani.it/enciclopedia/che-polivalente_\(Enciclopedia-dell'Italiano\)/](http://www.treccani.it/enciclopedia/che-polivalente_(Enciclopedia-dell'Italiano)/)
(Consultato il 19/09/2020)
- Firth J. R. (1957), *Papers in Linguistics 1934-1951*, Oxford University Press, London.
- Freddi G. (1970), *Metodologia e didattica delle lingue straniere*, Minerva Italica, Bergamo.
- Freddi G. (1994), *Glottodidattica. Fondamenti, metodi, tecniche*, UTET, Torino.
- Freddi G. (1999), *Psicolinguistica, sociolinguistica, glottodidattica*, UTET, Torino.
- Frosini G. (2020, a c. di), *Storia dell'italiano. La lingua, i testi*, Salerno, Roma.
- Galli de' Paratesi N. (1984), *Lingua toscana in bocca ambrosiana. Tendenze verso l'italiano standard: un'inchiesta sociolinguistica*, Il Mulino, Bologna.
- Gatta F. (2014), *Giornalismo*, in *Storia dell'italiano scritto, Vol III: Italiano dell'uso*, a c. di G. Antonelli, M. Motolese e L. Tomasin, Carocci, Roma, pp. 293-347.
- Geeraerts D. (2006), *Methodology in cognitive linguistics*, in *Cognitive linguistics: current applications and future perspectives*, a c. di G. Kristiansen, M. Achard, R. Dirven, F. J. R. Ruiz de Mendoza Ibáñez, Mouton de Gruyter, Berlin-New York.
- Giacalone Ramat A. (2003, a c. di), *Verso l'italiano. Percorsi e strategie di acquisizione*, Carocci, Roma.
- Giacalone Ramat A., Chini M., Andorno C. (2013), *Italiano come L2*, in *La linguistica italiana all'alba del terzo millennio (1997-2010)*, a c. di G. Iannàccaro, Bulzoni, Roma, pp. 149-206.
- Giorgi A., Pianesi, F. (1997), *Tense and Aspect. From Semantics to Morphosyntax*, Oxford University Press, New York/Oxford.
- Giuliano L., La Rocca G. (2008), *L'analisi automatica e semi-automatica dei dati testuali. Software e istruzioni per l'uso*, LED, Milano.
- Glisan E., Drescher V. (1993), *Textbook grammar: Does it reflect native speaker speech?*, «The Modern Language Journal», 77/1, pp. 23-33.
- Gossen C. T. (1954), *Studien zur syntaktischen und stilistischen Hervorhebung im modernen Italienisch*, Akademie Verlag, Berlin.
- Grandi, N. (2011), *Costruzione passiva*, in *Enciclopedia dell'italiano*, a c. di R. Simone, G. Berruto, P. D'Achille (2010-1). Disponibile online: [http://www.treccani.it/enciclopedia/costruzione-passiva_\(Enciclopedia-dell'Italiano\)/](http://www.treccani.it/enciclopedia/costruzione-passiva_(Enciclopedia-dell'Italiano)/) (consultato il 01/06/2019).
- Grandi N. (2018), *Sulla penetrazione di tratti neo-standard nell'italiano degli studenti universitari. Primi risultati di un'indagine empirica*, Griseldaonline (sez. Dibattiti), Disponibile alla pagina: <https://griseldaonline.unibo.it/article/view/9021> (consultato il 14/09/2020).
- Grandi N. (2019), *Che tipo, l'italiano neostandard!*, in *Le tendenze dell'italiano contemporaneo rivisitate. Atti del LII Congresso Internazionale di Studi della Società di Linguistica Italiana* (Berna, 6-8 settembre 2018), Officinaventuno, Milano, pp. 59-74.
- Granger S., Meunier F. (2008, a c. di), *Phraseology: An Interdisciplinary Perspective*, John Benjamins, Amsterdam.
- Grassi R., Bozzone Costa R., Ghezzi C. (2008, a c. di), *Dagli studi sulle sequenze di acquisizione alla classe di italiano L2*, Guerra, Perugia.
- Gries S. T., Stefanowitsch A. (2006, a c. di), *Corpora in cognitive linguistics: corpus-based approaches to syntax and lexis*, Mouton de Gruyter, Berlin-New York.
- Gries S. T., Wulff, S. (2005), *Do foreign language learners also have constructions? Evidence from priming, sorting, and corpora*, in «Annual Review of Cognitive Linguistics», 3, pp.182-200.

- Gries S. T., Yoon J. (2016, a c. di), *Corpus-based Approaches to Construction Grammar (Constructional Approaches to Language)*, John Benjamins, Amsterdam-Philadelphia.
- Gualdo R. (2007), *L'italiano dei giornali*, Carocci, Roma.
- Guidetti M.G., Lenzi G., Storchi, S. (2012), *Potenzialità e limiti dell'uso dei corpora linguistici per la didattica dell'italiano LS*, «Bollettino ITALS», 44. Disponibile online: <https://www.itals.it/potenzialit%C3%A0-e-limiti-dell%E2%80%99uso-dei-corpora-linguistici-la-didattica-dell%E2%80%99italiano-ls> (consultato il 05/01/2019)
- Hadley G, Hadley H. (2018), *Investigating data-driven learning with receptive skills: a contemporary approach*, contributo presentato alla 13a conferenza internazionale su *Teaching and Language Corpora (TaLC13)*, Cambridge University, 18-21 luglio 2018. Disponibile online: <https://www.educ.cam.ac.uk/events/conferences/talc2018/programme/BookofAbstractsFinalversion.pdf>(consultato il 20/12/2019).
- Harrich K. (2011), *Il notiziario ovvero l'importanza dell'uso di materiali autentici nell'insegnamento – apprendimento linguistico*, «Italiano LinguaDue», 2/2, pp. 321-335. Disponibile online: <http://riviste.unimi.it/index.php/promoitals/article/view/835> (consultato il 02/01/2020)
- Hilpert M. (2013), *Corpus-based approaches to constructional change*, in *The Oxford Handbook of Construction Grammar*, a c. di G.Trousdale, T. Hoffmann, Oxford University Press, Oxford, pp. 458-477.
- Hinkel E. (2005, a c. di), *Handbook of research in second language teaching and learning*, Erlbaum, Mahwah NJ.
- Hinkel E. (2011, a c. di), *Handbook of Research in Second Language Teaching and Learning: Volume 2*, Routledge, New York/ London.
- Hoffmann T., Trousdale G. (2013, a c. di), *The Oxford Handbook of Construction Grammar*, Oxford University Press, Oxford-New York.
- Holtus G. (1984), *Codice parlato e codice scritto*, in *Il dialetto dall'oralità alla scrittura. Atti del XIII Convegno per gli Studi Dialettali Italiani* (Catania-Nicosia, 28 settembre 1981), Pacini, Pisa, pp. 1-12.
- Holtus G., Radtke E. (1986-1990, a c. di), *Sprachlicher Substandard. Vol I: Sprachlicher Substandard; vol II: Standard und Substandard in der Sprachgeschichte und in der Grammatik; vol. III: Standard, Substandard und Varietätenlinguistik*, Niemeyer, Tübingen.
- Housen A., Pierrard M. (2005, a c. di), *Investigations in instructed second language acquisition*, Mouton de Gruyter, Berlin.
- Hunston S. (2002), *Corpora in Applied Linguistics*, Cambridge University Press, Cambridge.
- Hunston S, Francis G. (2000), *Pattern grammar. A corpus-driven approach to the lexical grammar of English*, John Benjamins, Amsterdam / Philadelphia.
- Ivanovska-Naskova R. (2014), *Il corpus parallelo italiano-macedone come strumento nella didattica dell'italiano LS*, in *Dal manoscritto al web: canali e modalità di trasmissione dell'italiano. Tecniche, materiali e usi nella storia della lingua*, a c. di E. Garavelli, E. Suomela-Härmä. Atti del XII Congresso SILFI - Società Internazionale di Linguistica e Filologia Italiana (Helsinki, 18–20 giugno 2012), Franco Cesati Editore, Firenze, pp. 471–479.
- Ivanovska-Naskova R. (2015), *I corpora e l'insegnamento dell'italiano LS*, in *Incroci. Studi sulla letteratura, la traduzione e la glottodidattica*, a c. di V. Zaccaro, R. Ivanovska-Naskova, Università degli Studi di Bari “Aldo Moro” e Università “Ss. Cirillo e Metodio” di Skopje, Bari/Skopje, pp. 105-123.

- Ivanovska-Naskova R. (2018), *L'insegnamento della grammatica dell'italiano LS attraverso corpora*, «Italice Wratislaviensia», 2018, 9/1, pp.
- Jakubiček M., Kilgarriff A., Kovář V., Rychlý P., Suchomel V. (2013), *The tenten corpus family*, in *7th International Corpus Linguistics Conference (CL2013)*. Atti del Convegno (Lancaster University 23-26 luglio 2013), UCREL – Lancaster University, Lancaster, pp. 125-127.
- Ježek E. (2011), *Verbi pronominali*, in Enciclopedia dell'italiano, a c. di R. Simone, G. Berruto, P. D'Achille (2010-1). Disponibile online: [www.treccani.it/enciclopedia/verbi-pronominali_\(Enciclopedia-dell'Italiano\)/](http://www.treccani.it/enciclopedia/verbi-pronominali_(Enciclopedia-dell'Italiano)/) (consultato il 05/01/2019)
- Johns T.F. (1991), *Should you be persuaded?-Two Samples of Data-Driven Learning Materials*, «English Language Research Journal», 4, pp.1-16.
- Jones M., Haywood S. (2004), *Facilitating the acquisition of formulaic sequences: An exploratory study in an EAP context*, in *Formulaic sequences*, a c. di N. Schmitt, John Benjamins Amsterdam, pp. 269-27.
- Kennedy C., Miceli, T. (2010), *Corpus-Assisted Creative Writing: Introducing Intermediate Italian Learners to a Corpus as a Reference Resource*, «Language Learning & Technology», 14/1, pp. 28-44.
- Kilgarriff A., Grefenstette G. (2003), *Introduction to the special issue on the Web as a Corpus*, in «Computational linguistics», 29/3, pp. 333-347.
- Koch P. (2014), *La scelta manzoniana tra selezione e ristandardizzazione*, in *Dall'architettura della lingua italiana all'architettura linguistica dell'Italia. Saggi in omaggio a Heidi Siller-Runggaldier*, a c. di P. Danler, C. Konecny, Lang, Frankfurt am Main et al., pp. 75-102.
- Koffka K. (1935), *Principles of Gestalt-Psychology*, Harcourt, Brace & Company, New York.
- Köhler W. (1929), *Gestalt Psychology*, Liveright, New York.
- Krashen S. D. (1982), *Principle and Practice in Second Language Acquisition*, Pergamon, Oxford.
- Krashen S. D. (1985), *The Input Hypothesis: Issues and Implications*, Longman, London.
- Lambrecht K. (1994), *Information Structure and Sentence Form. Topic, Focus, and the Mental Representations of Discourse Referents*, Cambridge University Press, Cambridge.
- Laudanna A., Thornton A.M., Brown G., Burani C., Marconi L. (1995), *Un corpus dell'italiano scritto contemporaneo dalla parte del ricevente*, in *JADT 1995. III Giornate internazionali di Analisi Statistica dei Dati Testuali*, a c. di S. Bolasco, L. Lebart, A. Salem, I. Cisu, Roma, pp. 103-109.
- Lavinio C. (1986), *Tipologia dei testi parlati e scritti*, «Linguaggi», 3/1-2, pp. 14-22.
- Leech G. (1997), *Teaching and language corpora: A convergence*, in *Teaching and language corpora*, a c. di A. Wichmann, S. Fligelstone, T. McEnery e G. Knowles, Longman, London, pp. 1-23.
- Lepschy A.L., Lepschy, G.C. (1988), *La lingua italiana: storia, varietà dell'uso, grammatica*, Bompiani, Milano.
- Levelt W. J. M. (1989), *Speaking. From intention to articulation*, MIT Press, Cambridge (Ma).
- Lewandowska-Tomaszczyk B., Dziwirek K. (2009 a c. di), *Studies in cognitive corpus linguistics*, Peter Lang, Frankfurt am Main.
- Liscia R. (2004, a c. di), *E-learning. Stato dell'arte e prospettive di sviluppo*, Apogeo.
- Lo Duca M.G. (1990), *Creatività e regole*, Il Mulino, Bologna.

- Lo Duca M.G., Solarino R. (2006), *Lingua italiana. Una grammatica ragionevole*, Unipress, Padova.
- Lo Duca, M.G. (2006), *Sillabo di italiano L2*, Carocci, Roma.
- Loporcaro M., Seiler S. (2009), *Evoluzione diacronica delle subordinate participiali in italiano*, in *Sintassi storica e sincronica dell'italiano. Subordinazione, coordinazione e giustapposizione, vol. I*, a cura di A. Ferrari. Atti del X congresso della Società Internazionale di Linguistica e Filologia italiana (Basilea 30 giugno - 3 luglio 2008), Cesati, Firenze, pp. 485-512.
- Lorenzetti L. (2002), *L'italiano contemporaneo*, Carocci, Roma.
- Lüdeling A., Kytö M. (2008-2009, a c. di), *Corpus linguistics. An international handbook. Vol. I e II*, Mouton de Gruyter, Berlin.
- Lyding V., Borghetti C., Dittmann H., Nicolas L., Stemle E. (2013), *Open corpus interface for Italian language learning*, in *Proceeding of ICT for Language Learning*, VI ed.. Atti del convegno (Firenze 14-15 novembre 2013). Disponibile online: https://www.researchgate.net/publication/257043915_Open_Corpus_Interface_for_Italian_Language_Learning (consultato il 05/01/2020)
- Lyding V., Chiocchetti E., Sérasset G., Brunet-Manquat F. (2006), *The LexALP Information System: Term Bank and Corpus for Multilingual Legal Terminology Consolidated*, in *Proceedings of the Workshop on Multilingual Language Resources and Interoperability (MLRI '06)*. Atti del convegno (Sydney, luglio 2006), Association for Computational Linguistics Stroudsburg, PA, pp. 25-31.
- Lyding V., Stemle E., Borghetti C., Brunello M., Castagnoli S., Dell'Orletta F., Dittmann H., Lenci A., Pirrelli V. (2014), *The PAISÀ Corpus of Italian Web Texts* in *Proceedings of the 9th Web as Corpus Workshop (WaC-9)* (EACL 2014). Atti del Convegno (Gothenburg, 26 aprile 2014), Association for Computational Linguistics, Stroudsburg, pp. 36-43.
- Maiden M. (2005). *Il problema del congiuntivo nell'insegnamento dell'italiano: verso una soluzione?*, in *Prospettive sull'italiano come Lingua Straniera*, a c. di A. L. Lepschy, A.R. Tamponi, Guerra, Perugia, pp. 109-117.
- Malinowska M. (1996), *Contenuti semantici del congiuntivo - ruolo della modalità e dei fattori extralinguistici nell'acquisizione della competenza comunicativa in italiano come lingua 2*, Viridis, Kraków.
- Mandelli M. (2006), *Quella 'e' a inizio frase... Scarti prosodici e interpuntivi tra elementi coordinati*, in *Parole frasi testi tra scritto e parlato*, a c. di A. Ferrari, in «Cenobio», 55/3, pp. 231-240.
- Manili P. (2015), *Il modo congiuntivo e l'italiano L2*, in «Gentes», 2, pp. 154-167.
- Marani T. (2016), *Didattica del congiuntivo e (neo)purismo linguistico*. Aggiornamenti, in «Rivista dell'Associazione dei docenti di italiano in Germania», 9, pp. 28-40.
- Marengo C. (1996), *Le parole dell'italiano. Lessico e dizionari*, Zanichelli, Bologna.
- Mauranen A. (2004), *Speech corpora in the classroom* in *Corpora and language learners*, a c. di G. Aston, S. Bernardini e D. Stewart, John Benjamins, Amsterdam/Philadelphia, pp.195-211.
- Mazzoleni M. (1992), *“Se lo sapevo non ci venivo”*: *l'imperfetto indicativo ipotetico nell'italiano contemporaneo*, in *Linee di tendenza dell'italiano contemporaneo*, a c. di Moretti B., Petrini D., Bianconi S., Bulzoni, Roma, pp. 171-190.
- McEnery A., Xiao R. (2005), *Help or help to: What do corpora have to say*, in «*English Studies*», 86/2, pp. 161-187.
- McEnery T. (2003), *Corpus Linguistics*, in *The Oxford Handbook of Computational Linguistics*, a c. di R. Mitkov, Oxford University Press, Oxford, pp. 448-463.

- McEnery T., Hardie A. (2012), *Corpus Linguistics: Method, Theory and Practice*, Cambridge University Press, Cambridge.
- McEnery T., Xiao R. (2011), *What corpora can offer in language teaching and learning*, in *Handbook of research in second language teaching and learning*, a c. di E. Hinkel, Routledge, New York, pp. 364-380.
- McEnery, T., Wilson, A. (2001[1996]), *Corpus Linguistics*, Edinburgh University Press, Edinburgh.
- Mengaldo P. V. (1994), *Storia della lingua italiana*, Il Novecento, Bologna.
- Meunier F., Granger S. (2008, a c. di), *Phraseology in Language Learning and Teaching*, John Benjamins, Amsterdam.
- Minciarelli, F., Comodi, A. (2005), *Sillabo per i cinque gradi del corso di lingua e cultura italiana per stranieri*, Guerra Edizioni, Perugia.
- Mioni A. M. (1983), *Italiano tendenziale: osservazioni su alcuni aspetti della standardizzazione*, in *Scritti linguistici in onore di Giovan Battista Pellegrini. Vol. I*, a c. di P. Benincà et al., Pacini, Pisa, pp. 495-517.
- Moneti A. (1988), *Sull'insegnamento del congiuntivo in italiano lingua seconda: proposta didattica*, in «Italice», 65/1, pp. 19-30.
- Moretti B., Petrini D., Bianconi S., (1992, a c. di), *Linee di tendenza dell'italiano contemporaneo*, Bulzoni, Roma.
- Moro A. (1998), “*Esserci*” e “*averci*”: *les clitiques sujets et une analyse en “petite proposition” pour “avoir”*, in “*Etre*” et “*avoir*”: *syntaxe, statique, typologie*, a c. di A. Rouveret, Presses Universitaires de Vincennes, Saint-Denis, pp. 155-170.
- Mortara Garavelli B. (2001), *Le parole e la giustizia*, Einaudi, Torino.
- Mudraya, O., Piao, S. S. L., Rayson P., Sharoff S., Babych B., Löfberg L. (2008), *Automatic extraction of translation equivalents of phrasal and light verbs in English and Russian*, in *Phraseology: An Interdisciplinary Perspective*, a c. di S. Granger, F. Meunier, John Benjamins, Amsterdam, pp. 293-309.
- Mukherjee J. (2002), *Korpuslinguistik und Englischunterricht: Eine Einführung*, Peter Lang, Frankfurt am Main.
- Nencioni G. (1976), *Parlato-parlato, parlato-scritto, parlato recitato*, in «Strumenti critici», 10, pp. 1-56, ora in G. Nencioni, *Di scritto e di parlato. Discorsi linguistici*, Zanichelli, Bologna, 1983, pp. 126-179.
- Nencioni G. (1987), *Costanza dell'antico nel parlato moderno*, in *Gli Italiani parlanti. Sondaggi nella lingua di oggi*, AA.VV., Accademia della Crusca, Firenze, pp. 7-25.
- Nesi H., Thompson P. (2014), *Using Sketch Engine with BAWE*. Disponibile alla pagina: <https://www.coventry.ac.uk/research/research-directories/current-projects/2015/british-academic-written-english-corpus-bawe/search-the-bawe-corpus/> (Consultato il 10/11/2017).
- Newman J. (2011), *Corpora and cognitive linguistics*, in «Brazilian Journal of Applied Linguistics», 11/2, pp. 521-559.
- Nunan D. (1989), *Understanding Language Classrooms. A Guide for Teacher-Initiated Research*, Prentice Hall, New York.
- Ondelli S., Pontrandolfo, G. (2014), *La negazione multipla nei testi giuridici: veramente non si può negare che sia un tratto caratteristico?*, in «Rivista internazionale di tecnica della traduzione», 16, pp.153–176.
- Ondelli S., Romanini F. (2018), *Norma interiorizzata e uso: un'indagine preliminare su parlanti italiani*, in «Italice Wratislaviensia», 9/1, 185-207.
- Ondelli S., Viale M. (2010), *L'assetto dell'italiano delle traduzioni in un corpus giornalistico. Aspetti qualitativi e quantitativi*, in «Rivista internazionale di tecnica della traduzione», 12, pp. 1–62.

- Orenha-Ottaiano A. (2012), *Collocations and the Design of Teaching Materials for Language Learners*, contributo presentato alla 10a conferenza internazionale su *Teaching and Language Corpora (TaLC10)*, Uniwersytet Warszawski, 11-14 luglio 2012. Disponibile online: <http://talc10.ils.uw.edu.pl/files/2011/09/ORENHA-OTTAIANO.pdf> (consultato il 20/12/2019).
- Palermo M. (1997), *L'espressione del pronome personale soggetto nella storia dell'italiano*, Bulzoni, Roma.
- Palermo M. (2011), *Soggetto*, in *Enciclopedia dell'italiano*, a c. di R. Simone, G. Berruto, P. D'Achille (2010-1). Disponibile alla pagina: [https://www.treccani.it/enciclopedia/soggetto_\(Enciclopedia-dell'Italiano\)/](https://www.treccani.it/enciclopedia/soggetto_(Enciclopedia-dell'Italiano)/) (consultato il 10/01/2019).
- Palermo, M. (2011), *Soggetto*, in *Enciclopedia dell'italiano*, a c. di R. Simone, G. Berruto, P. D'Achille (2010-1). Disponibile online: [http://www.treccani.it/enciclopedia/soggetto_\(Enciclopedia-dell'Italiano\)/](http://www.treccani.it/enciclopedia/soggetto_(Enciclopedia-dell'Italiano)/) (consultato il 31/05/2019).
- Palermo M. (2020), *Linguistica italiana*, Il Mulino, Bologna.
- Panunzi A., Cresti E., Gregori L. (2014), *RIDIRE. Corpus and Tools for the Acquisition of Italian L2*, in *Proceedings of the XVI EURALEX International Congress: The User in Focus*, a c. di A. Abel, C. Vettori, N. Ralli. Atti del convegno (Bolzano, 15-19 luglio 2014), EURAC research, Bolzano, 447-462.
- Patota G. (1990), *Sintassi e storia della lingua italiana: tipologia delle frasi interrogative*, Bulzoni, Roma.
- Pecman M. (2008), *Compilation, formalisation and presentation of bilingual phraseology: problems and possible solutions*, in *Phraseology in Language Learning and Teaching*, a c. di F. Meunier, S. Granger, John Benjamins, Amsterdam / Philadelphia, pp. 203-222.
- Pellegrini G. B. (1960), *Tra lingua e dialetto in Italia*, in «Studi mediolatini e volgari», 8, pp. 137-153.
- Pesce A. (2003), *Una lingua per l'infanzia*, in *Quaderni dell'Osservatorio Linguistico, vol. I*, a c. di F. Frasnèdi et al., Franco Angeli, Milano, pp.293-315.
- Picoral A., Staples S., Shin J., Swatek A. (2018), *Exploring variation and intertextuality in L2 undergraduate writing in English: Using the Corpus and Repository of Writing Online Platform for research and teaching*, workshop tenuto alla 13a conferenza internazionale su *Teaching and Language Corpora (TaLC13)*, Cambridge University, 18-21 luglio 2018. Disponibile online: <https://www.educ.cam.ac.uk/events/conferences/talc2018/programme/BookofAbstractsFinalversion.pdf> (consultato il 20/12/2019).
- Pieneman M. (1986), *L'effetto dell'insegnamento sugli orientamenti degli apprendenti nell'acquisizione di L2*, in *L'apprendimento spontaneo di una seconda lingua*, a c. di A. Giacalone Ramat, Il Mulino, Bologna, pp. 307-326.
- Pieneman M. (1998), *Language Processing and Second Language Development. Processability Theory*, Benjamins, Amsterdam/Philadelphia.
- Pienemann M. (2005), *Cross-Linguistic Aspects of Processability Theory*, John Benjamins, Amsterdam.
- Pistolesi E. (2015), *Diamesia: la nascita di una dimensione*, in *Parole, gesti, interpretazioni. Studi linguistici per Carla Bazzanella*, a c. di E. Pistolesi, R. Pugliese, B. Gili Fivela, Aracne, Roma, pp. 27-56.
- Poggi Salani T. (2010), *Italiano regionale*, in *Enciclopedia dell'italiano*, a c. di R. Simone, G. Berruto, P. D'Achille (2010-1). Disponibile alla pagina:

- [http://www.treccani.it/enciclopedia/italiano-regionale_\(Enciclopedia-dell%27Italiano\)/](http://www.treccani.it/enciclopedia/italiano-regionale_(Enciclopedia-dell%27Italiano)/) (consultato il 14/12/2017).
- Poggi Salani T. (2012 [1986]), *Per lo studio dell'italiano. Avviamento storico descrittivo*, Liviana, Padova.
- Pozzo G. (2008), *L'osservazione: uno strumento per conoscere cosa succede in classe*, in «Lend – Lingua e nuova Didattica», 37/1, pp. 15-28.
- Prandi M. (2002), *C'è un valore per il congiuntivo?*, in *Intorno al congiuntivo* a c. di L. Schena, M. Prandi, M. Mazzoleni, CLUEB, Bologna, pp. 29-44.
- Prandi M. (2020), *Le regole e le scelte. Grammatica italiana*, UTET, Torino.
- Radtke E. (1992), *Varietà dell'italiano*, in *La linguistica italiana degli anni 1976-1986*, a c. di A. A. Mioni, M. A. Cortelazzo, Bulzoni, Roma, pp. 59-74.
- Ramat P. (1993), *L'italiano lingua d'Europa*, in *Introduzione all'italiano contemporaneo. Vol. I*, a c. di A. A. Sobrero, pp. 3-39.
- Rastelli S. (2009), *Che cos'è la didattica acquisizionale*, Carocci, Roma.
- Rastelli S. (2010 a c. di), *Italiano di Cinesi, Italiano per Cinesi. Dalla prospettiva della didattica acquisizionale*, Guerra, Perugia.
- Rehm G., Santini M., Mehler A., Braslavski P., Gleim R., Stubbe A., Symonenko S., Tavasani M., Vidulin V. (2008), *Towards a reference corpus of web genres for the evaluation of genre identification systems*, in *Proceedings of the the sixth international conference on Language Resources and Evaluation (LREC 2008)*. Atti del convegno (Marrakech, 28-30 maggio 2008), ELRA - European Language Resources Association, Paris, pp. 351–358.
- Renzi L. (2012), *Come cambia la lingua. L'italiano in movimento*, Il Mulino, Bologna.
- Roggia C. E. (2006), *Costruzioni marcate tra scritto e parlato: la frase scissa*, in *Parole frasi testi tra scritto e parlato*, a c. di A. Ferrari, in «Cenobio», 55/3, pp. 222-230.
- Roggia C. E. (2009), *Le frasi scisse in italiano. Struttura informativa e funzioni discorsive*, Editions Slatkine, Genève.
- Römer U. (2004), *Comparing real and ideal language learner input: The use of an EFL textbook corpus in corpus linguistics and language teaching*, in *Corpora and language learners*, a c. di G. Aston, S. Bernardini e D. Stewart, John Benjamins, Amsterdam/Philadelphia, pp. 151-168.
- Römer U. (2008), *Corpora and language teaching*, in *Corpus Linguistics. An International Handbook Vol. 1*, a c. di A. Lüdeling e M. Kytö, Mouton de Gruyter, Berlin/New York, pp. 112-131.
- Rossi F. (2016), *Dislocazione a destra/a sinistra in due corpora di italiano scritto: tra grammaticalizzazione, ammiccamento e coesione*, in *La lingua variabile nei testi letterari, artistici e funzionali contemporanei*, a c. di G. Ruffino, M. Castiglione, Cesati, Firenze, pp. 847-859.
- Rossini Favretti R., Tamburini F., De Santis C. (2002), *CORIS/CODIS: A corpus of written Italian based on a defined and a dynamic model*, in *A Rainbow of Corpora: Corpus Linguistics and the Languages of the World*, a c. di A. Wilson, P. Rayson, T. McEnery, Lincom-Europa, München, pp. 27-38.
- Sabatini F. (1982), *La comunicazione orale, scritta e trasmessa: la diversità del mezzo, della lingua e delle funzioni*, in *Educazione linguistica nella scuola superiore. Sei argomenti per un curriculum*, a c. di A. M. Boccafurni, S. Serromani, Provincia di Roma e Consiglio Nazionale delle Ricerche, Roma, pp. 105-127.
- Sabatini F. (1985), *L'italiano dell'uso medio: una realtà tra le varietà linguistiche italiane*, in: *Geschprochenes Italienisch in Geschichte und Gegenwart*, a c. di G. Holtus, E. Radtke, Narr, Tübingen, pp. 154-184.

- Sabatini F. (1990), *Una lingua ritrovata: l'italiano parlato*, in «Studi latini e italiani», 4, pp. 215-234.
- Sabatini F. (1997), *Pause e congiunzioni nel testo. Quel "ma" all'inizio di frase...*, in *Norma e lingua in Italia: alcune riflessioni tra passato e presente*, a c. di I. Bonomi, Istituto Lombardo di Scienze e Lettere, Milano, pp. 113–146.
- Sabatini F. (1997), *Prove per l'italiano 'trasmesso' (e auspici di un parlato serio semplice)*, in *Gli italiani trasmessi: la radio. Atti del Convegno (Firenze, Villa Medicea di Castello, 13-14 maggio 1994)*, Accademia della Crusca, Firenze, pp. 11-30.
- Sabatini F. (2011), *L'Italiano nel mondo moderno, Saggi scelti dal 1968 al 2009, vol. II*, a c. di V. Coletti, R. Coluccia, P. D'Achille, N. De Blasi, D. Proietti, bibliografia degli scritti a c. di R. Cimaglia, Liguori, Napoli.
- Sabatini, F. (2004), *Risposta al quesito di Marina Hubacech sull'uso del verbo venire come ausiliare solo nei tempi semplici*, in «La Crusca per voi», 29, pp. 8-9. Disponibile online: www.accademiadellacrusca.it/it/lingua-italiana/consulenza-linguistica/domande-risposte/motivo-per-verbo-venire-utilizzato-ausiliare (consultato il 05/01/2019).
- Sabatini, F. (2016), *Lezione di italiano*, Mondadori, Milano.
- Salvi G. (1988), *La frase semplice*, in *Grande grammatica italiana di consultazione. Vol. I: La frase. I sintagmi nominale e preposizionale*, a c. di L. Renzi, Salvi G., Cardinaletti A., Il Mulino, Bologna, pp. 37-127.
- Sansò A. (2003), *Degrees of event elaboration. Passive construction in Italian and Spanish*, Franco Angeli, Milano.
- Santini M. (2005), *Genres in formation? An exploratory study of web pages using cluster analysis*, in *Proceedings of the 8th Annual Colloquium for the UK Special Interest Group for Computational Linguistics (CLUK05)*. Atti del convegno (Manchester, 11 gennaio 2005). Disponibile online alla pagina: <https://www.sheffield.ac.uk/dcs/research/groups/nlp> (consultato il 20/12/2019).
- Santini M., Mehler A., Sharoff S. (2010), *Riding the Rough Waves of Genre on the Web. Concepts and Research Questions*, in *Genres on the Web: Computational Models and Empirical Studies*, a c. di A. Mehler, S. Sharoff, M. Santini, Springer, Dordrecht, pp. 3–33.
- Santipolo M. (2014), *L'impiego dei materiali autentici per lo sviluppo della competenza sociolinguistica: riflessioni teoriche e spunti operativi*, «Cultura & Comunicazione», 5, pp.15-22.
- Schena L., Prandi M., Mazzoleni M. (2002, a c. di), *Intorno al congiuntivo*, CLUEB, Bologna.
- Scherer K. R. (1984), *On the nature and function of emotion: a component process approach*, in *Approaches to emotion*, a c. di K.R. Scherer, P. Ekman, Erlbaum, Hillsdale, pp. 293-317.
- Scherer K. R. (2009), *The dynamic architecture of emotion: Evidence for the component process model*, «Cognition and Emotion», 23/7, pp.1307-1351.
- Scherer K. R., Schorr A., Johnstone T. (2001, a c. di), *Appraisal processes in emotion: theory, methods, research*, Oxford University Press, New York.
- Schneider S. (1999), *Il congiuntivo tra modalità e subordinazione: uno studio sull'italiano parlato*, Carocci, Roma.
- Schumann J. H. (1999), *The Neurobiology of Affect in Language*, Blackwell, Oxford.
- Schumann J. H., Crowell S. E., Jones N. E., Lee, N., Schuchert, S. A., Wood, L. A. (2004), *The neurobiology of learning. Perspectives from second language acquisition*, Lawrence Erlbaum, Mahwah, (NJ).

- Selinker L. (1972), *Interlanguage*, «International Review of Applied Linguistics in Language Teaching», 10, pp. 209-241.
- Seracini F. (2018), *Phraseology in specialised translation learning: a corpus-based study*, contributo presentato alla 13a conferenza internazionale su *Teaching and Language Corpora* (TaLC13), Cambridge University, 18-21 luglio 2018. Disponibile online: <https://www.educ.cam.ac.uk/events/conferences/talc2018/programme/BookofAbstractsFinalversion.pdf> (consultato il 20/12/2019).
- Serianni L. (1986), *Il problema della norma linguistica dell'italiano*, in «Gli Annali della Università per Stranieri [di Perugia]», 7, pp. 47-69.
- Serianni L. (1988), *Grammatica italiana. Italiano comune e lingua letteraria. Suoni, forme, costrutti*, con la collaborazione di A. Castelvechi, UTET, Torino.
- Serianni L. (1988), *Grammatica italiana. Italiano comune e lingua letteraria. Suoni, forme, costrutti*, con la collaborazione di A. Castelvechi, UTET, Torino.
- Serianni L. (1989a), *Le varianti fonomorfologiche dei "Promessi Sposi" 1840 nel quadro dell'italiano ottocentesco*, in *Saggi di storia linguistica italiana*, L. Serianni, Morano, Napoli, pp. 141-213.
- Serianni L. (2006), *Prima lezione di grammatica*, Laterza, Roma/Bari.
- Serianni L., Trifone P. (1994, a c. di), *Storia della lingua italiana. Vol. II: Scritto e parlato*, Einaudi, Torino.
- Sgroi S.C. (2013), *Dove va il congiuntivo? Ovvero il congiuntivo da nove punti di vista*, UTET, Torino.
- Sharoff S. (2006), *Creating General-Purpose Corpora Using Automated Search Engine Queries*, in *Wacky! Working Papers on the Web as Corpus*, a c. di M. Baroni, S. Bernardini, Gedit, Bologna editors, pp. 63-98.
- Sharoff S. (2010), *Analysing similarities and differences between corpora*, in *7th Language Technologies Conference, Jožef Stefan Institute, Ljubljana*, pp. 5-11.
- Simone R. (1991), *Why linguists need variation. Reflections on Italian in "troubled contexts"*, in «Rivista di linguistica», 3/2, pp. 407-421.
- Simone R. (1993), *Stabilità e instabilità nei caratteri originali dell'italiano*, in *Introduzione all'italiano contemporaneo. Vol. I*, a c. di A. A. Sobrero, pp. 41-139.
- Simone R. (1993), *Stabilità e instabilità nei caratteri originali dell'italiano*, in *Introduzione all'italiano contemporaneo. Vol. I: Struttura e variazioni*, a c. di A.A. Sobrero, Laterza, Roma/Bari, pp. 41-100.
- Simone R., Berruto G., D'Achille P. (2010-1, a c. di), *Enciclopedia dell'italiano*, Treccani, Roma. (<http://www.treccani.it/>).
- Sinclair J. McH. (1996), *Preliminary recommendations on Corpus Typology, EAGLES Document EAG-TCWG-CTYP/P, Version of May, 1996*. Disponibile alla pagina: <http://www.ilc.cnr.it/EAGLES96/corpus/corpus.html> (consultato il 15/12/2019).
- Sinclair J. McH. (1997), *Corpus evidence in language description*, in *Teaching and language Corpora*, a c. di A. Wichmann, S. Fligelstone, T. McEnery e G. Knowles Longman, London, pp. 27-39.
- Sinclair J. McH. (2004, a c. di), *How to Use Corpora in Language Teaching*, John Benjamins, Amsterdam.
- Sinclair J. McH. (2005), *Developing Linguistic Corpora: a Guide to Good Practice. Corpus and Text — Basic Principles*. Disponibile alla pagina: <https://ota.ox.ac.uk/documents/creating/dlc/chapter1.htm> (consultato il 15/12/2019).
- Sinclair, J. McH. (1991), *Corpus, concordance, collocation*, Oxford University Press, Oxford.

- Ślapek D. (2017), *Argomenti grammaticali nei manuali e nei certificati di italiano LS*, «Rassegna Italiana di Linguistica Applicata», 49/1, pp. 109-127.
- Sobrero A. A. (1992a), *L'italiano di oggi*, Istituto della Enciclopedia Italiana Treccani, Roma.
- Sobrero A. A. (1992b), *Forme e norme*, in «Italiano e oltre», 7, p. 153.
- Sobrero A. A. (1993, a c. di), *Introduzione all'italiano contemporaneo. Vol. I: le strutture; Vol. II: la variazione e gli usi*, Laterza, Roma-Bari.
- Sornicola R. (1981), *Sul Parlato*, Bologna, Il Mulino.
- Sornicola R. (1981), *Sul parlato*, Il Mulino, Bologna.
- Sosnowski R. (2010), *Errori dovuti a interferenze negli scritti degli apprendenti polacchi (livello intermedio e avanzato). Ricerca corpus based*, in «Romanica Cracoviensia», 10/1, pp. 135-156.
- Spina S. (2010a), *The dictionary of Italian Collocations: Design and intergration in an Online Learning Environment*, in *Proceedings of the Seventh conference on International Language Resources and Evaluation (LREC'10)*. Atti del convegno (Malta, 17-23 maggio 2010), ELRA - European Language Resources Association, Paris, pp. 3202-3208.
- Spina S. (2010b), *AIWL: una lista di frequenza dell'italiano accademico*, in *Statistical Analysis of Textual Data, Proceedings of the 10th JADT Conference*, a c. di S. Bolasco, I. Chiari, L. Giuliano. Atti del convegno (Università La Sapienza, 9-11 giugno 2010), Editrice universitaria LED, Roma, pp. 1317-1325.
- Spina S. (2014), *Il Perugia Corpus: una risorsa di riferimento per l'italiano. Composizione, annotazione e valutazione in Proceedings of the First Italian Conference on Computational Linguistics (CLiC-it 2014)*, a c. di R. Basili, A. Lenci, B. Magnini, 1, Pisa University Press, Pisa, pp. 354-359.
- Spinelli B., Parizzi F. (2010 a c. di). *Profilo della lingua italiana. Livelli di riferimento del QCER A1, A2, B1, B2*, La Nuova Italia, Firenze.
- Stefanowitsch A., Gries S. T. (2006, a c. di), *Corpus-based approaches to metaphor and metonymy*, Mouton de Gruyter, Berlin-New York.
- Stubbs M. (2001), *Texts, corpora, and problems of interpretation: A response to Widdowson*, «Applied linguistics», 22/2, pp. 149-172.
- Svartvik, J. (1992). *Corpus linguistics comes of age*, in *Directions in corpus linguistics. Proceedings of Nobel Symposium 82*. (Stockholm 4-8 agosto 1991), a c. di J. Svartvik, Mouton de Gruyter, Berlin-New York.
- Tamburini F. (2000), *Annotazione grammaticale e lemmatizzazione di corpora in italiano*, in *Linguistica e informatica: multimedialità, corpora e percorsi di apprendimento*, a c. di R. Rossini, Bulzoni, Roma, pp. 57-73.
- Tavoni M. (2002), *Caratteristiche dell'italiano contemporaneo e insegnamento della scrittura*, in *Manuale dell'italiano professionale. Teoria e didattica*, a c. di F. Bruni, T. Raso, Zanichelli, Bologna, pp. 139-152.
- Taylor D. (1994), *Inauthentic authenticity or authentic in authenticity?*, «TESL-EJ», 1/2, pp. 1-12. Disponibile online: <https://www.tesl-ej.org/wordpress/issues/volume1/ej02/ej02a1/> (consultato il 24/11/2019).
- Tesi R., 2005, *Storia dell'italiano. La lingua moderna e contemporanea*, Zanichelli, Bologna.
- Titone R. (1976), *Psicodidattica*, La Scuola, Brescia.
- Tognini-Bonelli E. (2001), *Corpus Linguistics at Work*, John Benjamins, Amsterdam.
- Tomlin R. S. (1986), *Basic Word Order. Functional Principles*, Croom Helm, Londra/Sydney/Wolfeboro.

- Tono Y. (2018), *CEFR-Jx27: Developing corpus- and CEFR-based pedagogical resources and e-learning systems for 27 languages*, contributo presentato alla 13a conferenza internazionale su *Teaching and Language Corpora* (TaLC13), Cambridge University, 18-21 luglio 2018. Disponibile online: <https://www.educ.cam.ac.uk/events/conferences/talc2018/programme/BookofAbstractsFinalversion.pdf> (consultato il 20/12/2019).
- Trévisiol-Okamura P., Komur-Thilloy G. (2011, a c. di), *Discours, acquisition et didactique des langues: Les termes d'un dialogue*, Orizons, Paris.
- Vališová P. (2014), *Teaching Czech Verbs on Elementary Level: Comparing Textbook Corpus and National Corpus Data*, contributo presentato alla 11a conferenza internazionale su *Teaching and Language Corpora* (TaLC11), Lancaster University, 20-23 luglio 2014. Disponibile online: <http://ucrel.lancs.ac.uk/talc2014/doc/TALC2014-abstract-book.pdf> (consultato il 20/12/2019).
- Vanelli L. (1999), "Punti di crisi" nell'italiano contemporaneo, in *Insegnare italiano nella scuola del 2000*, a c. di U. Cardinale, Unipress, Padova, pp. 99-121.
- Vedovelli M., Villarini A. (2003), *Dalla linguistica acquisizionale alla didattica acquisizionale: le sequenze sintattiche nei materiali per l'italiano L2 destinati agli immigrati stranieri* in *Verso l'italiano. Percorsi e strategie di acquisizione*, a c. di A. Giacalone Ramat, Carocci, Roma, pp. 270-304.
- Vedovelli M. (2003), *Note sulla glottodidattica italiana oggi: problemi e prospettive*, «Studi Italiani di Linguistica Teorica e Applicata», 18/2, pp. 173-197.
- Véronique D. (2000 a c. di), *Didactique des langues étrangères et recherches sur l'acquisition*, in «Etudes de Linguistique Appliquée, 120 (numero monografico)», Trévisiol-Okamura, Komur-Thilloy (2011).
- Véronique D. (2005), *Les interrelations entre la recherche sur l'acquisition du français langue étrangère et la didactique du français langue étrangère*, in «Acquisition et Interaction en Langue Etrangère/AILE», 23, pp. 9-41.
- Viganò P. B. (2011), *I corpora e il loro sfruttamento in didattica*, «Italiano LinguaDue», 2, 115-128.
- Voghera, M. (1992), *Sintassi e intonazione nell'italiano parlato*, Il Mulino, Bologna.
- Wandruszka U. (1991), *Frase subordinate al congiuntivo*, in *Grande grammatica italiana di consultazione. Vol. II: I sintagmi verbale, aggettivale, avverbiale. La subordinazione.*, a c. di L. Renzi, Salvi G., Cardinaletti A., Il Mulino, Bologna, pp. 415-481
- Wible D. (2008), *Multiword expressions and the digital turn*, in *Phraseology in Language Learning and Teaching*, a c. di F. Meunier, S. Granger, John Benjamins, Amsterdam / Philadelphia, pp. 163-181.
- Widdowson H. G. (1991), *The description and prescription of language*, in *Georgetown University round table on language and linguistics. Linguistics and language pedagogy: The state of the art*, a c. di J. E. Alatis, Georgetown University Press, Washington D.C., pp. 11-24.
- Widdowson H. G. (2000), *On the limitations of linguistics applied*, «Applied linguistics», 21/1, pp. 3-25.
- Wilkins D. (1976), *Notional syllabus*, Oxford University Press, Oxford.
- Whittle A. (2015), *Focalizzare la forma: sviluppo della competenza linguistica nella classe multilingue della scuola primaria* in «Italiano LinguaDue», 2, pp. 13-35.
- Wunderlich C. (2018), *Creating a subject-specific corpus and academic word list for business informatics*, contributo presentato alla 13a conferenza internazionale su *Teaching and Language Corpora* (TaLC13), Cambridge University, 18-21 luglio

2018. Disponibile online:
<https://www.educ.cam.ac.uk/events/conferences/talc2018/programme/BookofAbstractsFinalversion.pdf> (consultato il 20/12/2019).
- Zamora P., Alessandro A. (2011), *La didáctica del italiano coloquial: viejos recelos, nuevos alicientes*, in «Revista electrónica de estudios filológicos», 21. Disponibile online: <https://www.um.es/tonosdigital/znum21/secciones/estudios-30-italiano.htm> (consultato il 10/10/2019).
- Zamora P., Alessandro, A., Hernández González, M. B. (2010), *L'italiano, a me mi piace. Corso per l'apprendimento dell'italiano colloquiale come LS/L2*, Editum, Murcia.
- Zanchetta E., Baroni, M. (2005), *Morph-it! A free corpus-based morphological resource for the Italian language*, in *Proceedings from the Corpus Linguistics Conference 1* (1). Atti del convegno (University of Birmingham, luglio 2005), Conference e-journal disponibile alla pagina: <https://www.birmingham.ac.uk/research/activity/corpus/publications/conference-archives/2005-conf-e-journal.aspx> (consultato il 20/11/2018).
- Zingaro A. (2018), *La "vita spericolata" della sintassi: il che polivalente nella canzone italiana tra gli anni '70 e '90 del Novecento*, in «ITALICA WRATISLAVIENSIA», 9/2, pp. 305-325.

Grammatiche

- Antonelli G., Picchiorri E. (2016), *L'italiano, gli italiani*, Einaudi scuola, Torino.
- Battaglia, S. - Pernicone, V. (1980 [1951]), *La grammatica italiana*, Loescher, Torino.
- Dardano, M. - Trifone, P. (1985), *La lingua italiana*, Zanichelli, Bologna.
- Fogarasi, M. (1983 [1969]), *Grammatica italiana del Novecento*, Bulzoni, Roma.
- Fogliato S. (2015), *L'italiano competente*, Loescher, Torino.
- Renzi L., Salvi G., Cardinaletti A. (2001-2), *Grande grammatica italiana di consultazione, Vol. I-III*, Il Mulino, Bologna.
- Salvi G., Vanelli L. (2004), *Nuova grammatica italiana*, Il Mulino, Bologna.
- Schmitt-Jensen, J. (1970), *Subjonctif et hypotaxe en italien: une esquisse de la syntaxe du subjonctif dans les propositions subordonnées en italien contemporain*, Odense University Press, Odense.
- Sensini, M. (1990), *La grammatica della lingua italiana*, Mondadori, Milano.
- Serianni, L. (1989b), *Grammatica italiana*, con la collaborazione di Castelvechi A., UTET, Torino.
- Serianni L., Della Valle V., Patota G. (2016), *Italiano plurale*, Bruno Mondadori, Milano.
- Sorrento, L. (1951), *Sintassi romanza. Ricerche e prospettive*, Istituto Editoriale Cisalpino Varese, Milano.

Sitografia di riferimento

CORPORA E PAGINE DI ISTITUTI DI RICERCA E STRUMENTI DI CONSULTAZIONE ESSI CORRELATI

Araneum ucts.uniba.sk, kontext.korpus.cz, www.sketchengine.co.uk. (consultati il 29/01/2018)

Aranea Corpora

http://ucts.uniba.sk/aranea_about/,

<http://ucts.uniba.sk/> (consultati il 29/01/2018)

Araneum Italicum http://ucts.uniba.sk/aranea_about/_italicum.html (consultato il 29/01/2018)

Athenaeum Corpus
<http://www.bmanuel.org/projects/at-HOME.html>,
http://www.corpora.unito.it/ateneo/ateneo_.php (consultato il 29/01/2018)

CHILDES <http://childes.psy.cmu.edu/> (consultato il 29/01/2018)

CHILDES-IT <http://childes.psy.cmu.edu/data/Romance/Italian/> (consultato il 29/01/2018)

CLIPS <http://www.clips.unina.it/it/> (consultato il 29/01/2018)

COLFIS
<https://www.istc.cnr.it/grouppage/colfis> (consultato il 06/12/2020)
<http://linguistica.sns.it/esploracolfis/home.htm> (consultato il 29/01/2018)
<http://www.ge.ilc.cnr.it/strumenti.php> (consultato il 29/01/2018)

CoLiTec <https://corpora.dipintra.it/> (consultato il 10/03/2018)

CorDIC <http://corporadidattici.lablita.it/> (consultato il 29/01/2018)

CORIS <http://corpora.dslo.unibo.it/TCORIS/> (consultato il 29/01/2018)

Corpus querying <https://www.sketchengine.eu/documentation/corpus-querying/> (consultato il 13/06/2018).

CQL Basics <https://www.sketchengine.eu/documentation/cql-basics/> (consultato il 13/06/2018).

Creative Commons <http://creativecommons.org/licenses/> (consultato il 28/02/2018)

e-allora <http://www.e-allora.net/> (consultato il 10/03/2018)

Esplora CoLFIS <http://esploracolfis.sns.it/EsploraCoLFIS/> (consultato il 16/03/2018)

ISST-TANL tagset http://www.corpusitaliano.it/static/documents/POS_ISST-TANL-tagset-web.pdf (consultato il 13/06/2018)

ITenTen <https://www.sketchengine.co.uk/itTenTen-italian-corpus/> (consultato il 29/01/2018)

IRIS Unito <https://iris.unito.it/handle/2318/120913#.WqaOuExFx1s> (consultato il 16/03/2018)

LABLITA <http://lablita.dit.unifi.it/> (consultato il 29/01/2018)

la Repubblica https://corpora.dipintra.it/public/run.cgi/first_form (consultato il 29/01/2018)

LexALP
http://lexalp.eurac.edu/projects/corpus_en.htm,
http://lexalpapps.eurac.edu:8080/htdocs2/lexalp/corp_lexalp/search_corp.php
(consultato il 29/01/2018)

Laboratorio di linguistica - Scuola Normale Superiore
<http://linguistica.sns.it/CoLFIS/Testi/libri/libri.htm>
(consultato il 29/05/2018).

Morph-it <http://dev.sslmit.unibo.it/linguistics/morph-it.php> (consultato il 29/01/2018)

NoSketch Engine <http://nlp.fi.muni.cz/trac/noske> (consultato il 29/01/2018)

NUNC
<http://www.bmanuel.org/projects/ng-HOME.html>
http://www.corpora.unito.it/index_nunc.php
(consultato il 29/01/2018)

PAISÀ <http://www.corpusitaliano.it/it/> (consultato il 29/01/2018)

Perugia Corpus <https://www.unistrapg.it/cqpweb/> (consultato il 29/01/2018)

Repubblica Scuola <http://scuola.repubblica.it/> (consultato il 28/05/2018)

RIDIRE <http://www.ridire.it/it.drwolf.ridire/home.seam> (consultato il 29/01/2018)

RIDIRE-CPI <https://github.com/lablita/ridire-cpi> (consultato il 29/01/2018)

Sketch Engine <http://www.sketchengine.co.uk/> (consultato il 29/01/2018)

SMS MS http://www.e-allora.net/SMS/ms_index.php (consultato il 29/01/2018)
SpiderLing <http://nlp.fi.muni.cz/trac/spiderling> (consultato il 29/01/2018)
TenTen <https://www.sketchengine.co.uk/documentation/tenten-corpora/> (consultato il 29/01/2018)
TreeTagger <http://www.cis.uni-muenchen.de/~schmid/tools/TreeTagger/> (consultato il 29/01/2018)
WaCky <http://wacky.sslmit.unibo.it/doku.php> (consultato il 29/01/2018)
Wikipedia Extractor http://wiki.apertium.org/wiki/Wikipedia_Extractor (consultato il 29/01/2018)

PIATTAFORME DI E-LEARNING E VERIFICA DELL'APPRENDIMENTO

A-tutor <https://atutor.ca> (consultato il 05/01/2019).
Blackboard <https://www.blackboard.com/it-it> (consultato il 05/01/2019).
Chamilo <https://chamilo.org/it/> (consultato il 05/01/2019).
Claroline <https://claroline.net/> (consultato il 05/01/2019).
Docebo <https://www.docebo.com/> (consultato il 05/01/2019).
Dokeos <https://www.dokeos.com/it/> (consultato il 05/01/2019).
E-Front <https://www.efrontlearning.com/> (consultato il 05/01/2019).
Edmodo www.edmodo.com (consultato il 05/01/2019)
Fidemia www.fidemia.com (consultato il 05/01/2019)
Ilias <https://www.ilias.de/> (consultato il 05/01/2019).
Moodle <http://moodle.org> (consultato il 05/01/2019)
Sakai <https://www.sakailms.org/> (consultato il 05/01/2019).
Schoology <https://www.schoology.com/> (consultato il 05/01/2019).
Socrative <https://socrative.com> (consultato il 13/01/2019)
Wizer.me www.wizer.me (consultato il 05/01/2019)

Indice dei nomi

- Acquaviva, P.; 60; 265
Aijmer, K.; 7; 79; 255; 265
Alessandro, A.; 177; 179; 285
Alhujaylan, H.; 68; 80; 265
Alisova, T.; 40; 265
Ammon, U.; 18; 19; 265
Amor, S.; 87; 265
Andorno, C.; 150; 265; 274
Antonelli, G.; 29; 33; 42; 60; 265; 274; 285
Arcangeli, M.; 183; 265
Arruffoli, A. M.; 266
Asor Rosa, A.; 265; 272
Aston, G.; 78; 79; 88; 265; 267; 277; 280
Babych, S.; 68; 80; 265; 278
Baker, M.; 72; 93; 265; 266
Balboni, P. E.; 4; 8; 147; 149; 151; 152; 155; 185; 199; 266
Bambini, V.; 97; 266
Bandini, A.; 266
Banfi, E.; 149; 150; 266
Barbera, M.; 81; 94; 95; 105; 106; 266
Barni, M.; 183; 184; 266
Baroni, M.; 7; 65; 66; 68; 73; 81; 101; 102; 103; 108; 109; 112; 266; 267; 282; 285; 293
Bazzanella, C.; 76; 267; 279
Begotti, P.; 87; 267
Bembo, P.; 11; 12; 32; 56; 267
Benincà, P.; 35; 36; 37; 267
Benko, V.; 93; 95; 267; 292
Benucci, A.; 183; 184; 185; 267
Berisso, M.; 42; 267
Bernardini, S.; 78; 79; 88; 101; 102; 265; 266; 267; 277; 280; 282
Bernini, G.; 150; 266
Berretta, M.; 36; 37; 50; 58; 150; 267; 268
Berruto, G.; 7; 11; 14; 15; 16; 17; 19; 20; 21; 23; 24; 25; 26; 29; 35; 36; 37; 41; 58; 59; 60; 61; 63; 64; 136; 177; 254; 266; 268; 271; 273; 274; 276; 279; 282
Bettoni, C.; 151; 268; 269
Bianconi, S.; 26; 277; 278
Biber, D.; 70; 74; 269
Bolasco, S.; 93; 269; 276; 283
Bolwler, L.; 93; 269
Bonomi, I.; 37; 42; 269; 281
Bonvino, E.; 86; 269; 271
Borghetti, C.; 113; 269; 277
Bortolini, U.; 98; 104; 269
Bozzone Costa, R.; 274
Braslavski, P.; 280
Breen, M. P.; 87; 269
Bresnan, J.; 151; 270
Brezina, V.; 68; 81; 269
Brown, G.; 276
Brunello, M.; 111; 113; 269; 277
Brunet-Manquat, F.; 277
Bruni, F.; 20; 270; 283
Buchholz, S.; 113; 270
Budassi, M.; 120; 134; 135; 136; 137; 165; 176; 270
Budinich, L.; 109; 270
Burani, C.; 276
Calleri, D.; 150; 270
Cardillo, G.; 183; 265; 270
Cardinale, U.; 62; 270; 284
Castagnoli, S.; 111; 113; 269; 277
Castellani, A.; 7; 30; 56; 270
Cerruti, M.; 16; 268
Charkova, D.; 68; 81; 270
Charkova, K.; 68; 81; 270
Chini, M.; 81; 149; 150; 270; 271; 274
Chiocchetti, E.; 277
Christ, O.; 80; 116; 271
Cignetti, L.; 36; 37; 49; 181; 271
Ciliberti, A.; 183; 271
Comastri, F.; 267
Comodi, A.; 87; 183; 184; 185; 271; 278
Conrad, S.; 70; 269
Coonan, C. M.; 8; 153; 154; 271
Cordin, P.; 32; 33; 270; 271
Corino, E.; 7; 68; 82; 83; 255; 266; 271
Cortelazzo, M.; 7; 20; 30; 40; 56; 60; 62; 271; 272; 280
Coseriu, E.; 11; 13; 30; 272
Cosme, C.; 68; 79; 272
Coxhead, A.; 68; 80; 272
Cresti, E.; 37; 81; 109; 272; 279
Curado-Fuentes, A.; 68; 81; 272
Curry, N.; 68; 81; 272
D'Achille, P.; 3; 6; 7; 11; 14; 20; 21; 26; 29; 30; 40; 50; 56; 62; 266; 268; 271; 272; 273; 274; 276; 279; 281; 282
D'Angelo, K.; 265
Daloiso, M.; 155; 185; 272
Danesi, M.; 8; 151; 153; 272
Dardano, M.; 6; 7; 11; 30; 42; 60; 272; 285
De Cesare, A. M.; 36; 272
de Haan, P.; 93; 273
De Mauro, T.; 6; 11; 20; 24; 84; 85; 97; 112; 273
De Roberto, E.; 47; 273
De Santis, C.; 42; 99; 100; 273; 280
Della Putta, P.; 149; 273

Dell'Orletta, F.; 277
 Di Biase, B.; 151; 268; 269
 Diadori, P.; 5; 91; 147; 153; 273
 Dittmann, H.; 277
 Doughty, C.; 149; 273
 Drescher, V.; 87; 274
 Dziwirek, K.; 76; 276
 Egerland, V.; 47; 273
 Ellis, R.; 149; 273
 Evans, V.; 76; 273
 Evert, S.; 8; 117; 256; 273
 Ferraresi, A.; 267
 Ferrari, S.; 36; 37; 265; 271; 272; 273; 277;
 280
 Ferraris, P.; 150; 270; 273
 Fiorentino, G.; 40; 273
 Firth, J. R.; 87; 274
 Francis, G.; 69; 265; 275
 Freddi, G.; 151; 152; 274
 Frison, L.; 35; 50; 267
 Frisson; 36
 Frosini, G.; 6; 11; 274
 Fusi, C.; 266
 Gablasova, D.; 68; 81; 269
 Galli de' Paratesi, N.; 12; 19; 25; 274
 Gatta, F.; 37; 274
 Geeraerts, D.; 76; 274
 Ghezzi, C.; 274
 Giacalone Ramat, A.; 81; 150; 265; 270;
 274; 284
 Gilquin, G.; 68; 79; 272
 Giorgi, A.; 43; 274
 Giugni, S.; 265
 Giuliano, L.; 93; 274; 283
 Gleim, R.; 280
 Glisan, E.; 87; 274
 Goodman, O.; 68; 81; 272
 Gossen, C. T.; 35; 274
 Grandi, N.; 61; 129; 274
 Granger, S.; 7; 79; 255; 272; 274; 278; 279;
 284
 Grassi, R.; 149; 274
 Greco, N.; 265
 Grefenstette, G.; 7; 65; 66; 67; 74; 75; 276
 Gregori, L.; 81; 98; 279
 Gries, S. T.; 76; 77; 274; 275; 283
 Gualdo, R.; 62; 275
 Guidetti, M. G.; 7; 68; 82; 255; 275
 Hadley, G.; 68; 81; 275
 Hadley, H.; 68; 81; 275
 Hardie, A.; 7; 67; 68; 69; 71; 72; 73; 75; 92;
 94; 266; 278
 Harrich, K.; 87; 275
 Haywood, S.; 80; 276
 Hernández González, M. B.; 177; 285
 Hilpert, M.; 76; 275
 Hinkel, E.; 7; 79; 149; 255; 275; 278
 Hoffmann, T.; 77; 275
 Holtus, G.; 17; 265; 268; 271; 275; 280
 Housen, A.; 149; 275
 Hunston, S.; 69; 72; 275
 Ivanovska-Naskova, R.; 7; 68; 81; 82; 84;
 85; 255; 275; 276
 Jakubíček, M.; 101; 276
 Ježek, E.; 132; 276
 Johns, T. F.; 77; 276
 Johnstone, T.; 281
 Jones, M.; 80; 276; 281
 Kennedy, C.; 7; 68; 82; 255; 276
 Kilgarriff, A.; 7; 65; 66; 67; 74; 75; 276
 Koch, P.; 18; 276
 Koffka, K.; 147; 276
 Köhler, W.; 147; 276
 Komur-Thillooy, G.; 149; 284
 Kovář, V.; 276
 Krashen, S. D.; 8; 87; 147; 148; 155; 276
 Kytö, M.; 67; 277; 280
 La Rocca, G.; 93; 274
 Lambrecht, K.; 36; 276
 Laudanna, A.; 96; 97; 276
 Lavinio, C.; 14; 276
 Leech, G.; 74; 78; 79; 108; 276
 Lenci, A.; 277; 283
 Lenzi, G.; 7; 68; 82; 255; 275
 Lepschy, A. L.; 42; 46; 181; 276; 277
 Lepschy, G. C.; 42; 46; 181; 276
 Levelt, W. J. M.; 151; 276
 Lewandowska-Tomaszczyk, B.; 76; 276
 Liscia, R.; 143; 276
 Lo Duca, M. G.; 150; 183; 184; 185; 276;
 277
 Löfberg, L.; 278
 Loporcaro, M.; 47; 277
 Lorenzetti, L.; 7; 13; 14; 17; 18; 19; 20; 21;
 22; 23; 26; 277
 Lucarelli, S.; 266
 Lucatorno, A.; 265
 Lüdeling, A.; 67; 277; 280
 Lyding, V.; 7; 68; 84; 104; 111; 112; 113;
 255; 277
 Maiden, M.; 42; 277
 Malinowska, M.; 42; 277
 Mancini, F.; 97; 273
 Mandelli, M.; 41; 277
 Manili, P.; 42; 277
 Marani, T.; 41; 42; 277
 Marconi, L.; 276
 Marellò, C.; 7; 68; 83; 93; 95; 105; 255; 266;
 268; 271; 277
 Marín Cervantes, I.; 68; 81; 269

Marsi, E.; 113; 270
 Mauranen, A.; 88; 277
 Mauroni, E.; 42; 269
 Mazzoleni, M.; 41; 42; 137; 267; 277; 280; 281
 McEnergy, A.; 7; 65; 67; 68; 69; 71; 72; 73; 75; 77; 78; 79; 92; 93; 94; 266; 276; 277; 278; 280; 282
 Mehler, A.; 280; 281
 Mengaldo, P. V.; 23; 278
 Meunier, F.; 7; 79; 255; 272; 274; 278; 279; 284
 Miceli, T.; 7; 68; 82; 255; 276
 Minciarelli, F.; 183; 184; 185; 278
 Mioni, A. M.; 11; 13; 14; 15; 278; 280
 Moneti, A.; 42; 278
 Moretti, B.; 26; 268; 277; 278
 Moro, A.; 51; 275; 278
 Mortara Garavelli, B.; 36; 268; 278
 Mudraya, O.; 68; 79; 278
 Mukherjee, J.; 93; 278
 Nencioni, G.; 7; 14; 29; 56; 99; 266; 278
 Nesi, H.; 8; 118; 256; 278
 Newman, J.; 76; 278
 Nicolas, L.; 277
 Nunan, D.; 8; 153; 278
 Ondelli, S.; 7; 29; 35; 37; 42; 54; 60; 61; 63; 272; 278
 Orenha-Ottaiano, A.; 68; 80; 279
 Palermo, M.; 5; 6; 11; 48; 91; 153; 273; 279
 Panunzi, A.; 81; 98; 108; 109; 272; 279
 Parizzi, F.; 183; 185; 283
 Patota, G.; 33; 48; 279; 285
 Pearson, J.; 93; 269
 Pecman, M.; 68; 80; 279
 Pellegrini, G. B.; 22; 278; 279
 Pesce, A.; 42; 279
 Petrini, D.; 26; 277; 278
 Pianesi, F.; 43; 274
 Piao, S. S. L.; 278
 Piccioni, L.; 267
 Picoral, A.; 68; 81; 279
 Pienemann, M.; 150
 Pienenman, M.; 279
 Pierrard, M.; 149; 275
 Pistolesi, E.; 17; 279
 Pizzuto, E.; 98; 269
 Poggi Salani, T.; 22; 24; 55; 269; 279; 280
 Pontrandolfo, G.; 60; 278
 Pozzo, G.; 8; 153; 280
 Prandi, M.; 7; 39; 42; 43; 44; 45; 46; 57; 58; 59; 60; 61; 63; 181; 254; 280; 281
 Proietti, D.; 7; 30; 50; 272; 281
 Radtke, E.; 17; 265; 268; 271; 275; 280
 Ramat, P.; 29; 280
 Rastelli, S.; 149; 280
 Rayson, P.; 278; 280
 Rehm, G.; 113; 280
 Renzi, L.; 7; 30; 31; 32; 33; 36; 37; 42; 46; 50; 51; 58; 181; 267; 271; 280; 281; 284; 285
 Reppen, R.; 70; 269
 Roggia, C. E.; 36; 50; 280
 Romanini, F.; 7; 29; 35; 37; 42; 54; 60; 61; 63; 278
 Römer, U.; 77; 87; 88; 280
 Rossi, F.; 37; 38; 273; 280
 Rossini Favretti, R.; 99; 100; 280
 Rychlý, P.; 276
 Sabatini, F.; 3; 5; 7; 8; 14; 24; 26; 27; 28; 29; 30; 31; 33; 34; 35; 36; 39; 40; 41; 42; 46; 47; 48; 49; 50; 51; 52; 53; 55; 56; 58; 59; 60; 61; 62; 63; 64; 84; 85; 90; 91; 92; 116; 121; 129; 130; 141; 145; 157; 159; 176; 178; 180; 181; 254; 256; 259; 273; 280; 281
 Salvi, G.; 32; 33; 35; 36; 37; 49; 50; 267; 271; 281; 284; 285
 Sansò, A.; 48; 281
 Santini, M.; 113; 280; 281
 Santipolo, M.; 87; 281
 Scaglioso, A. M.; 266
 Schena, L.; 42; 280; 281
 Scherer, K. R.; 155; 281
 Schneider, S.; 42; 181; 281
 Schorr, A.; 281
 Schumann, J. H.; 155; 281
 Seiler, S.; 47; 277
 Selinker, L.; 78; 148; 282
 Semplici, S.; 147; 273
 Seracini, F.; 68; 80; 282
 Sérasset, G.; 277
 Serianni, L.; 19; 32; 33; 34; 42; 44; 50; 265; 272; 282; 285
 Sgroi, C. S.; 42; 282
 Sharoff, S.; 113; 278; 281; 282
 Shin, J.; 81; 279
 Simone, R.; 7; 25; 30; 266; 268; 269; 271; 273; 274; 276; 279; 282
 Sinclair, J. McH.; 7; 73; 74; 77; 79; 87; 88; 93; 255; 265; 282
 Słapek, D.; 4; 8; 90; 91; 123; 142; 177; 178; 259; 283
 Sobrero, A. A.; 7; 25; 26; 56; 57; 61; 63; 254; 268; 280; 282; 283
 Solarino, R.; 150; 277
 Sornicola, R.; 37; 40; 135; 283
 Sosnowski, R.; 240; 283
 Spina, S.; 106; 107; 108; 283
 Spinelli, B.; 183; 185; 283

Sprugnoli, L.; 266
 Staples, S.; 81; 279
 Stefanowitsch, A.; 76; 274; 283
 Stemle, E.; 277
 Stewart, D.; 78; 79; 88; 265; 277; 280
 Storchi, S.; 7; 68; 82; 255; 275
 Strambi, B.; 266
 Stubbe, A.; 280
 Stubbs, M.; 87; 283
 Suchomel, V.; 276
 Svartvik, J.; 93; 283
 Swatek, A.; 81; 279
 Symonenko, S.; 280
 Tamburini, F.; 99; 100; 280; 283
 Tavoni, M.; 7; 30; 283
 Tavosanis, M.; 280
 Taylor, D.; 87; 283
 Tesi, R.; 7; 30; 283
 Thompson, P.; 8; 118; 256; 278
 Thornton, A. M.; 276
 Titone, R.; 151; 283
 Tognini-Bonelli, E.; 70; 265; 283
 Tomlin, R. S.; 37; 283
 Tono, Y.; 68; 78; 81; 284
 Trevisan, M.; 97; 266
 Trévisiol-Okamura, P.; 149; 284
 Trifone, P.; 42; 265; 272; 282; 285
 Troncarelli, D.; 5; 91; 147; 153; 273
 Trousdale, G.; 77; 275
 Ueyama, M.; 102; 267
 Vališová, P.; 68; 80; 284
 Vanelli, L.; 33; 34; 284; 285
 Vecchio, P.; 183; 270
 Vedovelli, M.; 8; 97; 149; 266; 273; 284
 Véronique, D.; 149; 284
 Viale, M.; 54; 272; 278
 Vidulin, V.; 280
 Viganò, P. B.; 82; 284
 Villarini, A.; 8; 149; 284
 Visigalli, M.; 149; 273
 Viviani, A.; 7; 30; 50; 272
 Voghera, M.; 26; 97; 273; 284
 Volpi, A.; 267
 Wandruszka, U.; 43; 284
 Whittle, A.; 151; 284
 Wible, D.; 68; 80; 86; 284
 Widdowson, H. G.; 87; 283; 284
 Wilkins, D.; 87; 284
 Wilson, A.; 7; 65; 72; 278; 280
 Wulff, S.; 77; 274
 Wunderlich, C.; 68; 81; 284
 Xiao, R.; 77; 78; 79; 277; 278
 Yoon, J.; 77; 275
 Zamora, P.; 177; 179; 285
 Zampolli, A.; 269
 Zanchetta, E.; 102; 109; 267; 285
 Zingaro, A.; 40; 285

Appendice 1

Panoramica generale dei principali dati raccolti sui corpora presi in esame (cfr. capitolo 4).

Corpora	Date dei contenuti	Numero di <i>token</i>	Informazioni generi testuali	Disponibilità	Licenze <i>Copyleft</i>
1. Araneum Italicum Maius	Scaricati dal web nel 2014	1,2 mld.	Dati non reperiti.	A seconda delle piattaforme, accessibile previa registrazione o abbonamento.	“The copyright status of the data is not clear”. (Benko, 2016: 4248)
2. Araneum Italicum Minus	Scaricati dal web nel 2014	119 mln.	Come sopra	Come sopra.	Come sopra.
3. Athenaeum	Dati non reperiti.	306 mila	Prosa accademica (scientifica e amministrativa).	Liberamente accessibile online.	<i>Creative Commons – Share Alike.</i>
4. CLIPS	Dati non reperiti.	Dati non reperiti.	Dati non reperiti.	Consultabile online previa registrazione gratuita.	Dati non reperiti.
5. COLFIS	Prodotti tra il 1992 e il 1994	3,7 mln.	Quotidiani, periodici, libri.	Liberamente accessibile online.	Non sono state concesse licenze per la totalità dei testi.
6. CorDIC	Dati non reperiti.	500 mila	Dominio d’uso: arti, burocratico, creativo, economia e stampa	Liberamente accessibile online.	Dati non reperiti.
7. CorIS/CoDIS	Prodotti tra gli anni ’80 e il 2017	150 mln.	Stampa, narrativa, prosa accademica, prosa giuridico-amministrativa, miscellanea ed <i>ephemera</i>	Liberamente accessibile online.	Non sono state concesse licenze, o perlomeno non per la totalità dei testi.

8. ItTenTen	Scaricati dal web nel 2016	5,8 mld	Dati non reperiti.	Accesso gratuito per 30 giorni. In alternativa, varie opzioni di abbonamento.	Dati non reperiti.
9. ItWaC	Scaricati dal web tra il 2005 e il 2007	2 mld.	Dati non reperiti.	Liberamente accessibile online.	“There is no easy way of determining whether the content of a particular page is copyrighted.” (Baroni et al. 2009: 224)
10. la Repubblica	Prodotti tra il 1985 al 2000	326 mln.	Prosa giornalistica.	Liberamente accessibile online.	Nessuna
11. LexALP	Dati non reperiti.	3000 ¹	Linguaggio giuridico.	Liberamente accessibile online.	Trattandosi di testi giuridici, non sussiste la questione del copyright.
12. LIF	Prodotti tra il 1947 e il 1968	500.000 ²	Teatro, romanzi, cinema, periodici, sussidiari.	Non accessibile né online, né offline.	Nessuna.
13. NUNC	Prodotti tra il 2002 e il 2004	Due sezioni: 127 mln. e 109 mln. di <i>token</i>	Forum di cinema, cucina, fotografia e motori.	Liberamente accessibile online.	La risorsa dispone di licenza <i>Creative Commons – Share Alike</i> , ma non i singoli testi che la compongono.
14. PAISÀ	Scaricati dal web nel	250 mln.	Pagine tratte dai	Liberamente	<i>CC-Attribuzione</i> ,

¹ Si riferisce al numero di documenti e non di token.

² Lemmi e non token.

	2010		progetti Wikimedia Foundation, guide.supereva.it, italy.indymedia.org e vari blog.	accessibile online.	<i>CC-Attribuzione-Condividi allo stesso modo, CC-Attribuzione-Non commerciale, o CC-Attribuzione-Non commerciale-Condividi allo stesso modo</i>
15. Perugia Corpus	Prodotti tra il 1990 e il 2012	26 mln.	Letteratura, saggistica, stampa, scritto accademico, scritto scolastico, scritto amministrativo, web,	Accessibile previa registrazione.	Nessuna.
16. RIDIRE	Scaricati dal web tra il 2009 e il 2013	1,5 mld.	Informazione, economia e business, amministrazione e legislazione sport, architettura e design, cucina, cinema, musica, moda, arti visive, religione, letteratura e teatro.	Liberamente accessibile online per le funzioni di ricerca; è richiesto il log-in per le funzioni di gestione, monitoraggio e aggiornamento	Dati non reperiti.
17. SMS	Dati non reperiti.	1.395 ³	Dati non reperiti.	Previa richiesta al gestore e collaborazione al progetto.	Dati non reperiti.

³ Si riferisce al numero di SMS e non di token.

Appendice 2

Panoramica generale delle stringhe di ricerca prodotte (cfr. capitolo 4)

FENOMENO GRAMMATICALE	SOTTOFENOMENO	STRINGA DI RICERCA
Pronome <i>gli vs le (a lei) / loro (a loro)</i>		[word="(?)gli"] [postag="V.*"]
Il partitivo preceduto da preposizione (es.: <i>con degli amici; per degli amici</i>)		[postag="E" & word!="prima dopo"] [postag="EA" & lemma="di"]
Dislocazione a sinistra	Doppio partitivo (Es.: <i>di amici ne ho molti</i>)	[postag!="S.* A.*"] [word="(?)di"] [postag="S.*" & word!="fatto seguito conseguenza norma solito"] [word="ne"] [postag="V.*"]
<ul style="list-style-type: none"> dislocazione a sinistra dell'elemento pronominale 	A me mi	[word="(?)a"] [word="me"][word="mi"]
	A te ti	[word="(?)a"] [word="te"][word="ti"]
	A lui gli	[word="(?)a"] [word="lui"][word="gli"]
	A lei le	[word="(?)a"] [word="lei"][word="le"]
	A lei gli	[word="(?)a"] [word="lei"][word="gli"]
	A noi ci	[word="(?)a"] [word="noi"][word="ci"]
	A voi vi	[word="(?)a"] [word="voi"][word="vi"]
	A loro gli	[word="(?)a"] [word="loro"][word="gli"]

Per cui in funzione di congiunzione causale-consecutiva		[word="(?)per"] [word="cui"]
Impiego di congiunzioni coordinanti e di connettivi testuali in posizione iniziale di discorso o di enunciato	<i>Ma</i> in posizione iniziale	[word="\.\ ;"] [word="(?)Ma"]
	<i>E</i> in posizione iniziale	[word="\.\ ;"] [word="(?)E"]
	<i>Allora</i> in posizione iniziale	[word="\.\ ;"] [word="(?)Allora"]
	<i>Comunque</i> in posizione iniziale	[word="\.\ ;"] [word="(?)Comunque"]
Cosa come interrogativo al posto di <i>che cosa</i>	Nell'interrogativa diretta	[cpostag="F"] [postag!="A.* D.* R.*"]? [word="(?)cosa"] [word!="?"]{0,10} [word="\?"]
	Nell'interrogativa indiretta	[lemma="chiedere domandare"] [cpostag!="F" & word!="che"]? [word="(?)cosa"]
Soggetto in posizione rematica rispetto al predicato	con discordanza di numero rispetto al soggetto: <i>niente</i> seguito da un sostantivo plurale (Es.: <i>niente soldi</i>)	[word="niente"][postag="S.*" & feats=".*num=p.*"]
	sostantivo plurale seguito da <i>non ce n'è</i> (Es.: <i>soldi non ce n'è</i>)	[postag="S.*" & feats=".*num=p.*"] [word="non"] [word="ce"][word="n"] [lemma="essere" & feats=".*num=s.*" & feats=".*per=3.*"]

Soggetto in posizione rematica	rispetto a verbi non composti (Es.: <i>Parla Luca</i> . Per ulteriori esempi si rinvia al capitolo 4)	[word="\."] [postag="V.*" & word="[A-Z].+" & lemma!="essere occorrere servire avere esistere" & deprel="ROOT" & feats=".*mod=i.*"] [postag="R.* A.* D.*"] {0,2} [postag="S.* P.*" & deprel="subj"]
	rispetto a verbi composti con l'ausiliare <i>essere</i>	[word="\."] [lemma="essere"] [postag="V.*" & feats=".*mod=p.*"] [postag="R.* A.* D.*"] {0,2} [postag="P.* S.*" & deprel="subj"]
	rispetto a verbi composti con l' ausiliare <i>avere</i> , con oggetto diretto tra participio e soggetto	[word="\."] [lemma="avere"] [postag="V.*" & feats=".*mod=p.*"] [postag="R.* A.* D.*"] {0,2} [postag="S.*" & deprel="obj"] [postag="R.* A.* D.*"] {0,2} [postag="P.* S.*" & deprel="subj" & word!="che"]
	rispetto a verbi composti con l'ausiliare <i>avere</i> (senza l'oggetto diretto tra participio e soggetto)	[word="\."] [lemma="avere"] [postag="V.*" & feats=".*mod=p.*"] [postag="R.* A.* D.*"] {0,2} [postag="P.* S.*" & deprel="subj" & word!="che"]
	rispetto a verbi passivi	[word="\."] [lemma="essere venire andare"] [lemma="essere andare" & feats=".*mod=p.*"] ? [postag="V.*" & feats=".*mod=p.*"] [postag="R.* A.* D.*"] {0,2} [postag="P.* S.*" & deprel="subj"]
	in frase subordinata	[postag="CS" & word!="[A-Z].*"] [postag="V.*" & feats!=".*mod=f.*"] {1,2} [postag="R.* A.* D.*"] {0,2} [postag="P.* S.*" & deprel="subj"]

Frase scissa	verbo <i>essere</i> + gruppo nominale + <i>che</i> + verbo coniugato (es.: <i>è Anna che parla</i> . Per ulteriori esempi si rinvia al capitolo 4)	<pre>[postag="F.* C.*"] [lemma="essere"] [postag="R.* A.* D.*"]{0,2} [postag="P.* S.*"] [word="che"] [postag="V.*"]</pre>
	verbo <i>essere</i> + verbo all'infinito + <i>che</i> + verbo coniugato (es. <i>è votare che è importante</i>)	<pre>[postag="F.* C.*"] [lemma="essere"] [postag="V.*"& feats=".*mod=f.*"] [word="che"] [postag="V.*"]</pre>
	verbo <i>essere</i> + avverbio + <i>che</i> + verbo coniugato (es. <i>è qui che abitiamo</i>)	<pre>[postag="F.* C.*"] [lemma="essere"] [postag="B.*"] [word="che"] [postag="V.*"]</pre>
	verbo <i>essere</i> + gruppo nominale preceduto da preposizione + <i>che</i> + verbo coniugato (es. <i>è a casa che dobbiamo andare</i> . Per ulteriori esempi si rinvia al capitolo 4)	<pre>[postag="F.* C.*"] [lemma="essere"] [postag="E.*"] [postag="A.* D.*"]{0,2} [postag="P.* S.*"] [word="che"] [postag="V.*"]</pre>

	verbo <i>essere</i> + gruppo nominale + a + verbo al modo infinito (es. <i>è lui a telefonare</i>)	[postag="F.* C.*"] [lemma="essere"] [postag="R.* A.* D.*"]{0,2} [postag="P.* S.*"] [word="a"] [postag="V.*" & feats=".*mod=f.*"]
Soggetto in forma pronominale per indicare partecipazione affettiva con i verbi non composti (Es. <i>mi bevo/berrò/bevvi (ecc) il caffè</i>)	1a pers. sing.	[word="(?)mi"] [lemma="mangiare leggere bere fare guardare vedere" & feats=".*num=s.*" & feats=".*per=1.*"] [postag="R.* A.* D.*"]{0,2} [postag="P.* S.*" & deprel="obj"]
	2a pers. sing.	[word="(?)ti"] [lemma="mangiare leggere bere fare guardare vedere" & feats=".*num=s.*" & feats=".*per=2.*"] [postag="R.* A.* D.*"]{0,2} [postag="P.* S.*" & deprel="obj"]
	3a pers. sing.	[word="(?)si"] [lemma="mangiare leggere bere fare guardare vedere" & feats=".*num=s.*" & feats=".*per=3.*"] [postag="R.* A.* D.*"]{0,2} [postag="P.* S.*" & deprel="obj"]
	1a pers. plur.	[word="(?)ci"] [lemma="mangiare leggere bere fare guardare vedere" & feats=".*num=p.*" & feats=".*per=1.*"] [postag="R.* A.* D.*"]{0,2} [postag="P.* S.*" & deprel="obj"]
	2a pers. plur.	[word="(?)vi"] [lemma="mangiare leggere bere fare guardare vedere" & feats=".*num=p.*" & feats=".*per=2.*"] [postag="R.* A.* D.*"]{0,2} [postag="P.* S.*" & deprel="obj"]

	3a pers. plur.	[word="(?)si"] [lemma="mangiare leggere bere fare guardare vedere" & feats=".*num=p.*" & feats=".*per=3.*"] [postag="R.* A.* D.*"] {0,2} [postag="P.* S.*" &deprel="obj"]
Soggetto in forma pronominale per indicare partecipazione affettiva con i verbi composti <i>(Es. mi sono/ero/fui/ sarei... bevuto i bevo/berrò/bevvi (ecc) il caffè)</i>	1a pers. sing.	[word="(?)mi"] [lemma="essere" & feats=".*num=s.*" & feats=".*per=1.*"] [lemma="mangiare leggere bere fare guardare vedere" & feats=".*mod=p.*"] [postag="R.* A.* D.*"] {0,2} [postag="P.* S.*" &deprel="obj"]
	2a pers. sing.	[word="(?)ti"] [lemma="essere" & feats=".*num=s.*" & feats=".*per=2.*"] [lemma="mangiare leggere bere fare guardare vedere" & feats=".*mod=p.*"] [postag="R.* A.* D.*"] {0,2} [postag="P.* S.*" &deprel="obj"]
	3a pers. sing.	[word="(?)si"] [lemma="essere" & feats=".*num=s.*" & feats=".*per=3.*"] [lemma="mangiare leggere bere fare guardare vedere" & feats=".*mod=p.*"] [postag="R.* A.* D.*"] {0,2} [postag="P.* S.*" &deprel="obj"]
	1a pers. plur.	[word="(?)ci"] [lemma="essere" & feats=".*num=p.*" & feats=".*per=1.*"] [lemma="mangiare leggere bere fare guardare vedere" & feats=".*mod=p.*"] [postag="R.* A.* D.*"] {0,2} [postag="P.* S.*" &deprel="obj"]

	2a pers. plur.	[word="(?)vi"][lemma="essere"& feats=".*num=p.*" & feats=".*per=2.*"] [lemma="mangiare leggere bere fare guardare vedere"& feats=".*mod=p.*"] [postag="R.* A.* D.*"] {0,2} [postag="P.* S.*" & deprel="obj"]
	3a pers. plur.	[word="(?)si"][lemma="essere"& feats=".*num=p.*" & feats=".*per=3.*"] [lemma="mangiare leggere bere fare guardare vedere"& feats=".*mod=p.*"] [postag="R.* A.* D.*"] {0,2} [postag="P.* S.*" & deprel="obj"]
Concordanza a senso tra il soggetto collettivo singolare seguito da un'espressione partitiva e il predicato plurale (es.: <i>Un milione di elettori hanno votato</i>)		[postag="S.*" & feats=".*num=s.*"] [word="di"] [postag="S.*" & feats=".*num=p.*"] [postag="V.*" & feats=".*num=p.*" & feats!=".*mod=p.*"]
Ci “attualizzante”	C'+ verbo <i>avere</i> + sostantivo (Es.: <i>C'ha voglia</i>)	[word="(?)c"] [lemma="avere"] [postag="S.*"]
	Ci + verbo <i>avere</i> + sostantivo (Es.: <i>Ci ha voglia</i>)	[word="(?)ci"] [lemma="avere"] [postag="S.*"]
	C'+ verbo <i>avere</i> + da + verbo (Es.: <i>c'ho da fare</i>)	[word="(?)c"] [lemma="avere"] [word="da"] [postag="V.*"]

	Ci+ verbo <i>avere</i> + da + verbo (Es.: <i>ci ho da fare</i>)	[word="(?)ci"][lemma="avere"][word="da"][postag="V.*"]
--	---	--

<i>Che</i> polivalente	<i>che</i> esplicativo-consecutivo; <i>che</i> causale	[postag="V.*" & lemma!="ricordare dimostrare raccontare parere sembrare sentire dire confidare convenire sperare vedere volere sapere scrivere rispondere scoprire ribadire promettere mettere insistere dichiarare oservare sostenere immaginare temere ritenere premettere augurare pensare pretendere calcolare lasciare ripetere scommettere riconoscere significare vedere prevedere constatare supporre importare capire trovare decidere precisare aggiungere ammettere notare credere chiedere aspettare considerare bisognare occorrere avvenire bastare essere"] [word="che"] [postag="V.*"]
	<i>Che</i> pseudorelativo	[postag="P.*"&word!="si"] [lemma="vedere sentire guardare scorgere notare osservare mirare ascoltare udire intendere percepire"] [word="che"] [postag="V.*"]
	<i>Che</i> consecutivo-presentativo con avverbi di dubbio	[word="\.\.;\!\?\\,"] [word="(?)che"] [word="se forse chissà eventualmente probabilmente poi magari"] [] {0,1} [postag="V.*"]
Impiego dell'indicativo al posto del congiuntivo in alcune subordinate	dopo i verbi di opinione/sentimento	[lemma="pensare credere ritenere sembrare parere supporre dubitare immaginare temere sperare dispiacersi stupirsi augurarsi aspettarsi volere"] [word="che"] [word!="\."]{0,5} [postag="V.*" & feats=".*mod=i.*"]

	dopo i verbi di opinione/sentimento preceduti da negazione	[word="non"][lemma="pensare credere ritenere sembrare parere supporre dubitare immaginare temere sperare dispiacersi stupirsi augurarsi aspettarsi volere sapere"] [word="che"] [word!="\."]{0,5} [postag="V.*" & feats=".*mod=i.*"]
	dopo la costruzione “verbo <i>essere</i> + aggettivo + che”:	[lemma="essere"] [cpostag="A"] [word="che"] [word!="\."]{0,5} [postag="V.*" & feats=".*mod=i.*"]
	dopo la costruzione “verbo <i>avere</i> + paura/timore + che”:	[lemma="avere"][word="paura timore"][word="che"][word!="\."]{0,5}[postag="V.*" & feats=".*mod=i.*"]
	nell’interrogativa indiretta	[lemma="chiedere domandare"] [word="se"] [word!="\."]{0,5} [postag="V.*" & feats=".*mod=i.*"]
	nelle ipotetiche	[postag="CS" & word="(?i)se"] [word!="ne stess(o a i e)"]{0,5} [postag="V.*" & feats=".*mod=i.*" & feats=".*ten=[i].*"] [word!="\."]{0,5} [postag="V.*" & feats=".*mod=i.*" & feats=".*ten=[i].*"]

Appendice 3

Valutazione dell'unità didattica sui verbi riflessivi "affettivi"

Cosa pensi delle attività di oggi? Rispondi a questo questionario. Le tue risposte sono utili all'insegnante per programmare le future attività per gli studenti. **Grazie!** 😊

Tre domande su di te:

1. Qual è la tua lingua madre? _____
2. Quali altre lingue conosci? _____
3. Da quanto tempo studi l'italiano? _____

I contenuti di questa unità:

4. Prima delle attività di oggi, avevi già trovato degli esempi di uso dei verbi riflessivi in senso affettivo (esempio: Mi bevo un caffè)? _____
5. Se sì, in quale contesto? Segna con una crocetta una o più risposte:
 - A. Parlando con gli italiani
 - B. TV, giornali, libri ecc.
 - C. Sui social network
 - D. Chat/messaggi
 - E. Altro: _____
6. Esiste un fenomeno simile nella tua lingua?

7. La tua lingua madre o altre lingue che conosci ti hanno aiutato/a a imparare/capire meglio l'argomento di grammatica di oggi?

Opinioni sulle attività:








Metti una crocetta nelle tabelle qui sotto per dare la tua opinione. Risposte possibili:

- 😞😞 - **0: no**
- 😞/😊 - **1: poco**
- 😊 - **2: abbastanza**
- 😊😊 - **3 - sì**








È stata utile questa attività?	😞😞 0	😞/😊 1	😊 2	😊😊 3
8. Leggere il testo "L'appuntamento"				
9. Esercizi di collegamento (parole del testo – significato)				
10. Rispondere alle domande sul tema del testo				
11. Cercare nel testo i verbi che descrivono la cura del corpo				
12. Completare la coniugazione dei verbi riflessivi				
13. Ripassare le regole sulla posizione dei pronomi				
14. Distinguere le diverse azioni espresse dai verbi riflessivi (<i>azione del soggetto su se stesso/ su parti del corpo o vestiti/accessori che appartengono al soggetto; azione reciproca tra più soggetti; riflessivo "affettivo"</i>)				

15. Scegliere tra la forma riflessiva-affettiva e la forma non riflessiva				
16. Contenuto extra: i riflessivi nei modi di dire				
17. Fare esercizi basati su frasi vere della vita quotidiana (tratte da blog e altre pagine web) che non sono state inventate per “fare grammatica”				

Se vuoi, qui puoi spiegare meglio le tue risposte, aggiungere informazioni o consigli:

Dopo queste attività ...	  0	 /  1	 2	  3
18. ho un'idea più chiara dell'uso dei riflessivi				
19. riesco a distinguere meglio le situazioni formali da quelle informali				
20. ho capito che in alcune situazioni informali è accettabile usare i riflessivi “affettivi”				
21. ho ancora dei dubbi sull'uso dei riflessivi				

Se vuoi, qui puoi spiegare meglio le tue risposte, aggiungere informazioni o consigli:

Opinioni sulla piattaforma:	  0	 /  1	 2	  3
22. Il login e l'accesso alle diverse parti dell'unità è stato facile				
23. Le istruzioni delle attività erano comprensibili				
24. Le attività si visualizzano bene sullo schermo				
25. È stato utile avere esercizi con correzione automatica				

Se vuoi, qui puoi spiegare meglio le tue risposte, aggiungere informazioni o consigli:

Valutazione dell'unità didattica sull'uso del congiuntivo vs indicativo

Cosa pensi delle attività di oggi? Rispondi a questo questionario. Le tue risposte sono utili all'insegnante per programmare le future attività per gli studenti. **Grazie!** 😊

Tre domande su di te:

1. Qual è la tua lingua madre? _____
2. Quali altre lingue conosci? _____
3. Da quanto tempo studi l'italiano? _____

I contenuti di questa unità:

4. Prima delle attività di oggi, avevi già trovato degli esempi di uso dell'indicativo al posto del congiuntivo in situazioni informali?

5. Se sì, in quale contesto? Segna con una crocetta una o più risposte:
 - Parlando con gli italiani
 - Nei mass media
 - Sui social network
 - Chat/messaggi
 - Altro _____
6. Esiste un fenomeno simile nella tua lingua?

7. La tua lingua madre o altre lingue che conosci ti hanno aiutato/a a imparare/capire meglio l'argomento di grammatica di oggi?

Opinioni sulle attività:

È stata utile questa attività?

Metti una crocetta per dare la tua opinione su quanto sono state utili le attività.

Risposte possibili:

- 😞😞 **0: no**
- 😞/😊 **1: poco**
- 😊 **2: abbastanza**
- 😊😊 **3: sì**

ATTIVITÀ	😞😞 0	😞/😊 1	😊 2	😊😊 3
8. Cercare nei 4 brevi testi gli esempi d'uso del congiuntivo				
9. Ripassare i casi in cui il congiuntivo è obbligatorio				
10. Fare ipotesi sulla scelta tra congiuntivo e indicativo in base al contesto formale/informale.				
11. Fare esercizi basati su frasi vere della vita quotidiana (tratte da blog e altre pagine web) che non sono state inventate per "fare grammatica"				

Se vuoi, qui puoi spiegare meglio le tue risposte, aggiungere informazioni o consigli:

Dopo queste attività ...	SÌ	NO
12. ho un'idea più chiara di quando è obbligatorio il congiuntivo		
13. riesco a distinguere meglio le situazioni formali da quelle informali		
14. ho capito che in alcune situazioni informali è accettabile non usare il congiuntivo		
15. ho ancora dei dubbi sulla scelta tra congiuntivo e indicativo		

Se vuoi, qui puoi spiegare meglio le tue risposte, aggiungere informazioni o consigli:

Opinioni sulla piattaforma:	SÌ	NO
16. Il login e l'accesso alle diverse parti dell'unità è stato facile		
17. Le istruzioni delle attività erano comprensibili		
18. Le attività si visualizzano bene sullo schermo		
19. È stato utile avere esercizi con correzione automatica		
20. Mi è piaciuta l'idea delle attività online		

Se vuoi, qui puoi spiegare meglio le tue risposte, aggiungere informazioni o consigli:

Appendice 4

Le unità didattiche

Si ricorda che le unità didattiche sono collocate sulla piattaforma *Wizer.me*. Alcune imperfezioni relative alla grafica e all'impaginazione dei seguenti documenti sono dovute alla conversione in PDF delle schermate delle pagine web.

Mi preparo per uscire.

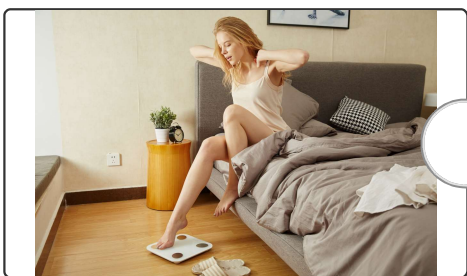
Obiettivi:

Questa unità ci aiuterà a:

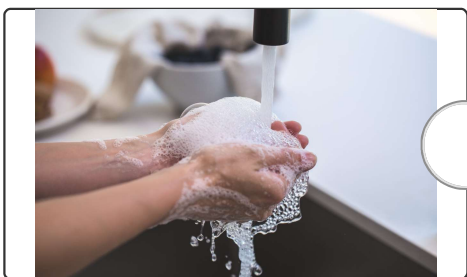
- Arricchire il **lessico** sulle azioni della **vita quotidiana**;
- Ripassare i **pronomi riflessivi**;
- **Per approfondire:** uso dei riflessivi in senso intensivo - affettivo nell'italiano neostandard.

Che cosa fai prima di uscire di casa?

Collega le immagini ai verbi che le descrivono.



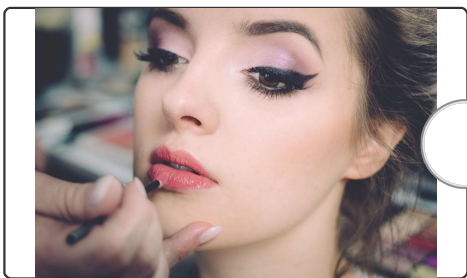
radersi/farsi la barba



alzarsi



lavarsi



vestirsi

pettinarsi

truccarsi

Che cosa mi metto?

L'APPUNTAMENTO

Leggi il testo, per avere un'idea generale del tema. Non preoccuparti di capire tutto, cerca di cogliere il senso generale.



Sono già le otto e sono ancora sotto la doccia, mi devo asciugare i capelli: questi maledetti capelli, sono sempre spettinata! Accappatoio, mousse e tagliaunghie. Devo stare calma altrimenti sudo e devo di nuovo fare una doccia.

Otto e un quarto. Crema per il corpo. Mi lavo i denti? Sì, me li lavo, ma mi fumo una sigaretta. Speriamo che fumi anche lui. Ecco, lo smalto per le unghie, potevo mettermelo prima, adesso non c'è più tempo: lo farò la prossima volta, sempre che ci sia una prossima volta. Stai calma. Sì, sto calma. Mentre mi asciugo, per ottimizzare il tempo, preparo il vestito. È arrivato il momento dell' "Oddio, cosa mi metto?" Gonna o pantalone? Avevo deciso per la gonna e allora mi metto la gonna, no? E se poi mi porta in un locale alternativo? Con musica alternativa? Con gente alternativa? Io cosa ci faccio con la mia gonnellina? Meglio un classico pantalone con giacca. Distinta, professionale, va bene sempre e non impegna, come direbbe mia madre. Ecco, deciso. Accendo il phon a massima potenza, testa in giù, tv accesa ad alto volume.

Otto e quarantacinque. I capelli sono ancora bagnati ma possono andare. Orecchini di zircono, una cosa fine. Mi trucco: fondotinta, blush, mascara e rossetto. Sudo. Mi metto il deodorante. Mi vesto. Devo dimagrire, ma non adesso, non starò a guardare se ho 2 chili in più e se lo fa è un cretino. E poi neanche lui sarà perfetto. Scarpe con tacco (niente scarpe da ginnastica: vi devo proprio lasciare a casa stasera). Spruzzo il profumo.

Sono le nove. Sono una donna puntuale. Io sono puntuale, lui no.

Cellulare: "Dimmi"

"Mi sono infossato davanti a casa tua, sai ... la nebbia... non ho visto il bordo della strada... puoi scendere con una corda per darmi una mano a spingerla su?"

La prossima volta scelgo un uomo che mi venga a prendere con un trattore!

Adattamento del brano tratto da:

<https://notimetolose.myblog.it/?s=appuntamento&submit=Cerca> (consultato il 12/03/2019)

Collega le parole nella colonna di sinistra con il loro significato

a. mousse	○	cristallo che somiglia al diamante
b. tagliaunghie	○	mezzo di trasporto agricolo
c. ottimizzare	○	strumento per tagliare le unghie
d. zircono	○	finire in una buca nel terreno
e. trattore	○	schiuma per capelli
f. infossarsi	○	sfruttare al meglio

Rispondi alle seguenti domande:

Quali emozioni prova la protagonista del testo prima dell'appuntamento? Quali parole ti fanno capire questa cosa?

Write your answer...

Answer recorder (optional) -  Voice

Cosa succede all'uomo?

Write your answer...

Answer recorder (optional) -  Voice

Qual è la reazione della protagonista?

Write your answer...

Answer recorder (optional) -



Voice

Analisi

Trova nel testo tutte le parole che riguardano:

ABBIGLIAMENTO e ACCESSORI	PRODOTTI PER IL CORPO	VERBI che descrivono la cura del corpo
<div style="border: 2px dashed orange; height: 60px;"></div>	<div style="border: 2px dashed orange; height: 60px;"></div>	<div style="border: 2px dashed orange; height: 60px;"></div>

Riflessione

Guardiamo questi esempi:

Lavarsi i capelli, accendersi una sigaretta, lavarsi i denti, mettersi un vestito, infossarsi con la macchina, incontrarsi ecc. sono tutte azioni

- 1) che il soggetto fa su se stesso (mi lavo, mi pettino);
- 2) che riguardano direttamente il soggetto (il ragazzo si è infossato con la macchina; Giulia si sente un po' male);
- 3) che coinvolgono due o più persone: es. incontrarsi, sposarsi, eccetera.

Quasi tutti i verbi possono avere una forma riflessiva.

Di quali pronomi hanno bisogno questi verbi?

Completa la tabella con i pronomi e i verbi mancanti:

LAVARSI	METTERSI	VESTITI
io mi lavo		
	tu ti metti	
		lui/lei/Lei si veste

noi ci laviamo		
	voi vi mettete	
		loro si vestono

Ripassiamo le regole sulla posizione dei pronomi diretti e indiretti.

I pronomi:

1. si trovano normalmente **prima** del verbo;
2. si trovano dopo **il verbo e si uniscono a questo** solo nel caso di:
 - a. infinito
 - b. imperativo
 - c. gerundio
 - d. participio
3. con i verbi *volere, dovere, potere, sapere, cominciare a, stare per e finire di + l'infinito* di un verbo (esempio: comincio a capire), il pronome può trovarsi **o prima del verbo coniugato** (la comincio a capire) o **dopo il verbo all'infinito**, unito a questo (comincio a capirla).

La posizione dei pronomi

1. Nelle seguenti frasi ci sono degli esempi delle regole grammaticali sulla posizione dei pronomi. Scegli a quale regola corrispondono le frasi.

Esempio: cosa mi metto? Il pronome si trova prima del verbo

È questa la pietra che vuoi? L'ho trovata nella foresta. Prendila, è tua !

a Prima del verbo

b Unito al verbo all'imperativo

- c Unito al verbo all'infinito d Unito al verbo al gerundio
e Unito al verbo al participio.

Nel 1970, un surfer dal nome di Frank Nasworthy iniziò a sviluppare delle nuove ruote per lo skateboard. Usandole la sensazione era magnifica paragonata alle precedenti ruote d'argilla.

- a Prima del verbo b Unito al verbo all'imperativo
c Unito al verbo all'infinito d Unito al verbo al gerundio
e Unito al verbo al participio

Nessuno sta dicendo che non devi ascoltare la Spears. Ascoltala, ma non venite a dirci che è la più grande pop star di tutti i tempi!

- a Prima del verbo b Unito al verbo all'imperativo
c Unito al verbo all'infinito d Unito al verbo al gerundio
e Unito al verbo al participio

Nessuno sta dicendo che non devi ascoltare la Spears. Ascoltala, ma non venite a dirci che è la più grande pop star di tutti i tempi!

- a Prima del verbo b unito al verbo all'imperativo
c unito al verbo all'infinito d unito al verbo al gerundio
e unito al verbo al participio

I capelli? Devo ancora asciugarli.

Qual è un altro modo per costruire questa frase? Prova a cambiare la posizione del pronome all'interno di questa frase. In caso di dubbi, rileggi le regole sulla posizione dei pronomi.

Write your answer...

Answer recorder (optional) -  Voice

Alcuni automobilisti di passaggio, vistolo uscire di strada, si erano fermati per dargli soccorso.

-
- | | | | |
|---|-------------------------------|---|-------------------------------|
| a | Prima del verbo | b | unito al verbo all'imperativo |
| c | unito al verbo all'infinito | d | unito al verbo al gerundio |
| e | unito al verbo al participio. | | |

Sono nascosta al capolinea degli autobus e finalmente lo vedo. Comincio a seguirlo a breve distanza.

Qual è un altro modo per costruire questa frase? Prova a cambiare la posizione del pronome all'interno di questa frase. In caso di dubbi, rileggi le regole sulla posizione dei pronomi.

Write your answer...

Answer recorder (optional) -  Voice

Osserviamo ora questi casi:

1. Mi metto la gonna >> mettersi;
2. Mi vesto >> vestirsi;
3. Ci incontriamo > incontrarsi;
4. Mi asciugo i capelli >> asciugarsi;
5. Mi fumo una sigaretta >> fumarsi.
6. Mi trucco >> truccarsi.

Hanno tutti una struttura simile, con un **verbo che all'infinito termina con il suffisso -si**, ma le azioni che descrivono questi verbi appartengono a **quattro categorie diverse**.

Inserisci nella tabella gli esempi 1-6 che trovi qui di seguito.
NB: Ogni esempio appartiene ad una sola delle 4 categorie (a, b, c, d).

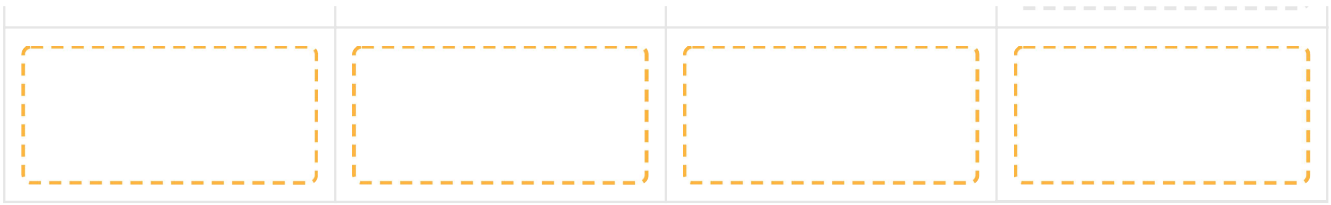
1. Mi metto la gonna;
2. Mi vesto;
3. Ci incontriamo;
4. Mi asciugo i capelli;
5. Mi fumo una sigaretta;
6. Mi trucco.

a) Azione reciproca tra più soggetti

b) azione su **parti del corpo** o **vestiti/accessori** che appartengono al soggetto

c) azione del soggetto **su se stesso**

d) il verbo non è un vero verbo riflessivo. Il pronome riflessivo viene **usato colloquialment** e solo per **sottolineare la partecipazione psicologica** di chi compie l'azione



Clicca qui per approfondire: i “riflessivi affettivi”

<http://www.accademiadellacrusca.it/it/lingua-italiana/consulenza-linguistica/domande-risposte/mangiarsi-pizza-fumarsi-sigaretta-ascoltarsi>
(Consultato il 19/03/2019)

La **definizione d** descrive i casi di “**riflessivi affettivi**”, in cui l’uso del pronome riflessivo serve solo ad esprimere il coinvolgimento emotivo, l’intensità con cui viene fatta l’azione. “Si tratta di un uso proprio del registro familiare e colloquiale [...] generalmente ammesso nel parlato ma inopportuno negli scritti più formali e sorvegliati.”



Clicca qui per approfondire: il dativo etico



[http://www.treccani.it/enciclopedia/dativo-etico_\(Enciclopedia-dell'Italiano\)/](http://www.treccani.it/enciclopedia/dativo-etico_(Enciclopedia-dell'Italiano)/) (Consultato il 19/03/2019)

Applicazione pratica

CIBO E BEVANDE

Completa le frasi con i seguenti verbi: *faccio – fai – mangiamo – facciamo – beve – mangiava – berremo – fanno – mangio – berremo.*

La domenica metto la sveglia alle 13 , mi alzo con calma , accendo la tv , mi metto a tavola praticamente in mutande, e me la vedo mentre mi _____ qualcosa di gustoso

Stasera mi _____ il cous-cous con un sacco di verdure prelibate.

Ma lo trovi il tempo per mangiare? Ti _____ un panino davanti alla tv.

Era un criceto normale, che si _____ i suoi semi di girasole, rompeva la carta igienica e si arrampicava

Non è possibile che una persona che si _____ una birra nel parco debba essere multata in questo modo.

Ci _____ un ottimo pasto.

Ti voglio libera presto, oggi, in questo minuto. Poi ci _____ un bicchiere di vino.

Ci _____ una birra

Fai del tuo meglio e se ci riuscirai, ci _____ una tazza di tè.

Quanti di noi si _____ il cous-cous invece del riso o della pasta, ogni tanto?

TEMPO LIBERO

Completa le frasi con i seguenti verbi: *leggono - guarderanno - vedevo - guardo - leggevo - guardate - vediamo - leggo - leggete.*

Se voglio vedere una storia intricata e originale mi _____ un altro genere di film, non un horror... mio parere personale.

Nei momenti di svago mi _____ film e giocavo un po'.

Insomma, mi _____ questo articolo e scopro che la fascia oraria del programma che voglio vedere in tv è quella delle 12 -14.

“Ciao Owen, sono io, Nadine! Volevo sapere che facevi, ci _____ un film?”

Se volete, vi _____ l'articolo su Repubblica.

A: Cosa stavi facendo?

B: Ero qui sul divano, che mi _____ i commenti a questo articolo.

Io ho degli amici francesi e italiani che domani sera si _____ la partita insieme.

Il corteo con la torcia olimpica si è spostato a destra verso la piazza. Se vi _____ una mappa da via Barberia, la strada stretta, attraverso piazza Malpighi.

Ci sono magistrati che fanno udienza quando vogliono, che non si _____ i fascicoli prima di andare in tribunale.

Applicazione pratica - 2

Quante cose possiamo fare? Completa le frasi con i pronomi riflessivi e le voci del verbo *fare* alla forma riflessiva:

Esempio:

E se alla fine (farsi, tu) _____ i conti, è meglio che giocare in Borsa » **ti fai** i conti.

Ci sono persone che vivono in 15 dentro un appartamento! È necessario che tutti noi _____ un discorsetto su chi ha permesso e sta permettendo tutto questo.

Noi _____ una diretta in streaming

La mattina presto, quando vado a correre, _____ dei giri completi, quindi vedo tutto il parco.

Adesso noi due _____ una bella chiacchierata.

"L'acqua è molto alta, qui non si tocca con i piedi! Bambini, state attenti, quando _____ il bagno!"

Ci sono politici di settant' anni che (farsi, loro) _____ il trapianto di capelli!

Mentre gli altri escono dai sacchi a pelo, io _____ un bagno nel lago, che a quest' ora è gelido.

(Fare, io) _____ il mio viaggetto fino a Milano e quando arrivo in Stazione Centrale mi trovo davanti uno splendido spettacolo.

Quel programma televisivo è orribile. Se (fare, io) _____ un pisolino è meglio, molto meglio.

Seguiamo i consigli di Laura e (fare, noi) _____ un bel giro lungo la costa tirrenica della Sicilia.

Coniuga al passato prossimo il verbo riflessivo all'infinito che trovi tra parentesi.

NB: l'ausiliare dei verbi riflessivi è sempre *essere*.

TEMPO LIBERO

Esempio: Da poco (farsi, lui) _____ una vacanza in una beauty farm

>> Da poco **si è fatto** una vacanza in una beauty farm.

Lucia: "Ieri sera (guardarsi, io) _____ la registrazione del webcast ".NET ed il test del software "".

Tra il calore delle quattro mura di Giuly, le tre amiche (farsi) _____ i pop-corn e Giuly si è rilassata.

Marco: "Ieri sera (vedersi, io) _____ la puntata di Ballarò registrata martedì."

Francesca: "Al cinema (vedersi, io) _____ Avatar e si usavano gli occhialini per il 3D.

Giovanni: "Io (leggersi) _____ il libro in due giorni isolandomi dal resto del mondo."

Luca chiede a Marta: "Ma il 4 giugno dove stavi, in piazza ? (farsi, tu) _____ un giro al Pink Paint Party?"

I giornalisti (bersi) _____ un whiskey prima di scrivere questo pseudo-articolo?

Sara: "Ho incontrato Jorge Garcia, l'attore di Lost, e (farsi, noi) _____ una foto insieme."

Marco: "Ho conosciuto Arisa ai provini di X Factor e recentemente ci siamo rivisti e (farsi, noi) _____ una chiacchierata ridendo del fatto che allora non eravamo nessuno."

Roberta: "Per smaltire il pranzo io e le mie amiche (farsi, noi) _____ una bella passeggiata e devo dire che la città di Comacchio mi è piaciuta. "

BELLEZZA

Elisa: "Avevo comprato uno spazzolino da denti giallo, una crema idratante verde, (farsi, io) _____ le unghie rosse e i capelli di due toni più scuri."

Quando mi pettino le sopracciglia, mi dicono che (farsi, io) _____ il lifting agli occhi. Rifatta, io? Tutta roba mia!

A: "Sei andato in montagna?"

B: "No, (farsi, io) _____ una lampada"

Brad Pitt è considerato il più bono ma non è così! In verità (farsi, lui) _____ la plastica facciale.

Applicazione pratica -3

Qual è la funzione del verbo riflessivo? Scegli la risposta giusta:

1) Mi lavo i capelli due o tre volte a settimana.

- a azione reciproca tra più soggetti
- b azione compiuta su parti del corpo/ vestiti/accessori del soggetto
- c azione del soggetto su se stesso
- d azione descritta da un verbo riflessivo "affettivo"

2) Ma il 4 giugno dove stavi, in piazza? Ti sei fatta un giro al *Pink Paint Party*?

- a azione reciproca tra più soggetti
- b azione compiuta su parti del corpo/ vestiti/accessori del soggetto
- c azione del soggetto su se stesso
- d azione descritta da un verbo riflessivo "affettivo"

3) Mi sono messa le scarpe e sono uscita.

- a azione reciproca tra più soggetti
- b azione compiuta su parti del corpo/ vestiti/accessori del soggetto

- c azione del soggetto su se stesso
- d azione descritta da un verbo riflessivo "affettivo"

4) Mi sono letto il libro in due giorni isolandomi dal resto del mondo.

- a azione reciproca tra più soggetti
- b azione compiuta su parti del corpo/ vestiti/accessori del soggetto
- c azione del soggetto su se stesso
- d azione descritta da un verbo riflessivo "affettivo"

5) Il 29 luglio 2000 a Malibu Jennifer Aniston si è sposata con l'attore Brad Pitt con cui era fidanzata da due anni e mezzo.

- a azione reciproca tra più soggetti
- b azione compiuta su parti del corpo/ vestiti/accessori del soggetto
- c azione del soggetto su se stesso
- d azione descritta da un verbo riflessivo "affettivo"

6) Per digerire il pranzo ci siamo fatte una bella passeggiata e devo dire che la città di Comacchio mi è piaciuta.

- a azione reciproca tra più soggetti
- b azione compiuta su parti del corpo/ vestiti/accessori del soggetto
- c azione del soggetto su se stesso
- d azione descritta da un verbo riflessivo "affettivo"

7) Nuovi disordini sono scoppiati a metà pomeriggio a Ginevra: gruppi di dimostranti mascherati si sono scontrati con la polizia.

- a azione reciproca tra più soggetti
- b azione compiuta su parti del corpo/ vestiti/accessori del soggetto
- c azione del soggetto su se stesso
- d azione descritta da un verbo riflessivo "affettivo"

8) Secondo te ogni volta che mi guardo un film mi leggo prima anche il relativo libro?

- a azione reciproca tra più soggetti
- b azione compiuta su parti del corpo/ vestiti/accessori del soggetto
- c azione del soggetto su se stesso
- d azione descritta da un verbo riflessivo "affettivo"

9) Alessia si è messa per l'occasione un abito costosissimo.

- a azione reciproca tra più soggetti
- b azione compiuta su parti del corpo/ vestiti/accessori del soggetto
- c azione del soggetto su se stesso
- d azione descritta da un verbo riflessivo "affettivo"

10) Mentre mi asciugo, per ottimizzare il tempo preparo i vestiti.

- a azione reciproca tra più soggetti
- b azione compiuta su parti del corpo/ vestiti/accessori del soggetto
- c azione del soggetto su se stesso
- d azione descritta da un verbo riflessivo "affettivo"

11) Attori comici straordinari, Laurel e Hardy si stimavano moltissimo ed erano amici e complici sul set.

-
- a azione reciproca tra più soggetti
 - b azione compiuta su parti del corpo/ vestiti/accessori del soggetto
 - c azione del soggetto su se stesso
 - d azione descritta da un verbo riflessivo "affettivo"

12) Se voglio un film di alto livello, mi guardo un film di Kubrick o di Nolan.

-
- a azione reciproca tra più soggetti
 - b azione compiuta su parti del corpo/ vestiti/accessori del soggetto
 - c azione del soggetto su se stesso
 - d azione descritta da un verbo riflessivo "affettivo"

13) Scienziati provenienti da più di 100 paesi si sono incontrati per scambiarsi esperienze e per conoscere le notizie più recenti sia dall'occidente che dall'oriente.

-
- a azione reciproca tra più soggetti
 - b azione compiuta su parti del corpo/ vestiti/accessori del soggetto
 - c azione del soggetto su se stesso
 - d azione descritta da un verbo riflessivo "affettivo"

14) Da poco si è fatto una vacanza in una beauty farm.

-
- a azione reciproca tra più soggetti
 - b azione compiuta su parti del corpo/ vestiti/accessori del soggetto
 - c azione del soggetto su se stesso
 - d azione descritta da un verbo riflessivo "affettivo"

15) Inizialmente Nathan e Lucas si odiano, ma nel corso della serie televisiva la loro relazione migliora molto.

- a azione reciproca tra più soggetti
- b azione compiuta su parti del corpo/ vestiti/accessori del soggetto
- c azione del soggetto su se stesso
- d azione descritta da un verbo riflessivo "affettivo"

Applicazione pratica - 4

Completa le seguenti frasi. Scegli, in base al contesto informale o formale, se usare la forma riflessiva-affettiva o la forma non riflessiva dei verbi dati:

bere - fare - fare - fare - guardare - guardare - leggere - leggere - leggere - vedere

NB: gli spazi vuoti possono contenere più di una parola.

1) L'insegnante dice: "Se (voi) _____ un po' di storia del movimento operaio italiano, scoprite che il problema della salute nelle fabbriche è stato sollevato per la prima volta proprio dagli stessi operai."

2) In un commento scritto su Facebook: "Se cerchi una soluzione, basta che _____ un giro su internet e avrai tutte le risposte ai tuoi dubbi."

3) Un giornalista scrive la recensione di un libro: "Se (voi) _____ l'ultimo libro di Bendetti sulla Cecenia, avrete dei motivi per preoccuparvi".

4) Un'attrice in un programma televisivo trash di intrattenimento: "Mi dicono che io alcuni mesi fa _____ il lifting agli occhi. Rifatta, io? Questa è tutta roba mia!"

5) Un amico racconta: "Alcuni anni fa al cinema io _____ "Avatar" e si usavano ancora gli occhialini per il 3D."
(NB: coniugare il verbo al passato prossimo)

6) In un programma televisivo culturale: "Francamente oggi i libri pubblicati sono troppi . E se (noi) _____ con attenzione l' articolo apparso sull'ultimo numero del settimanale "Il Domenicale", abbiamo il quadro completo della situazione."

7) Un amico ti dice al telefono: " Se la sera non esci, poi (tu) _____ un panino davanti alla tv."

8) In un libro sulla mitologia greca: "È importante ricordare la leggenda di Prometeo, che ogni giorno _____ un'aquila che mangiava il suo fegato".

9) Due persone al bar commentano il giornale di oggi: "Guarda, secondo me i giornalisti _____ un whiskey prima di scrivere questo pseudo-articolo!"
(NB: coniugare il verbo al passato prossimo)

10) Un amico racconta: "Da piccoli io ed altri _____ i film d'azione dove si picchiavano di brutto."

Contenuti extra: i verbi riflessivi nei modi di dire

Uso del verbo *vedersi* con il significato di *immaginarsi*.
Coniuga i verbi al presente.

Esempio:

Fiorello ha detto che gli piacerebbe rifare “Fantastico”:
(vedersi, tu) _____ in uno spettacolo di questo tipo?

>> **Ti vedi** in uno spettacolo di questo tipo?

(Significato: ti immagini in uno spettacolo di questo tipo?)

Il mio sogno? Tra 4 anni (vedersi, io) _____ vestito di rosso in Australia per la prima gara di Mondiale Superbike.

Giornalista: “Dove (vedersi, tu) _____ dopo Cultura Moderna?”

Attrice: “Amo vivere giorno per giorno senza fare troppi programmi.”

Il massimo sarebbe abitare in una casa di legno sugli scogli, lontano da tutti, con all'orizzonte solo il mare. Magari con grandi vetrate per il sole nelle giornate di maggio, una piccola veranda e poco più. (

Vedersi, io) _____ già lì adesso.

Giornalista: “(Vedersi, tu) _____ nel campo dello spettacolo?”

Attrice: “Mi piacerebbe far ridere la gente, far divertire. Recitare in una sitcom o in un programma comico: (vedersi, io) mi vedo così.”

Alcune persone usano la legge per i propri comodi. Già (vedersi, io) _____ le persone che si inventeranno i sistemi più astrusi per sfruttare economicamente la legge.

Leggi le seguenti frasi, che contengono il verbo *mangiarsi* usato in senso metaforico, poi completa la tabella:

1) Ti dico da medico che uno stomaco vuoto non preme sul diaframma e quindi sì, respiri e canti meglio (ma questo vale per tutti perché se ti mangi un vitello intero, sei appesantito anche a salire le scale).

2) Se non hai Sky, i canali nazionali sono solo quelli, quindi “o ti mangi quella minestra, o ti butti dalla finestra”).

3) Ci sono vecchie donne di teatro che conservano un fascino, un sguardo fiero ed altero. Indifferenti al tempo che si mangia il corpo, conservano un gesto fatale che è negato agli altri, che può essere un semplice sistemarsi i capelli o accendersi una sigaretta.

4) Ci fanno una multa da 600 mila Euro, ci mangiamo il capitale e la società è fallita: cosa faremo?

La seguente tabella contiene 4 significati metaforici del verbo *mangiarsi*. Scrivi i numeri delle frasi (1, 2, 3 o 4) in cui il verbo *mangiarsi* è usato con quel significato:

--	--

Significati metaforici	Numero della frase (1, 2, 3 o 4)
a. rovinare, distruggere.	
b. esaurire, sprecare;	
c. mangiare molto;	
d. non avere alternative: o così o niente;	

Produzione scritta e orale

Produzione scritta

Scrivi un breve racconto di una tua giornata tipica. Prova a riutilizzare il lessico che hai trovato nel testo e anche i verbi riflessivi di tipo "intensivo-affettivo" (es. mi mangio un panino).

Write your answer...

Answer recorder (optional) -  Voice

Produzione orale

A coppie: ciascuno prepara 5 domande sulla vita quotidiana da fare all'altra persona. Provate a riutilizzare il lessico che avete trovato nel testo e anche i verbi riflessivi di tipo "intensivo-affettivo" (es. mi mangio un panino). L'obiettivo è quello di raccontare la propria vita quotidiana, le proprie abitudini e interessi.

Write your answer...

Answer recorder (optional) -  **Voice**

Congiuntivo_01

Leggi questi 4 testi e cerca i verbi al congiuntivo.

Testo n.1: Arte

Sebbene quasi tutta la produzione di Kaspar David Friedrich sia basata sul paesaggio, l'artista disegnava e dipingeva nel suo studio, basandosi unicamente sui suoi numerosi disegni realizzati en plein air (...). Questo processo di progettazione è molto evidente nel "Monaco in riva al mare" (...). Tramite alcune fonti, si può immaginare che il pittore abbia iniziato questo dipinto a Dresda nel 1808.

https://it.wikipedia.org/wiki/Monaco_in_riva_al_mare

Testo n.2: Acqua su Marte

Generalmente si ritiene che l'acqua (allo stato liquido) sia sinonimo di vita; perciò l'esplorazione di Marte con la missione Mars Exploration Rover non è stata finalizzata direttamente alla ricerca di forme vitali, ma alla scoperta del prezioso liquido.

https://it.wikipedia.org/wiki/Vita_su_Marte

Testo n.3: Attraversando l'Africa

Fausto Moroni è biologo, è appassionato di fotografia, da anni frequenta l'Africa, che conosce bene. (...) Collabora con l'agenzia fotografica Homo Ambiens, specializzata in fotografia di natura (...). L'autore impiega due reflex: una Nikon F-70 e una Nikon F-30. (...) A chi gli chiede se sia interessato anche alla fotografia digitale, risponde subito ribadendo che ogni innovazione deve essere colta con curiosità.

http://guide.supereva.it/viaggi_safari/interventi/2003/03/131025.shtml

Testo n.4:

La scrittura deve servire innanzitutto a se stessi, per riflettere e trovare la coscienza di sé, come la meditazione nelle filosofie orientali. A proposito del fatto che la scrittura dev'essere per prima cosa un piacere in sé, vorrei ricordare cosa mi hanno detto Natalia Aspesi: "Penso che scrivere sia un grande piacere personale che ripaga di ogni delusione".

http://guide.supereva.it/case_editrici/interventi/2005/07/218030.shtml

Quando è usato il congiuntivo? Scrivi gli esempi che hai trovato nei testi:

Con un verbo di opinione	Con una forma impersonale	Con una congiunzione subordinante	Nella frase interrogativa indiretta

Congiuntivo-02

Perché ci serve il congiuntivo?

Il congiuntivo è il modo verbale che in generale si usa per esprimere la volontà, le opinioni, le incertezze, la possibilità e i sentimenti.

Esistono quattro tempi del congiuntivo: il presente, il passato, l'imperfetto e il trapassato. **Per il momento studiamo solo due tempi: il congiuntivo presente e passato.**

Presente

	Essere	Avere	Prima coniugazione	Seconda coniugazione	Terza coniugazione	
			Am-are	ved-ere	apr-ire	cap-ire
io	sia	abbia	ami	veda	apra	capisca
tu	sia	abbia	ami	veda	apra	capisca
lui/lei/ Lei	sia	abbia	ami	veda	apra	capisca
noi	siamo	abbiamo	amiamo	vediamo	apriamo	capiamo
voi	siate	abbiate	amiate	vediate	apriate	capiate
loro	siano	abbiano	amino	vedano	aprano	capiscano

Passato

Congiuntivo presente di essere/avere + participio passato

	Essere	Avere	Un verbo con l'ausiliare <i>essere</i>	Un verbo con l'ausiliare <i>avere</i>
			and-are	ved-ere
io	sia stato	abbia avuto	sia andato	abbia visto
tu	sia stato	abbia avuto	sia andato	abbia visto
lui/lei/ Lei	sia stato	abbia avuto	sia andato	abbia visto
noi	siamo stati	abbiamo avuto	siamo andati	abbiamo visto
voi	siate stati	abbiate avuto	siate andati	abbiate visto
loro	siano stati	abbiano avuto	siano andati	abbiano visto

Casi in cui il congiuntivo è obbligatorio:

1. Nelle frasi principali

Quando diamo del Lei a qualcuno, usiamo il congiuntivo se vogliamo **consigliare** a qualcuno di fare qualcosa, quando vogliamo **invitare** qualcuno a fare qualcosa o **dare un ordine** in modo gentile o vogliamo **fare delle richieste**:

-
1. Signora, (scusare) _____ , va in stazione?
 2. Dottore, (venire) _____ qui.
 3. Buonasera, signora! (Entrare), _____ prego!
 4. Signora, (dare a me) _____ _____ un etto di formaggio, per favore

(se invece diamo del tu ad una persona, in questi stessi casi –consiglio, invito, richiesta, ordine– si usa invece l'**imperativo**:

Scusa, vai in stazione?

Vieni qui.

Entra, prego.

Accomodati!

Sedetevi!

Dammi un etto, per favore

2. Nelle frasi secondarie

Quando il verbo della frase principale esprime:

opinione (penso che, credo che, ritengo che, mi sembra che, mi pare che, suppongo che, ho idea che, ecc.)

dubbio o incertezza (non sapevo che, ho il dubbio che, dubito che, immagino che, non sono sicuro che, non sono convinto che)

sentimento (sono contento/felice/sorpreso che, mi dispiace che, mi stupisco che, disapprovo che, sono sicuro che, non sono sicuro che ecc.)

speranza (spero che, mi aspetto che, mi auguro che, ecc.)

timore (temo che, ho paura che, ho il terrore che ecc.)

volontà (voglio che, richiedo che, ordino che, pretendo che ecc.)

...il verbo della frase secondaria vuole il congiuntivo

	Frase principale	Frase secondaria con il congiuntivo
opinione	<i>Credo</i>	<i>che questo libro <u>sia</u> più interessante dell'altro.</i>
dubbio o incertezza	<i>Non credo</i>	<i>che <u>sia</u> giusto.</i>
opinione	<i>Mi sembra</i>	<i>che <u>stia</u> piovendo.</i>
sentimento	<i>Mi dà fastidio</i>	<i>che mi <u>trattino</u> male</i>
sentimento	<i>Mi dispiace</i>	<i>che tu <u>non sia</u> venuta.</i>
speranza	<i>Spero</i>	<i>che tu <u>sia</u> felice.</i>
volontà	<i>Voglio</i>	<i>che <u>finiate</u> il libro per domani.</i>
opinione	<i>Penso</i>	<i>che Bologna <u>sia</u> più grande della mia città.</i>

Completa le seguenti frasi coniugando i verbi al congiuntivo

1. Spero che Elisa (tornare) _____ presto a trovarci.
2. Credo che i miei genitori mi (venire) _____ a trovare per Natale.
3. Temo che tuo fratello (essere) _____ in ritardo come la volta scorsa.
4. Non credo che domani i negozi (essere) _____ aperti.

5. Ho paura che ora (andare) _____ via tutti.

6. Mi sembra che Luisa (sposarsi) _____ l'anno prossimo.

7. Mi pare che ora lui (bere) _____ troppo.

**** nota:**

Osservate queste frasi: qual è la differenza?

Penso di venire con voi alla festa.

Vs. Penso che Lucia venga con voi alla festa.

Credono di avere dimenticato il libro in aula.:

Vs. Credono che l'insegnante abbia dimenticato il libro in aula.

Mi sembra di averlo già visto.

Vs. Mi sembra che lo abbiate già visto.

.....
 Quando si usano i verbi di opinione (penso che, credo che, ritengo che, mi sembra che, mi pare che, suppongo che, ecc.), se **il soggetto della frase principale è lo stesso della frase secondaria**, non si usa il congiuntivo: si usa **DI+ INFINITO**:

RIASSUNTO:

soggetto frase principale = soggetto frase secondaria → **DI + INFINITO**

Penso di venire con voi alla festa

soggetto frase principale **DIVERSO** dal soggetto frase secondaria → **CONGIUNTIVO**

Penso che Lucia venga con voi alla festa.

(io penso =soggetto n.1, Lucia= soggetto n.2)

.....
3. quando nella frase principale c'è una forma impersonale:

esempi:

a. se nella frase principale c'è **un verbo impersonale** (SEMBRA che, PARE che, BISOGNA che, BASTA che, OCCORRE che, PUO' DARSI che, ecc.) o **forme impersonali** con il **SI** che esprimono incertezza (esempio: SI DICE che), la frase secondaria vuole il **congiuntivo**:

Sembra/pare che sia molto famoso.

Bisogna che ci vada domani.

Può darsi che ci vada subito.

Si dice che tu sia bravo.

Completa le seguenti frasi coniugando i verbi al congiuntivo

1.

Il lavoro non deve essere rifatto. Basterà che Lei (cambiare) _____ solo questa parte.

2. Bisogna che lei (capire) _____ la situazione e che (prendere) _____ una decisione.

3. Da quello che ho sentito, sembra che lui non (stare) _____ bene.

4. Si dice che quella donna (avere) _____ sofferto molto nella vita.

b. se nella frase principale c'è **È + AGGETTIVO + CHE/ COSA**, la frase secondaria vuole il **congiuntivo**

*È **necessario** che partiate subito.*

*È **meglio** che andiate via subito.*

*È **difficile** che vengano.*

*Non è **chiaro** cosa voglia da noi.*

*È **assurdo** che la biblioteca sia chiusa di sabato.*

*È **triste** che non siate venuti.*

Completa le seguenti frasi coniugando i verbi al congiuntivo:

1. È probabile che Carla ci (cercare) _____ mentre siamo fuori.

2. È possibile che Luisa ci (telefonare) _____ e non ci (trovare)

_____ in casa.

3. È difficile che i tuoi amici (dimenticarsi) _____ di venire: te l'hanno promesso.

4. Nelle frasi INTERROGATIVE INDIRECTE

In queste frasi il congiuntivo corrisponde a uno stile più formale

È utile?

Non so se sia utile / non so se è utile

Perché Emilio è partito?

Mi chiedo perché sia partito. / Mi chiedo perché è partito

Anna si è offesa?

Non capisco se si sia offesa. / Non capisco se si è offesa

Trasforma queste frasi in uno stile più formale: coniuga i verbi al congiuntivo.

1. Rebecca è tornata a casa, sua madre la vede pensierosa, mentre beve una tazza di caffè: le chiede se ha voglia di parlare di ciò che è accaduto.
2. Inutile dirlo, sto facendo una vita infernale, mi domando se ci sono possibilità che guarisca.
3. Vado a riportare l'oggetto al proprietario del negozio, che mi chiede come mai io non sono entusiasta.
4. Silver incontra il professor Matthews, al quale chiede se sente la mancanza di sua sorella.
5. Quando Massimo le ha chiesto se le piace Daniele, ha sorriso e ha detto di trovarsi molto bene con lui.

1. Rebecca è tornata a casa, sua madre la vede pensierosa, mentre beve una tazza di caffè: le chiede se _____ voglia di parlare di ciò che è accaduto.

2. Inutile dirlo, sto facendo una vita infernale, mi domando se ci _____ possibilità che guarisca.

3. Vado a riportare l'oggetto al proprietario del negozio, che mi chiede come mai io non _____ entusiasta.

4. Silver incontra il professor Matthews, al quale chiede se _____ la mancanza di sua sorella.

4. Silver incontra il professor Matthews, al quale chiede se _____ la mancanza di sua sorella.
5. Quando Massimo le ha chiesto se le _____ Daniele, ha sorriso e ha detto di trovarsi molto bene con lui.

5. dopo alcuni connettivi

Connettivi temporali:

- Andiamo a fare la spesa **prima che** chiudano i negozi
- Rimarrò con mia madre **fino a che non stia** bene.

1. La domanda va spedita alla segreteria una settimana prima che (scadere) _____ i termini della consegna.
2. Tagliate le foglie e fatele asciugare al sole, finché non (diventare) _____ da verde scuro a gialle.
3. Devi avvisarli prima che loro (partire) _____ .
4. Non si può entrare nel sito fino a che non (venire) _____ inserita una password.
5. Prima che (partire) _____ il treno devo riuscire a timbrare il biglietto.

Connettivi concessivi:

indicano che una causa non ha prodotto un certo effetto.

- Esco **benché** [*sebbene/ nonostante/ malgrado*] sia stanco.
- Non mi sembra che Luca si sia comportato correttamente, **nonostante** [*benché, quantunque, sebbene*] abbia delle spiegazioni.
- **Malgrado** [*sebbene/ nonostante/ malgrado*] ci sia il sole, fa ancora molto freddo

Nota: “**anche se**” ha lo stesso significato di **benché/ sebbene/ nonostante/ malgrado, per quanto (formale), quantunque (formale), ma non vuole il congiuntivo** = anche se ero stanco.)

1. Sebbene non (piovere) _____ più, preferisco portare la bambina a scuola in macchina.
2. Nonostante (guadagnare) _____ molto, lui e sua moglie risparmiano sempre su tutto.

3. Benché (divertirsi) _____ , a Maria non piace andare alle feste.

4. Andrò in macchina malgrado(preferire) _____ andare a piedi.

Connettivi condizionali:

l'azione principale dipende da una condizione espressa dal connettivo: *purché, a condizione che, a patto che, sempre che, laddove, nell'eventualità che, casomai, ove (formale), qualora (formale), ammesso che (formale), posto che (formale);*

- Vengo a teatro *purché* [a condizione che/ a patto che...] non costi troppo.
- *Casomai* [nel caso in cui] Giovanni non fosse in casa, cercalo in ufficio.
- Ti presterò i soldi che mi hai chiesto, *purché* [oppure a patto che, a condizione che, sempre che] tu me li restituisca presto.
- La segreteria ha stabilito che l'orario del corso di italiano verrà cambiato *a patto che* [oppure a condizione che, purché, sempre che] lo richiedano gli studenti.

1. Cerco un nuovo lavoro a condizione che (essere) _____ interessante.
2. Ti presto il dizionario a patto che tu me lo (restituire) _____ entro venerdì.
3. Le regaleremo una vacanza purché nostra figlia (superare) _____ l'ultimo esame

1. Cerco un nuovo lavoro a condizione che (essere) _____ interessante.

2. Ti presto il dizionario a patto che tu me lo (restituire) _____ entro venerdì.

3. Le regaleremo una vacanza purché nostra figlia (superare) _____ l'ultimo esame

Connettivi eccettuativi:

indicano un'eccezione o un limite: *salvo che, a meno che (non), eccetto che, tranne che.*

- *Ci andremo domenica a meno che non/ salvo che non piovano.*
- Ti aiuterò a finire il lavoro, *a meno che* [eccetto che, salvo che, tranne che] tu non voglia fare tutto da solo.
- Non mi disturbate per nessun motivo, *tranne che* [a meno che, eccetto che, salvo che] arrivi il direttore in persona.

1. Maria non si laureerà a meno che non (studiare) _____ molto.

2. Maria andrà a Roma tranne che (trovare) _____ un lavoro a Seattle.

3. Carla va al mare salvo che (fare) _____ brutto.

Connettivi finali:

indicano lo scopo, il fine per cui si compie l'azione della frase principale: *in modo che, affinché, perché, così che, sì che*

- Ti scrivo *in modo che* [*affinché/ perché*] tu *sappia* cos'è accaduto.

Faremo sì che non succeda più. (*FAR SÌ CHE = fare in modo che*)

1. Mi impegnerò personalmente affinché i lavori del recupero completo della chiesa (potere) _____ avvenire in tempi brevi.

2. Ti dico questo perché tu (capire) _____ bene la situazione.

Riflessione: congiuntivo o indicativo?

Leggi questo brano:

Anch' io ho letto il libro "Tutta da rifare" di Giorgia Wurth... sicuramente ho letto libri di portata più elevata, però credo che il messaggio che manda è molto più profondo e sottile... ragazze non siate superficiali... la trama potrebbe risultare ad alcuni anche banale, ma è il messaggio che c' è sotto che conta!!!

Ormai tutte le ragazzine pensano a rifarsi, vogliono il seno nuovo come regalo per i 18 anni... lo non ho niente contro la chirurgia, anzi penso che sia geniale per risolvere problemi "seri" di una persona... ma l'esagerazione che ormai è tanto e a mio parere TROPPO diffusa non porta mai a nulla di buono...

Tratto da: <http://www.cineblog.it/post/20702/tutta-da-rifare-il-romanzo-di-giorgia-wurth-il-book-trailer/1>

Osserva i verbi delle due frasi sottolineate:

1. penso che sia geniale per risolvere problemi "seri" di una persona

a) Qual è il verbo della frase principale? _____

b) Quale modo verbale è presente nella frase secondaria? _____

c) La frase ti sembra corretta?

Write your answer...

Answer recorder (optional) -  Voice

2. credo che il messaggio che manda è molto più profondo

a) Qual è il verbo della frase principale? _____

b) Quale modo verbale è presente nella frase secondaria? _____

c) La frase ti sembra corretta?

Write your answer...

Answer recorder (optional) -  Voice

Cosa è successo in questo testo?

Come sappiamo è obbligatorio usare il congiuntivo: a volte anche gli

Cosa è successo in questo testo?

Come sappiamo, è obbligatorio usare il congiuntivo; a volte anche gli italofoni nativi sbagliano ad usare il congiuntivo ;-) ma nelle situazioni più informali nella lingua sia parlata, sia scritta si può accettare anche l'indicativo.

Nella sezione successiva affronteremo tutti i casi che abbiamo trattato finora (verbi di opinione, forme impersonali, interrogative indirette ecc.) e vedremo come è possibile scegliere tra l'uso dell'indicativo e del congiuntivo a seconda della situazione formale o informale.

Mi chiedo se è vero

Indicativo vs. congiuntivo con l'interrogativa indiretta

Nelle frasi INTERROGATIVE INDIRETTE, il congiuntivo corrisponde a uno **stile più formale**

È utile?
utile

Non so se sia utile /non so se è

Perché Luca è partito? Mi chiedo perché sia partito/ perché è partito

Anna si è offesa? Non capisco se si sia offesa/ Non capisco se si è offesa

Leggi le seguenti frasi:

1) Cominciamo ad andare in giro per la città. Già nei primi momenti mi chiede se io sono la giornalista straniera che è venuta a visitare la città.

2) Se si intende lasciare l'auto in una città o in un posto diverso da quello dove si noleggia, bisogna dirlo subito, e chiedere se sono previsti sovrapprezzi in questo caso.

3) La ragazza chiede se c'è qualcuno che proviene da fuori e Clem solleva la sua mano.

4) Mi chiede se per caso io ho letto "I Libri di Natale" di Dickens.

Ora trasforma le frasi precedenti per adattare ad una situazione formale: stai scrivendo un articolo sulla tua esperienza come studente/ssa in scambio all'estero che sarà pubblicato sul sito della tua università.

Quindi coniuga i verbi al congiuntivo.

1) Cominciamo ad andare in giro per la città. Già nei primi momenti mi chiede se io (essere) _____ la giornalista straniera che è venuta a visitare la città.

2) Se si intende lasciare l'auto in una città o in un posto diverso da quello dove si noleggia, bisogna dirlo subito, e chiedere se (essere) _____ previsti sovrapprezzi in questo caso.

3) La ragazza chiede se c'è qualcuno che (provenire) _____ da fuori e Clem solleva la sua mano.

4) Mi chiede se per caso io (avere) _____ letto "I Libri di Natale" di Dickens.

Situazioni informali:

2) stai raccontando ai tuoi amici un film che hai visto: coniuga i verbi all'indicativo.

1) Kurt si rivolge alla ragazza e le chiede se (essere) _____ tutto vero. Mylee è costretta a confermare in lacrime.

2) Nick gli chiede se lui (volere) _____ inviare un messaggio ad Annie ed insieme fanno volare verso il cielo l'aereo, con su scritto semplicemente "Hey Annie!".

3) Cindy si rivolge a Charlie, visibilmente nervoso, e gli chiede se (andare) _____ tutto bene, ma viene però rudemente respinta.

4) Nilde raggiunge Alberto al lavoro e gli chiede se per caso (avere, lui) _____ dei problemi psicologici visto che frequenta Eloisa.

5) La ragazza chiede se (esserci) _____ qualcuno che proviene da fuori e Clem solleva la sua mano.

6) Liz Taylor, rimasta evidentemente affascinata da Depardieu, gli chiede se (essere, lui) _____ sposato. L' attore risponde di sì e la Taylor commenta "Peccato".

3) Scrivi un commento in una pagina web dedicata a **MUSICA E LIBRI**: coniuga i verbi all'indicativo.

1) Marco Mengoni risponde così ai redattori di Mondoteen che, nel corso di un' intervista, gli chiedono se (avere, lui) _____ copiato o meno il pezzo di Morgan.

2) Mi chiedo se per caso (avere, tu) _____ letto " I Libri di Natale " di Dickens

3) Qui potete anche mangiare qualcosa di messicano, asiatico o medio-orientale, ma se avete una band, potete anche provare a chiedere se i proprietari vi (fare) _____ suonare nel locale.

4) Mi chiedete se (avere, io) _____ letto i suoi libri.

Penso che è molto simpatica!

Indicativo vs. congiuntivo con i verbi di opinione/sentimento.

Con **verbi di opinione** e con **verbi che indicano sentimento** (piacere, dispiacere, sorpresa, ecc.) nelle situazioni più informali nella lingua sia parlata, sia scritta si può accettare anche l'indicativo.:

Credo che tu abbia ragione. / Credo che hai ragione.

Mi pare che abbia il morbillo/mi pare che ha il morbillo

Mi dispiace che vada via. /Mi dispiace che va via.

Situazione informale: sei tra amici, al bar, sui social oppure stai semplicemente esprimendo i tuoi sentimenti: **coniuga i verbi all'indicativo presente.**

Stai guidando la tua auto e ti arrabbi perché...

.. fai un incidente per colpa di un'altra persona:

"Basta, mi sono stancato! Se vuoi che ti (pagare, io) _____ il danno, tira fuori il modulo per la denuncia!"

...c'è troppo traffico:

"Le corriere dovrebbero andare più piano, chi guida non pensa che (esserci) _____ chi deve passare!"

Su Facebook si parla di RELAZIONI e SENTIMENTI:

1) Anch'io ho dei problemi. Se pensi che non ne (avere, io) _____, sono affari tuoi. Non ho fatto minacce.

2) Beato chi crede che la vita (essere) _____ perfetta e che ogni cosa che accade (essere) _____ prevedibile, chi crede che la compassione (esistere) _____ senza averla mai provata per qualcuno

3) Questa persona ha preso in giro tutti noi, tutti quelli che veramente soffrono e non ha pensato che (essere, loro) _____ tanti.

4) Prima di agire dovremmo fare un calcolo mentale de chiederci quali saranno le conseguenze del nostro comportamento. A meno che tu non (ritenere) _____ che per imparare occorre prima sbagliare.

5) Chiedo per l'ultima volta scusa, mi dispiace, ma ho veramente esagerato, punto. Ora non voglio che non mi (eliminare, voi) _____ dal gruppo....anzi è giusto così... quindi ora me ne vado.

Stai parlando di SPORT con i tuoi amici:

1) CALCIO - Se questi giocatori sono scarsi, allora i giocatori del Siena, dell'Empoli, del Livorno, della Reggina, che sono? Perciò penso che forse (essere) _____ lui che ha qualche problema di comunicazione con la squadra.

2) MOTO - Kallio sta andando benino, è arrivato ora dalla 250 e ha conosciuto solo questa Motogp non sa che cosa (significare) _____ guidare una moto dove sei tu che la porti e non lei che porta te.

3) WRESTLING - Vickie Guerrero non ritiene che Jeff Hardy (essere) _____ degno di battersi per il titolo massimo, e nel momento in cui glielo dice, scatena in Jeff Hardy un'incontrollabile rabbia.

4) FORMULA 1- Ho l'impressione che loro non (gradire) _____ molto il glamour della Red Bull (donne bellissime ad ogni weekend di gara, serate e non solo, in discoteca, ecc.), forse hanno paura che (distrarre, loro) _____ troppo i piloti dal loro vero lavoro.

5) FORMULA 1- Per quanto riguarda la Ferrari, sono contento che (vincere) _____ ovunque.

6) CALCIO - Tutti voi interisti scrivete queste cose perché avete paura che il Milan si (confermare) _____ anche quest'anno campione d'Europa...

Alcune persone al bar discutono di POLITICA:

1) Gli americani non sanno che cosa (essere) _____ la mafia.

2) Scusami per lo sfogo, ma mi sono accorto che certa gente non vuole che (venire) _____ scoperti i mali dell' Italia.

3) Oggi è più facile immaginare che anche il Primo Ministro Olmert possa continuare con la linea politica di Sharon, ma io non penso che (essere) _____ così chiaro qual è la linea di Sharon.

4) Ma non pensare che tu e quelli come te (essere, voi) _____ democratici e liberali. Siete l'opposto.

5) Beato chi crede che i problemi economici (essere) _____ sempre risolvibili grazie all'aiuto degli altri!

Scegli 5 frasi degli esercizi precedenti e trasformale per adattarle ad una situazione formale: coniuga i verbi al congiuntivo e, se vuoi, fai anche altri cambiamenti alla frase per renderla più formale:

Esempio di una situazione informale:

"Basta, mi sono stancato! Se vuoi che ti pago il danno, tira fuori il modulo per la denuncia!"

Esempi di situazioni formali:

"Basta, mi sono stancato! Se vuoi che ti paghi il danno, tira fuori il modulo per la denuncia!"

"La prego, per favore, sono esausto. Se vuole che io Le paghi il danno, prenda il modulo per la denuncia."

Ecc.

Write your answer...

Answer recorder (optional) -  Voice

Uso del congiuntivo: formalità vs. informalità

Scegli la risposta che spiega perché, secondo lo standard, è obbligatorio il congiuntivo nelle seguenti frasi:

1) Quando facevo in moto quel percorso, mi sembrava una cosa lunga. Adesso mi sembra che non siano passati due minuti.

- a C'è un verbo di opinione/sentimento nella frase principale
- b C'è una forma impersonale nella frase principale
- c C'è una congiunzione subordinante
- d È una frase interrogativa indiretta

2) Mi chiedo se sia ancora necessaria una manifestazione come *Miss Italia*.

- a C'è un verbo di opinione/sentimento nella frase principale
- b C'è una forma impersonale nella frase principale
- c C'è una congiunzione subordinante
- d È una frase interrogativa indiretta

3) Il regista croato ancora oggi, sebbene la Jugoslavia non esista più, vive ad Amsterdam.

- a C'è un verbo di opinione/sentimento nella frase principale
- b C'è una forma impersonale nella frase principale
- c C'è una congiunzione subordinante
- d È una frase interrogativa indiretta

4) Spero che queste vacanze servano a tutti per chiarirsi le idee, per leggersi dentro e capire cosa è meglio per ciascuno, senza i pensieri del lavoro e della vita quotidiana.

- a C'è un verbo di opinione/sentimento nella frase principale
- b C'è una forma impersonale nella frase principale
- c C'è una congiunzione subordinante
- d È una frase interrogativa indiretta

5) Ci sono troppi dischi fatti bene che sono inutili. Bisogna che i nuovi produttori capiscano che non basta fare il disco, servono melodie, cantanti, ma anche idee geniali!

- a C'è un verbo di opinione/sentimento nella frase principale
- b C'è una forma impersonale nella frase principale
- c C'è una congiunzione subordinante
- d È una frase interrogativa indiretta

6) Alexandre chiede se sia il caso di andare in albergo o da qualche altra parte.

- a C'è un verbo di opinione/sentimento nella frase principale
- b C'è una forma impersonale nella frase principale
- c C'è una congiunzione subordinante
- d È una frase interrogativa indiretta

7) Robert spiega all'amico che ha imparato a fingere perfettamente nonostante non sia affatto facile, ma Gregor gli dice che non può andare avanti in quel modo.

- a C'è un verbo di opinione/sentimento nella frase principale
- b C'è una forma impersonale nella frase principale
- c C'è una congiunzione subordinante
- d È una frase interrogativa indiretta

8) Ci domandiamo quale possa essere il ruolo, oggi, di istituzioni fondamentali per quel genere di tematica, come i musei, le biblioteche , gli archivi.

- a C'è un verbo di opinione/sentimento nella frase principale
- b C'è una forma impersonale nella frase principale
- c C'è una congiunzione subordinante
- d È una frase interrogativa indiretta

9) In pratica, saranno affidate a imprese private alcune parti dell'Amazzonia per un periodo da un minimo di 5 fino a un massimo di 40 anni, purché garantiscano di sviluppare politiche sostenibili.

- a C'è un verbo di opinione/sentimento nella frase principale

- b C'è una forma impersonale nella frase principale
- c C'è una congiunzione subordinante
- d È una frase interrogativa indiretta

10) Non credo che le mie parole rendano l'idea.

- a C'è un verbo di opinione/sentimento nella frase principale
- b C'è una forma impersonale nella frase principale
- c C'è una congiunzione subordinante
- d È una frase interrogativa indiretta

Indicativo o congiuntivo? Scegli in base al contesto informale o formale della frase:

1) L'insegnante dice: "Mi chiedo se per caso tu (avere) _____ letto "I Libri di Natale" di Dickens."

2) Un giornalista scrive: "Liz Taylor, rimasta evidentemente affascinata da Depardieu, gli chiede se (essere) _____ sposato. L'attore risponde di sì e la Taylor commenta "Peccato!"".

3) Hai fatto un incidente stradale e hai danneggiato l'auto di un'altra persona: "Basta, mi sono stancato! Se vuoi che io ti (pagare) _____ il danno, prendi il modulo per la denuncia!"

4) In un commento scritto su Facebook: "Se a una persona va bene tutto, è ovvio che non (distinguere) _____ Dj Francesco dai Pink Floyd e quindi non dà a certa musica il giusto valore."

5) All'ufficio informazioni: "Se si vuole lasciare l'auto in città o in un posto diverso da quello dove si noleggia, bisogna dirlo subito, e chiedere se (essere) _____ previsti sovrapprezzi in questo caso."

6) In un commento su Facebook: "Ma non pensare che tu e quelli come te (essere) _____ democratici e liberali. Siete l' opposto!"

7) In un articolo di politica internazionale: "La generale preoccupazione dell'Europa e degli USA si registra più che altro intorno all'avvio del progetto di arricchimento dell' uranio al 20% da parte dell' Iran. Che nessuno (desiderare) _____ gestire una situazione turbolenta e instabile è evidente."

8) In un articolo sul wrestling: "Vickie Guerrero non ritiene che Jeff Hardy (essere) _____ degno di battersi per il titolo massimo e, nel momento in cui glielo dice, scatena in lui un'incontrollabile rabbia"

9) Due persone discutono di calcio al bar: "Tutti voi tifosi dell'Inter scrivete queste cose perché avete paura che il Milan si (confermare) _____ anche quest'anno campione d'Europa..."

10) Un giornalista sportivo in tv, parlando di Formula 1: "Ho l'impressione che i

piloti non gradiscano molto il glamour della squadra Red Bull, ossia donne bellissime ad ogni weekend di gara, serate in discoteca, eccetera. Forse hanno paura che tutto questo li (distrarre) _____ troppo dal loro vero lavoro.”

