

La particella “ne” nella scrittura creativa in italiano LS. Indagine su un corpus di apprendenti greci

Katerina Florou

Dipartimento di Lingua e Letteratura Italiana, Università Nazionale e Capodistriaca di Atene

kathyflorou@ill.uoa.gr

Ioanna Tyrou

Dipartimento di Lingua e Letteratura Italiana, Università Nazionale e Capodistriaca di Atene

iotyrou@ill.uoa.gr

Abstract

L'analisi di *learner corpora* offre la possibilità di recuperare stringhe di parole ricorrenti (gruppi di parole) nei testi di apprendenti. Negli ultimi decenni, tale metodo è stato utilizzato anche per identificare le sequenze lessicali che caratterizzano diversi generi testuali. Lo scopo del presente studio è quello di analizzare le caratteristiche delle sequenze di parole ricorrenti nei testi scritti dagli studenti del Dipartimento di Lingua e Letteratura Italiana dell'Università Nazionale e Capodistriaca di Atene, al fine di individuare pattern comuni nelle produzioni, grazie ai quali trarre delle generalizzazioni. A questo scopo, sono state interrogate quattro sottosezioni di SCRICREA, un *learner corpus* che raccoglie produzioni di studenti del corso di scrittura creativa, attraverso una ricerca per *gruppo lessicale*. I risultati preliminari di questo procedimento consentono di formulare una serie di osservazioni sui gruppi lessicali estratti. In primo luogo, i gruppi lessicali recuperati e analizzati confermano che gli studenti utilizzano un numero ridotto, ma molto diffuso di *pattern* prefabbricati. In secondo luogo, è possibile ricondurre all'interferenza della L1 la principale causa dei pochi ma ripetuti errori nelle sequenze. Tali osservazioni possono avere un impatto immediato non solo sui metodi di insegnamento, ma anche sul sillabo dell'insegnamento della lingua italiana, così come già verificatosi per altre lingue.

Parole chiave: Corpora degli apprendenti, gruppi lessicali, italiano come LS

The analysis of learner corpora offers the possibility to retrieve recurring word strings (bigrams, trigrams, etc.) in learners' texts. In recent decades, this method has also been used to identify lexical bundles that characterize different text genres. The aim of this study is to analyze the characteristics of recurring word sequences in texts written by students of the Department of Italian Language and Literature at the National and Kapodistrian University of Athens, in order to identify common patterns in the productions, thereby making generalizations. For this purpose, four sub-corpora of SCRICREA, a learner corpus that collects productions from students of

the creative writing course, were queried through an n-gram search. The preliminary results of this process allow for a series of observations on the extracted lexical bundles. Firstly, the retrieved and analyzed lexical groups confirm that students use a limited but widespread number of prefabricated patterns. Secondly, the main cause of the few but repeated errors in the sequences can be attributed to L1 interference. These observations can have an immediate impact not only on teaching methods but also on the syllabus of Italian language teaching, as has already occurred for other languages.

Keywords: Learner corpora, lexical bundles, Italian language as a foreign language

Introduzione

Questa ricerca si propone di esplorare e confrontare elementi convergenti e divergenti nel discorso scritto prodotto da studenti greci che stanno apprendendo la lingua italiana. Attraverso l'analisi del linguaggio utilizzato dagli stessi studenti, si intende individuare pattern linguistici distintivi e ricorrenti. Studi nell'ambito dell'analisi degli errori [21] sono stati condotti sulle convinzioni degli insegnanti di lingua italiana in merito alle strutture linguistiche che gli studenti tendono a utilizzare o evitare. Un'analoga indagine potrebbe essere intrapresa anche in questo contesto. A tale scopo, sono stati impiegati learner corpora e strumenti di elaborazione della lingua, che hanno permesso di organizzare in modo sistematico tutti i testi inclusi nel corpus degli studenti e di studiare i gruppi lessicali che potrebbero rappresentare il linguaggio degli apprendenti.

È opportuno sottolineare che questo corpus si focalizza sulla lingua italiana come lingua straniera (LS) e su un metodo educativo moderno, la scrittura creativa che, sebbene insegnato di regola, ma solo poche volte è stato oggetto di sviluppo di corpora di apprendenti, perché ogni testo viene affrontato in modo esperienziale ed espone lo studente a una varietà di informazioni provenienti da diverse fonti per concentrarsi su diversi aspetti della materia, alla ricerca di nuove idee e conoscenze che siano preziose e importanti per lui, in modo che sia portato a creare il proprio significato. Successivamente, saranno delineati i passaggi metodologici adottati nella ricerca: dall'elaborazione iniziale del corpus, all'estrazione dei dati analitici, fino alla formulazione delle conclusioni.

Questo studio non solo contribuirà ad una migliore comprensione del processo di acquisizione e utilizzo della lingua italiana da parte degli studenti universitari greci, di tutti gli anni di studio nel dipartimento, ma potrà anche offrire spunti utili per il miglioramento delle strategie didattiche e per l'arricchimento dei contenuti del curriculum di insegnamento della lingua italiana come LS.

Stato dell'arte: nei Corpora

Durante gli ultimi due decenni, i corpora hanno consolidato la loro posizione nel campo della linguistica computazionale e dell'insegnamento delle lingue straniere e hanno formato un settore autonomo (Linguistica dei Corpora) [33][33]. Ci sono diversi tipi di corpora che si utilizzano a seconda dei bisogni di ogni ricercatore. I corpora di apprendenti sono ampie raccolte di testi orali o scritti prodotti da studenti di una lingua straniera [26]. Di conseguenza, sempre più studi ora si rivolgono ai corpora come fonte di dati primari. Mentre le metodologie basate sui corpora sono aumentate in sofisticazione, l'uso dei dati dei corpora è anche associato a una serie di problemi irrisolti [10]. I corpora di apprendenti possono essere studiati riguardo agli errori degli studenti, oppure possono essere confrontati con altri corpora di apprendenti o con corpora di

madrelingua. Tali ricerche sono state effettuate, ad esempio, per studiare la frequenza d'uso dei verbi [43][43], l'uso dei connettivi [8], il vocabolario [32], l'uso delle preposizioni [17], l'uso dei verbi modali [3].

Stato dell'arte: nei gruppi lessicali

Ci sono moltissimi studi che si concentrano sulla lingua scritta e in particolare quella argomentativa, usando il metodo dei gruppi lessicali per analizzare il testo [40]. Questo metodo, utilizzato negli ultimi venti anni, offre la possibilità di evidenziare tutte le combinazioni ripetute di parole utilizzando strumenti automatici o semiautomatici. Secondo Vilkaite [58], i ricercatori stanno concettualizzando e definendo il linguaggio formulare in modi diversi, nonostante stiano studiando lo stesso fenomeno. Molti hanno anche l'abitudine di usare i vari termini in modo intercambiabile. Infatti, il linguaggio formulare è composto da diverse categorie (ad esempio, idiomi, cliché, proverbi, collocazioni, gruppi lessicali e verbi frasali), ciascuna con le proprie caratteristiche e comportamenti. Il più vasto settore di ricerca linguistica è quello delle collocazioni. In diversi studi si sono analizzate le collocazioni con lo scopo di evidenziare la lingua dello studente della LS, come per esempio per gli apprendenti dello spagnolo [59], utilizzando l'analisi degli errori, o per l'apprendimento dell'inglese con lo studio di sovrauso o sottouso di certe collocazioni [47], oppure per l'italiano imparato dai cinesi [48]. Inoltre, ci sono studi sulle collocazioni che fanno riferimento al loro scopo educativo [24]. Nello studio di Gablasova e colleghi [24] si fa riferimento alla comparazione di diversi corpora d'apprendimento che rappresentano diversi registri, modalità e tipi testuali, attraverso le collocazioni. Mentre le collocazioni, secondo Biber e gli altri [13], sono associazioni tra parole lessicali che si riferiscono alle connessioni che esistono tra parole che co-occorrono più frequentemente del previsto per caso, i gruppi lessicali sono identificati utilizzando un approccio basato sulla frequenza: essi sono le sequenze di parole che si verificano con una maggiore frequenza in un sotto-corpus di testi appartenenti a un singolo registro [12]. In questo campo di ricerca si ritrovano studi che investigano i gruppi lessicali del discorso accademico -orale o scritto- di studenti che imparano l'inglese come lingua straniera [48]; [39]; [12] o imparano l'italiano come LS [22] o in articoli scientifici in lingua inglese nelle discipline di biologia o di linguistica [44].

Gruppi lessicali con la particella ne in italiano

Nell'ormai lontano 1974, Seuren [46] propose per la prima volta una teoria sulla particella del “ne” nel discorso come pronome clitico (le teorie precedenti si basavano su altre lingue di origine latina e giungevano a conclusioni incomplete); tuttavia, nel suo studio non vi è alcun riferimento all'uso della particella. Da allora, le grammatiche italiane menzionano il “ne” e forniscono informazioni sul suo utilizzo, che, come riassume Pirvu nella sua ricerca [41], potrebbero essere suddivise in tre principali categorie:

- a) “Ne” ha il valore di avverbio di luogo, indica il luogo (reale o figurato) da qui si viene e significa “da qui”, “da lì”.
- b) “Ne” viene poi largamente usata come pronome clitico, e in questo caso corrisponde al sintagma formato dalla preposizione di o da + un pronome personale di 3° persona, singolare e plurale maschile e femminile, nel caso di un referente umano, o un pronome dimostrativo, nel caso di un referente non umano.

- c) “Ne” è inoltre usata in molte locuzioni cristallizzate, in cui i suoi altri valori appaiono indeboliti. Uno di questi usi è con verbi intransitivi (di stato o di movimento) in cui ne, combinandosi con i clitici pseudoriflessivi mi, ti, si, ci, vi, dà luogo a forme più espressive di quelle semplici: andarsene (rispetto ad andare, andar via), dormirsene, starsene, restarsene, tornarsene, vivrsene ecc.

Per gli usi sopra citati non esiste un corrispettivo in greco. Sebbene per la maggior parte dei pronomi della lingua italiana il parlante di madrelingua greca possa trovare un equivalente, per la particella “ne” una trasposizione diretta della struttura non produrrebbe risultati. Inoltre, dall’analisi degli usi emerge che il “ne” viene impiegato all’interno di gruppi lessicali, il che potrebbe predefinire le condizioni in cui viene utilizzato. Si osserva, inoltre, che i verbi con cui appare e la sua posizione lo rendono un elemento molto frequente nel discorso italiano. Un’analisi di questi gruppi lessicali, così come si presentano nelle produzioni degli studenti di italiano come lingua straniera, sarebbe utile per determinare la difficoltà (o meno) nel loro uso, gli eventuali errori e, soprattutto, le possibili strategie per affrontarli. Il presente lavoro, pur attingendo a tutta l’esperienza di ricerca precedente, differisce perché analizza un genere di discorso scritto che è altamente raccomandato come materia accademica: la scrittura creativa.

Stato dell’arte: nella scrittura creativa

La scrittura creativa promuove lo sviluppo della lingua in termini di grammatica, vocabolario e fluidità [7]; [50], ma supporta, con approcci appropriati, lo sviluppo individuale di ogni studente [18]; [45] e il miglioramento dell’acquisizione della lingua straniera in ambiti più personali, socio-culturali e psicologici [19]. La scrittura creativa, come metodo, implica l’uso di espressioni immaginative per trasmettere sentimenti, pensieri o idee, incoraggiando gli studenti ad esplorare nuovi modi di interagire con il linguaggio. Questo approccio non solo potenzia le capacità di scrittura, ma promuove anche un miglioramento delle competenze linguistiche. Inoltre, esso facilita un’esperienza di scrittura autentica e significativa per gli apprendenti di LS [54]; [57]; [60]; [56], che può rappresentare “un mezzo per l’esplorazione e l’espressione personale” [28]. Tale approccio massimizza le opportunità di apprendimento grazie a una partecipazione attiva [25] e contribuisce a rafforzare il senso di empowerment legato all’identità linguistica e letteraria in LS [36]; [32]; [19]; [37], migliorando così l’atteggiamento verso la scrittura creativa [1]; [11]; [29]. I media tecnologici e i metodi digitali, integrati con attività di scrittura creativa, hanno dimostrato un impatto positivo sia sul pensiero creativo che sulla scrittura. Tali benefici includono il miglioramento di aspetti come l’accuratezza, la flessibilità, la fluidità e l’originalità [37]; [20], nonché un incremento della produttività scritta e delle prestazioni generali di scrittura [2]; [50]; [4]. Inoltre, questi strumenti favoriscono la motivazione, la soddisfazione, le abilità di scrittura e la capacità di identificare e correggere errori grammaticali [6]. Si osserva anche un ampliamento del vocabolario [53] e maggiori opportunità di interazione con parlanti nativi [26]. L’uso di tecnologie digitali contribuisce al miglioramento della fluidità nella lingua straniera [7] e allo sviluppo di strategie cognitive di alto livello, come l’autoregolazione emotiva e cognitiva, l’analisi dei problemi, il brainstorming e il testing delle soluzioni. Questi processi sono spesso associati a risultati superiori nelle valutazioni e a giudizi positivi [15]; [16]; [61]. Inoltre, materiali multimodali e collaborativi, utilizzati in ambienti digitali (ad esempio Google Docs, Google Forms), hanno incentivato la produzione orale e scritta creativa, sia in classe che in contesti esterni. Attività come la creazione di storie digitali, poster, video o post di blog si sono rivelate strumenti efficaci per rendere la scrittura un’esperienza di apprendimento motivante [55].

Domanda di ricerca

In questa ricerca ci concentriamo sullo studio dei gruppi lessicali appartenenti ad un elemento grammaticale osservato e riportato dagli insegnanti della lingua italiana come problematico nell'uso dagli studenti greci. In dettaglio, usiamo quadrigrammi per indagare la particella “ne” nella lingua italiana così come “ne” è stato registrato l'uso da parte degli studenti universitari greci. Lo studio dell'uso di “ne” è impegnativo dato che gli studenti di italiano di lingua madre greca trovano difficile adottare il significato di una parola per la quale non esiste un equivalente in greco. La domanda di ricerca è quanto questa difficoltà può influenzare l'uso di “ne”? Viene usato in modo corretto o in modo eccessivo (p. es. con certi verbi o gruppi di parole)? E infine, quanto l'uso di *ne* si influenza dell'obbiettivo della produzione scritta ogni volta.

Metodologia

Il Corpus di apprendenti

In questa ricerca viene utilizzato lo SCRICREA Corpus, un Learner Corpus costituito dalle produzioni scritte degli studenti universitari del corso di scrittura creativa del Dipartimento di Lingua e Letteratura Italiana dell'Università di Atene. Il Corpus ha un totale di 900.000 parole ed è la raccolta degli scritti degli ultimi 4 anni. È articolato in 8 sottosezioni che, con un punto di partenza ogni volta diverso, guida gli studenti alla produzione di un testo diverso. L'obbiettivo iniziale del Corpus era quello di raccogliere scritti e valutarli per la creatività della scrittura sulla base di una rubrica di criteri, come quantità di idee (fluidità mentale), tipologia delle idee (flessibilità mentale), rarità delle idee (originalità del pensiero), immagine (visualizzazione verbale), stile (voce) dell'autore, descrizione del personaggio (caratterizzazione) e narrazione. Le tematiche date per la produzione di otto unità sono:

UNITÀ	TEMA-OBBIETTIVO
1a unità	Continuazione della storia (il giovane gambero)
2a unità	“Inventare” un omino, di qualsiasi materiale e farlo agire, creare relazioni e casualità, a seconda del materiale di cui è fatto (come vetro, plastica, legno, gelato, ecc.) (Giacomo di Cristallo)
3a unità	Intervista immaginaria del protagonista o di un'altra persona della storia, oppure il calendario breve con elementi di testo o la registrazione dell'autobiografia dell'eroe della storia (il pozzo di Cascina Piana)
4a unità	Cambiamento del punto di vista, oppure l'interpolazione nella storia- (la bella sconosciuta)
5a unità	Un cortometraggio a tema musicale per la scrittura del racconto (basato sul tema del video) rivolto a bambini piccoli o grandi. Oppure, la scelta di un “oggetto-amico” per il protagonista della storia e scrittura di un testo per i ragazzi, o la narrazione della storia con la “voce” dell'oggetto - protagonista (un palloncino) (Way on clouds)
6a unità	Cambiamento di genere oppure la tecnica con inserimenti/aggiunzioni nella storia o la tecnica del “Cosa accadrebbe se...”(l'uomo che rubava il Colosseo)

7a unità	Tecnica del cambiamento del testo - sostituzione delle parole (cambiamento del verbo del titolo) (il re che doveva morire)
8a unità	Selezione di una canzone e scrittura di una storia che precede o segue il tema della canzone oppure la tecnica di scrittura libera

Tavola 1: I titoli delle otto unità della Scrittura Creativa

Per raggiungere l'obiettivo di questo studio e per arrivare alle risposte delle domande di ricerca sovrapposte sono state scelte 4 unità (2a, 5a, 6a e 7a) dal Corpus SCRICREA per avere un genere testuale simile da studiare, la narrativa. I testi che appartengono al corpus sono le produzioni scritte di 105 studenti. La distribuzione si vede in seguito:

Anno accademico	No di studenti/esse
2018-19	27
2019-20	40
2020-21	38
2021-22	40

Tavola 2: Numero degli studenti e le studentesse partecipanti

Quindi, in totale, usiamo 462.760 parole provenienti da 4 sottosezioni, scritti in quattro anni diversi, con la distribuzione mostrata di seguito:

Anno	2a unità	5a unità	6a unità	7a unità
2019	34.995	33.261	25.101	19.600
2020	33.853	36.893	28.681	37.878
2021	37.399	42.008	34.416	23.121
2022	20.524	20.244	17.396	17.390
TOTALE	126.771	132.406	105.594	97.989

Tavola 3: Numero di parole per le unità scelte distribuito ad ogni anno

Come si può vedere i volumi dei quattro sub-corpora sono comparabili fra di loro e il corpus in tutto ha un volume medio per trarre delle conclusioni sull'uso di gruppi lessicali. Ci sono però, studi recenti che si basano su learner corpora [48]; [49] e su corpora speciali [44] dello stesso volume. Senza dubbio, il presente non è un corpus esteso. Di sicuro, non si possono generalizzare gli eventuali risultati per tutti gli studenti di italiano come LS. Tuttavia, le caratteristiche uniche di questo corpus italiano, come il genere comune delle produzioni scritte e l'omogeneità degli studenti che vi hanno contribuito, lo rendono, nonostante le sue dimensioni modeste, uno strumento ben adatto per studiare lo sviluppo della conoscenza collocazionale nella scrittura creativa di studenti di italiano come LS. Come sottolineato da Paquot e Granger [40], "i piccoli corpora possono anche essere di considerevole valore". Indipendentemente dal valore potenziale dei piccoli corpora di studenti, le dimensioni ridotte sono forse da aspettarsi se la produzione di questi testi è fatta senza l'aiuto di grammatiche e vocabolari.

Per garantire un livello uniforme in tutti i testi ed evitare discrepanze, abbiamo utilizzato due strumenti di misurazione basati sul Quadro Comune Europeo di Riferimento (QCER). Gli studenti coinvolti nella ricerca dichiarano di possedere una competenza linguistica compresa tra i livelli B2 e C1, tuttavia risulta difficile verificarlo, poiché non hanno una consapevolezza precisa del proprio livello secondo il QCER. Ciononostante, la tematica della scrittura creativa e i relativi descrittori sono del tutto coperti dal QCER [42]. Nei casi in cui è stato necessario, abbiamo utilizzato il READ-IT [38] dell’Istituto di Linguistica Computazionale “Antonio Zampolli” e il MALT-IT2 [23] dell’Università per Stranieri di Perugia. Nonostante entrambi gli strumenti siano destinati a testi di parlanti nativi in questo caso sono stati utilizzati per misurare il livello dei testi scritti dagli studenti dell’italiano. READ-IT è stato creato per analizzare i testi misurando diversi indici, come la leggibilità, la densità del vocabolario, la lunghezza delle parole e delle frasi, ecc. Utilizzando gli stessi indici, lo strumento MALT-IT2 cerca di rilevare automaticamente il livello dei testi in italiano, operando in conformità al Quadro Comune Europeo di Riferimento per le Lingue (QCER).

Per un’analisi accurata delle strutture in esame, è necessario uno strumento per l’elaborazione della lingua dei corpora [14]. Le applicazioni degli strumenti NLP (Natural Language Processing) sono aumentate di utilizzo negli ultimi anni [30]. Esistono vari usi efficaci del NLP che assistono negli ambienti educativi, come l’elaborazione di dati empirici, l’elaborazione di corpora di apprendenti o altre analisi linguistiche necessarie per l’insegnamento delle lingue [5]. Nel caso di questo studio, ciascuno dei suddetti corpora è stato analizzato utilizzando AntConc [9]. Dopo aver categorizzato i testi in 4 singoli sub-corpora, in formato txt, abbiamo utilizzato la funzione N-grams di AntConc [48]. Abbiamo scelto i quadrigrammi per avere un quadro più completo delle collocazioni e soprattutto per il caso in cui tre parole non siano sufficienti per capire l’uso di “ne”.

I risultati della ricerca

Come previsto i gruppi lessicali che hanno avuto la frequenza più alta complessivamente e soprattutto in tre dei quattro sub-corpora sono state quelle con verbi come andare, avere, accorgere, valere e rendersi. Nella tabella osserviamo diverse combinazioni in ogni sottosezione del corpus con gli stessi verbi:

Pattern	N-grams e numero di apparizioni							
Verbo+ne	Sub-corpus 2		Sub-corpus 5		Sub-corpus 6		Sub-corpus 7	
andarsi+ne	ne andò di casa, ne andò felice, ne andava al mare, ne andato quando pioveva	17	ne andremo e mi, ne sono andati e, ne andarono e Giuseppe, ne andarono ridendo celino, ne andato ma lo	17	ne andai e gli, ne andai nel d, ne andava da solo, ne andava felice salutando, ne andrò quando non	12	ne andarono per la, ne andarono tutti via, ne andavano tutti rimaneva,	5

							ne ando nella sua	
avere+ne	ne aveva bisogno la, ne aveva bisogno quello, ne aveva molti non, ne aveva solo in, ne avevo un altro	8	ne aveva bisogno senza, ne aveva molti era, ne aveva solo una, ne aveva tanti un, ne aveva voglia era	19	ne avessi uno uguale, ne aveva più le, ne aveva un solo, ne avevano bisogno il	11	ne aveva bisogno ascoltav a, ne aveva più nessun, ne avevano abbastan za del	8
accorgersi+ne	ne accorgeva, me ne accorsò di tutto	2	ne accorge che, ne accorse, ne accorto	3	ne accurse e io, ne accorse di essere	2	ne accorger ebbe, ne accorger e, ne accorgi	3
valere+ne	ne vale la pena, ne vaveva la pena	2	-		ne vale la pena, ne vaveva la pena	2	Ne valevo la pena	1
rendere+ne	-		ne rendesse conto	1	ne rendi conto	1	ne rende proprio conto	1
rimanere+ne	-		-		ne rimani, ne rimase	2	Ne rimase perpless o	1
parlare+ne	ne parlavano	1	ne parlare, ne parlava, ne parlavano, ne parliamo	4	-		ne parla nemmen o, ne parlavan o con ammiraz ione	2

Tavola 4: Gli n-grams dello SCRICREA corpus, relativi alla particella “ne”

Per quanto riguarda il verbo “andarsi”, è un dato di fatto, come si può vedere dal suo uso, che venga insegnato agli studenti di italiano come combinato con la particella “ne” poiché, attraverso la maggior parte dei libri di testo, viene presentato come parte integrante del verbo per dare un senso del luogo di movimento, simile a “partire”. Come si vede, il verbo “avere” è accompagnato

in diversi casi da un aggettivo numerico o quantitativo e in altri dal sostantivo “bisogno”. Sono entrambi usi che, a quanto pare, sono stati consolidati dagli studenti di italiano.

Per i verbi che meno spesso richiedono il complemento introdotto da “ne”, dobbiamo osservare che la loro frequenza non è pari a quella dei verbi precedenti, anche se non mancano del tutto nel corpus. Gli studenti usano meno e solo pochissimi di loro formulano frasi con “valerne” forse perché si considerano a un livello inferiore per poter utilizzare frasi idiomatiche. Lo stesso vale per “rendersi conto” e “accorgersi + ne”, ma in questo caso anche il significato è importante: non sono verbi attesi in un testo di scrittura creativa, soprattutto quando si tratta di narrazione.

A questo punto, sarebbe opportuno considerare anche la frequenza dei rispettivi verbi in un corpus di parlanti nativi. Nel CORIS Corpus, sviluppato da un gruppo di ricerca dell’Università di Bologna, osserviamo combinazioni di verbi contenenti la particella “ne”. Per garantire un confronto valido, utilizziamo il sub-corpus Monitor 2017-20, che è il più recente [52]. Notiamo che andarsi con la particella “ne” e avere con la particella “ne” sono più frequenti rispetto ad altri pattern, come evidenziato nella tabella. Tuttavia, si riscontra una certa frequenza anche per i gruppi di parole accorgersi + ne e parlare + ne. Quest’ultimo, però, non si osserva nel corpus SCRICREA. Per quanto riguarda gli altri verbi, possiamo notare, più che altro, la loro bassa frequenza nel seguente resoconto.

Pattern	Coris Corpus Sub Corpus: Monitor 2017- 20 (n. words:14M)	%	SCRICREA Corpus (n.words: 462.760)	%
Andarsi+ ne	620	0,0044	51	0,0110
Avere+ne	1839	0,0131	46	0,0099
Accorgersi+ne	134	0,0009	10	0,0021
Valere+ne	35	0,0002	5	0,0010
Rendere+ne	82	0,0005	3	0,0006
Rimanere+ne	26	0,0001	3	0,0006
Parlare+ne	277	0,0019	7	0,0015

Tavola 5: Confronto delle apparizioni dei verbi in un Corpus di nativi

Sebbene i risultati dimostrino in tutti i sub-corpora combinazioni simili di “ne” con i verbi, non li incontriamo con la stessa frequenza. Una caratteristica comune è che i primi due verbi, cioè “andarsene” e “averne”, sono abbastanza più frequenti, ma questo non è il caso dei verbi successivi in tutti e quattro i corpora. Ne risulta, quindi, che l’obiettivo del tema assegnato influisce effettivamente sulla scelta delle collocazioni contestuali di un elemento linguistico – nel nostro caso ne – ma non condiziona l’uso dell’elemento stesso quando questo non fa parte di una struttura verbale.

Gli errori osservati nei quadrigrammi descritti nella tabella sono di bassa frequenza ma influenzano la prestazione scritta dello studente. Ad esempio, il trasferimento fonologico “se ne andato” (invece di “se n’è andato”) è un errore che compare 3 volte ed è problematico, così come l’uso ridondante di “ne” quando “bisogno” è seguito da complemento indiretto (p.e. “ne hanno bisogno per”) o, erroneamente, complemento diretto (p.e. “ne avevo bisogno il”). Per quanto riguarda “parlare”, d’altra parte, di solito crea problemi soprattutto con la preposizione

complementare “di”. In tutti i casi apparsi nel corpus, però, non c’è né sovrauso della preposizione “di”, dal momento che appare “ne”, né uso errato di qualche altra preposizione.

Conclusioni

Osservando i primi risultati, arriviamo a varie conclusioni. In primo luogo, quasi come previsto, gruppi lessicali specifici corrispondono alla tematica dei testi che li contengono e soprattutto al genere testuale [51] dal momento che si tratta di verbi ed espressioni che si usano soprattutto nella narrativa. In seguito, possiamo vedere le “ossessioni” degli studenti in modelli linguistici ripetitivi. Una possibile spiegazione di questa osservazione è la fiducia che provano quando usano le stesse frasi [35], che sono probabilmente quelle che conoscono meglio. Tra l’altro, le sequenze prefabbricate sono in genere limitate a quelle combinazioni di parole che sono relativamente fisse nella forma e non composizionali nel significato [12]. La terza osservazione è relativa alla lingua madre e all’interferenza osservata nei pochi, ma ripetuti, errori degli studenti [40], come per esempio quando usano la preposizione “per” dopo il gruppo lessicale “avere bisogno” traducendo dal greco e dalla frase equivalente che si accompagna dalla preposizione “για”.

L’analisi di gruppi lessicali rivela una tecnica in grado di distinguere modelli linguistici su uno o più elementi (“ne” nel nostro caso), partendo da quelli comuni a tutti i sub-corpora e passando a quelli che contraddistinguono la varietà linguistica di uno studente universitario da un altro [1]. Questa conclusione è utilizzabile nell’insegnamento, che è l’obiettivo finale di tutta la nostra ricerca. La sua applicabilità si vede non solo nel metodo di insegnamento, ma anche nel curriculum dell’insegnamento dell’italiano, così come è già stato fatto con altre lingue [34].

Più nello specifico, i modelli osservati in questo studio potrebbero essere oggetto di ricerca da parte degli studenti stessi, poiché la ricerca dei corpora degli studenti universitari è un modo per migliorare e perfezionare la lingua scritta in italiano come LS. Ma allo stesso tempo, la ricerca sui gruppi lessicali in un corpus di parlanti madrelingua (dello stesso genere testuale) può essere introdotta come strategia di ricerca in classe con lo scopo di individuare le corrispondenze nelle due lingue e far realizzare allo studente la fase dell’interlingua in cui si trova. In questo modo, non solo si usufruisce dell’utilizzo delle applicazioni digitali, ma si può imparare ad usare la ricerca scientifica per l’apprendimento delle lingue [5]. In modo più specifico, a seguito di una tale ricerca, potrebbero essere proposte microattività didattiche che coinvolgono strumenti automatici (tipo NLP), i quali, tramite la ricerca, aiuterebbero gli studenti a individuare gruppi lessicali nei corpora di parlanti nativi o nei propri testi. La funzione delle concordanze è molto utile sia come esercizio che come strumento di riferimento. D’altra parte, tenendo presente i gruppi lessicali che creano difficoltà o che gli studenti tendono a evitare, l’insegnante può concentrarsi su testi autentici che possano fungere da spunto per la scrittura creativa e che contengano tali strutture.

Concludendo, si può dire che nell’apprendimento delle lingue e, in particolare, nell’apprendimento dell’italiano come LS, c’è bisogno di studi che indaghino fino a che punto gli studenti riescano a controllare le funzioni discorsive dei gruppi lessicali attraverso l’acquisizione naturale, così come studi che valutino l’efficacia degli approcci pedagogici per l’insegnamento esplicito dei gruppi lessicali [12].

Ringraziamenti

Vorremmo ringraziare il docente Kostas Mylonas e i dottorandi Fabio Zanda e Andrea Malorgio per il loro aiuto e le loro osservazioni durante la ricerca. Un ringraziamento speciale ai nostri studenti (e futuri insegnanti dell'italiano) per il loro contributo indispensabile.

Bibliografia

- [1] Aarts, J., & Granger, S. (2014). Tag sequences in learner corpora: A key to interlanguage grammar and discourse. In *Learner English on computer*, pp. 132-141. Routledge. Available from https://scholar.google.com/scholar?hl=en&as_sdt=0%2C5&q=Aarts+granger+2014&btnG=
- [2] Abd Alwahab Ibrahim Abd Alwahab Gena (2022). The Effects of Using RAFT Strategy on Developing EFL Writing Skill among Secondary School Students (thesis).
- [3] Abdulqader, R., Saifadin, A., Sharif, I., & Abdulla, S. (2023). Applying Techniques in Creative Writing by EFL Students. 10.21271/zjhs.27.4.24. *Zanco Journal of Human Sciences*, pp.347-354.
- [4] Aijmer, K. (2002). Modality in advanced Swedish learners' written interlanguage. In S. Granger, J. Hung, & S. Petch-Tyson (Eds.), *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching*, pp. 55-76. John Benjamins Publishing Company, Amsterdam/ Philadelphia. <https://doi.org/10.1075/llt.6.07aij>
- [5] Al-Jarf, R. (2015). Discourse and Creativity Issues in EFL Creative Writing on Facebook. *Int. J. Signs Semiot. Syst.*, 4, 54-81.
- [6] Alhawiti, K. M. (2024). *Natural language processing and its use in education.* "International Journal of Advanced Computer Science and Applications", 5(12). <https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=c0e3dd3bcb8990a956db697bd7f5cc26c5390eb0#page=84>
- [7] Alkhaldi, Ali. (2023). The Impact of Technology on Students' Creative Writing: A Case Study in Jordan. *Theory and Practice in Language Studies*, 13, 586-592. 10.17507/tpls.1303.06.
- [8] Alkhaldi, A. A., Ozdemir, E. A, & Alhasan, R. F. (2022). ESP creative writing from engineering students' perspectives: A case study. *International Journal of English Language and Literature Studies*, 11(3), 136-146.
- [9] Altenberg, B., & Tapper, M. (1998). *The use of adverbial connectors in advanced Swedish learners'.* In S. Granger (Ed.), *Learner English on Computer*, pp. 80-93. Longman, London and New York.
- [10] Anthony, L. (2022). AntConc (Version 4.1.4) [Computer Software]. Tokyo, Japan: Waseda University. Available from <https://www.laurenceanthony.net/software>

- [11] Arppe, A., Gilquin, G., Glynn, D., Hilpert, M., & Zeschel, A. (2010). Cognitive Corpus Linguistics: Five points of debate on current theory and methodology. In *Corpora* 5. pp. 1-27. <https://doi.org/10.3366/cor.2010.0001>
- [12] Athanasopoulou, A., & Douzina, M. (2022). Redefining the students' contact with literature via creative writing and reading. *Innovation in Language Learning*, 1–5. <https://conference.pixel-online.net>
- [13] Biber, D., & Barbieri, F. (2007). Lexical bundles in university spoken and written registers. *English for specific purposes*, 26(3), 263-286.
- [14] Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (2000). *Longman grammar of spoken and written English*.
- [15] Chambers, A. (2005). Integrating Corpus Consultation in Language Studies. *Language Learning and Technology*, 9(2), 11-125.
- [16] Dawilai, S., Kamyod, C., & Prasad, R. (2019). Effectiveness Comparison of the Traditional Problem-Based Learning and the Proposed Problem-Based Blended Learning in Creative Writing: A Case Study in Thailand. *Wireless Personal Communications*, 118, 1853 - 1867.
- [17] Dawilai, S., Kamyod, C., & Champakaew, W. (2018). Proposed Problem-based Blended Learning in Creative Writing: Self-Regulated Learning in EFL Learners. *International Journal of Applied Engineering Research*, ISSN 0973-4562, 13 (7), 4834-4841.
- [18] Diez Bedmar, M., & Casas Pedrosa, V. (2006). The use of prepositions by Spanish learners of English at University level: Main problems. In *7th Conference on Teaching and Language Corpora (TaLC) Proceedings*, pp. 42-44, Paris.
- [19] Durdas, A., Harbuza, T., Radchenko, Y., & Starosta, H. (2022). Teaching Foreign Language Efficiently: The Role of Creative Writing. *Continuing Professional Education: Theory and Practice*. 3, 33-38. 10.28925/1609-8595.2022.3.4.
- [20] Efverlund, Y., & Wachtmeister, W. (2021). The Effectiveness of using L2 Creative Writing in the Classroom to support Autonomous Learning and Motivation. *Independent Project with Specialization in English Studies and Education*, <https://www.diva-portal.org/smash/get/diva2:1526788/FULLTEXT02>
- [21] El-sherbeny, A. (2019). Using CoRT Based Program to Improve Creative writing Skills of EFL Secondary Stage Students. 63-41. 108. *مجلة كلية التربية بالمنصورة*. 10.21608/maed.2019.132095.
- [22] Florou, K. (2019): Learner Translator Corpus: Italogreco or another way to confirm teachers' intuitions. *Journal of Education and Learning*; Vol. 8, No. 5; 2019 <https://doi.org/10.5539/jel.v8n5p75>
- [23] Florou, K. (2023). Teaching the usage of the verb complements by analyzing our students written productions? A study in the lexical bundles of a learner corpus of the Italian language. In *Conference Proceedings. The Future of Education 2023*.
- [24] Forti, L., Milani, A., Piersanti, L., Santarelli, F., Santucci, V., & Spina, S. (2019). Measuring text complexity for Italian as a second language learning purposes. In

Proceedings of the Fourteenth Workshop on Innovative Use of NLP for Building Educational Applications, 360-368, Florence, Italy, August. Association for Computational Linguistics.

- [25] Gablasova, D., Brezina, V., & McEnery, T. (2017). Collocations in corpus-based language learning research: Identifying, comparing, and interpreting the evidence. *Language learning*, 67(S1), 155-179.
- [26] Giacomini, L. (2015). Using “Storybird” in Young Learners’ Creative Writing Class. *English Teaching Forum*, 35-37.
- [27] Granger, S. (2004). Computer Learner Corpus Research: Current Status and Future Prospects. In M. Conn, & T. Upton (Eds.), *Applied Corpus Linguistics: A Multidimensional Perspective*, 123-145. Amsterdam and Atlanta: Rodopi. https://doi.org/10.1163/9789004333772_008
- [28] Granger, S. (1998). *The computer learner corpus: A versatile new source of data for SLA research*. In S. Granger (Ed.), *Learner English on Computer*, 3-17. Longman, London and New York.
- [29] Hanauer, D. I. (2014). Appreciating the beauty of second language poetry writing. In D. Disney (Ed.), *Exploring second language creative writing: Beyond Babel*, 11-22. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- [30] Kosmopoulou, A.M. (2020). L’écriture creative: un outil didactique et pedagogique alternative au service de l’éducation interculturelle en classe de FLE. Diploma thesis, Hellenic Open University.
- [31] Lauriola, I., Lavelli, A., & Aiolfi, F. (2022). An introduction to deep learning in natural language processing: Models, techniques, and tools. *Neurocomputing*, 470, 443-456.
- [32] Lenko-Szymanska, A. (2005). The role of L1 influence and L2 instruction in the choice of rhetorical strategies by EFL learners. In B. Lewandowska-Tomaszczyk (Ed.), *PALC’2005. Practical Applications in Language Corpora*. Peter Lang.
- [33] Maloney, I. (2022). The Impact of Creative Writing on L2 Writer Confidence. *International Journal of TESOL Studies*. 10.46451/ijts.2022.04.03.
- [34] McEnery, T., & Hardie, A. (2012). *Corpus Linguistics: Method, theory and practice*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511981395>
- [35] Meunier, F. (2012). Formulaic language and language teaching. *Annual Review of Applied Linguistics*, 32, 111–129. doi: 10.1017/S0267190512000128. Available from <https://doi.org/10.1017/S0267190512000128>
- [36] Nesselhauf, N. (2005). *Collocations in a Learner Corpus*. Amsterdam, the Netherlands: John Benjamins. doi:10.1075/scl.14
- [37] Nondabula, N. & Nomlomo V. (2023). Learners’ experiences of creative writing in English First Additional Language: Pedagogical implications. *Journal for Language Teaching*, 57(1). <https://doi.org/10.56285/jltVol57iss1a5884>

- [38] Nurmieva R.R. & Soboleva N.P. (2018). Using the technology of creative writing to develop students' creative thinking in teaching English. *The Journal of Social Sciences Research*, 1, 458-463.
- [39] Orletta Dell', F., Montemagni, S., & Venturi, G. (2011). READ-IT: Assessing readability of Italian texts with a view to text simplification. In *Proceedings of the second workshop on speech and language processing for assistive technologies*, 73-83.
- [40] Pan, F., Reppen, R., & Biber, D. (2016). Comparing patterns of L1 versus L2 English academic professionals: Lexical bundles in Telecommunications research journals. *Journal of English for Academic Purposes*, 21, 60-71.
- [41] Paquot, M., & Granger, S. (2012). Formulaic language in learner corpora. *Annual Review of Applied Linguistics*, 32, 130-149. doi:10.1017/S0267190512000098. Available from https://scholar.google.com/scholar?hl=en&as_sdt=0%2C5&q=paquot+granger+2012&btnG=
- [42] Pîrvu, E. (2014). La particella italiana ne e le modalità della sua traduzione in romeno.
- [43] Redazione, L. (2023). QUADRO COMUNE EUROPEO DI RIFERIMENTO PER LE LINGUE: APPRENDIMENTO, INSEGNAMENTO, VALUTAZIONE. VOLUME COMPLEMENTARE: Language Policy Programme. Education Policy Division. Education Department. Council of Europe. *Italiano LinguaDue*, 15(2).
- [44] Ringdom, H. (1998). Vocabulary frequencies in advanced learner English: A cross-linguistic approach. In S. Granger (Ed.), *Learner English on Computer*, 41-52, Longman, London and New York.
- [45] Samraj, B. (2024). Disciplinary differences in lexical bundles use: A cautionary tale from methodological variations. *Journal of English for Academic Purposes*, 101399.
- [46] Seliem, S. & Mohamed, M. & Ali, R. (2020). The Effect of RAFT Strategy on Developing EFL Creative Writing Skills for the Third Year Governmental Language Preparatory School Students. *مجلة واجتماعية تربوية دراسات مجلة*, 6311-355. doi 10.21608/jsu.2020.160538
- [47] Seuren, P. A. (1974). Pronomi clitics in italiano. In *Sesto Congresso Internazionale di Studi* (pp. 309-327). Bulzoni.
- [48] Shih, H. H. (2000). Collocation deficiency in a learner corpus of English: From an overuse perspective. In *Proceedings of the 14th Pacific Asia Conference on Language, information and Computation*, 281-288.
- [49] Shin, Y. K., & Won, D. O. (2024). To what extent do L2 learners produce genre-appropriate language? A comparative analysis of lexical bundles in argumentative essays and speeches. *Journal of English for Academic Purposes*, 69, 101389.
- [50] Singh, N., Bernal, G., Savchenko, D., & Glassman, E. L. (2022). Where to Hide a Stolen Elephant: Leaps in Creative Writing with Multimodal Machine Intelligence. *ACM Transactions on Computer-Human Interaction*. Siyanova-Chanturia, A. (2015). On the

- ‘holistic’ nature of formulaic language. *Corpus Linguistics and Linguistic Theory*, 11(2), 285-301.
- [51] Smith, C. A. (2014). Creative writing as an important tool in second language acquisition and Practice. *The Journal of Literature in Language Teaching*, 2(1), 11-18.
- [52] Spina, S., & Tanganelli, E. (2012). Collocations as an index for distinguishing text genres. *Corpus*, 11, 73-89.
- [53] Tamburini, F. (2022). I corpora del FICLIT, Università di Bologna: CORIS/CODIS, BoLC e DiaCORIS. In *Corpora e Studi Linguistici. Atti del LIV Congresso Internazionale di Studi della Società di Linguistica Italiana* (pp. 189-197). Società di Linguistica Italiana.
- [54] Tenaja, J.F.M., & Oco, R.M. (2023). Utilization of Social Media and Learners’ Creative Writing Skills. *Asian Journal of Education and Social Studies*, 44 (3), 1-14, AJESS.100435 ISSN: 2581-6268.
- [55] Tyrou, I. (2023). The experience of writing creative texts in the acquisition of foreign language. In the *International Conference Advances in second-foreign language acquisition (ASeFoLA)*, 9-10 June, Nicosia, Cyprus.
- [56] Tyrou, I. & Mylonas, K. (2023). Teaching Italian as a foreign language and elements of its culture using digital media. Poster presentation. In the *International Conference Advances in second-foreign language acquisition (ASeFoLA)*, 9-10 June, Nicosia, Cyprus.
- [57] Tyrou, I. (2022). La Scrittura Creativa nell’Insegnamento delle Lingue Straniere: Suggestimenti ed Esempi. Il Caso Universitario [Creative Writing in Foreign Language Teaching: Suggestions and Examples. The University Case.] In *5th International Conference on Creative Writing*, Palermo, 1866-1883.
- [58] Urlaub, P. (2011). Developing literary reading skills through creative writing in German as a second language. *Die Unterrichtspraxis/Teaching German*, 44 (2), 98-105. <https://doi.org/10.1111/j.1756-1221.2011.00099.x>
- [59] Vilkaitė, L. (2016). Formulaic language is not all the same: Comparing the frequency of idiomatic phrases, collocations, lexical bundles, and phrasal verbs. *Taikomoji kalbotyra*, (8), 28-54.
- [60] Wanner, L., Ramos, M. A., Vincze, O., Nazar, R., Ferraro, G., Mosqueira, E., & Prieto, S. (2013). Annotation of collocations in a learner corpus for building a learning environment. *Twenty years of learner corpus research. Looking back, moving ahead*, 493-503.
- [61] Youssif, M. (2021). Using Literature Circles-based Program for Developing EFL Creative Short Story Writing Skills among Secondary Stage Students, 32, 120-79, *مجلة كلية التربية*. <https://doi.org/10.21608/jfeb.2021.233065>
- [62] Αλεξόπουλος, Π., Γάκης, Π., & Ξεστέρου, Μ. (2022). Ψηφιακές εφαρμογές και παιχνίδια δημιουργικής γραφής για τη διδασκαλία της λογοτεχνίας και της ελληνικής ως δεύτερης ξένης γλώσσας. [Creative writing digital applications and games for teaching literature and Greek as a second foreign language]. In the *7th Panhellenic Scientific Conference “Integration and Use of ICT in the educational process”*, Patras, 693-704.